CONTROLLING TRANSPOSITION IN THE MALE GAMETOPHYTE OF TRANSGENIC PLANTS

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

Michael Richard Roberts BSc. (Durham)

Department of Botany

University of Leicester

March 1992

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Controlling Transposition in The Male Gametophyte of Transgenic Plants.

Thesis submitted for the degree of Doctor of Philosophy.
Michael Richard Roberts (BSc.)

Abstract

Investigations were carried out to determine the feasibility of a transposon tagging experiment in flax, *Linum usitatissimum*. The excision of the maize transposable element *Activator* (*Ac*) from the genome of transgenic flax callus was demonstrated, whilst a *Dissociation* element (*Ds*) was found to be stable. However, reintegration of excised *Ac* elements was not detected, and this barrier to gene tagging led to an examination of procedures which might improve the general applicability of transposon tagging.

A recombinant *Ac* transposase gene was constructed in order to achieve a high germinal transposition frequency in transgenic plants; this feature is an essential component of an efficient transposon tagging strategy. The *Ac* construct was produced by fusing the promoter of an anther-specific gene to the transposase coding region. The anther-specific gene, APG, was cloned from *Arabidopsis thaliana*, following the identification of four putative microspore-specific mRNAs from *Brassica napus*. Of these mRNAs, one, termed I3, was analysed in detail and found to encode a novel oleosin protein, and was apparently confined to developing pollen. The I3 cDNA was used as a molecular probe to clone the APG gene, which encodes a proline-rich protein of unknown function. A small gene family encoding proteins with high sequence similarity to the APG protein was identified in *B. napus* via the isolation of three cDNAs termed CEX1, CEX2 and CEX6.

Promoter fragments of the APG gene were demonstrated to drive expression of a β-glucuronidase reporter gene in the male gametophyte, tapetum, stomium and anther wall of *Nicotiana tabacum* and *Arabidopsis* during the microspore development stage of gametogenesis. The restriction of transposition to these cells would permit the production of a seed population containing a wide range of unique transposon inserts which would be stable in during vegetative growth. Such applications of the APG/*Ac* fusion are discussed.

Acknowledgements

Thanks are due to my supervisors, Dr. John Draper, Dr. Rod Scott and Dr. Amar Kumar, for coming up with the ideas which became the results presented in this thesis. Thanks also go to Dr. Gary Foster for not allowing me to condemn the APG clones to the back of a freezer for eternity. These people are also thanked for critical reading of this manuscript. I thank Rod Scott and Fran Robson for the gift of the cDNA clones G2, #14, #17 and I3, and Rob Blundell and Sue Robinson for their contributions to the sequencing of the APG clones.

Special thanks to all my friends who have guided (or misguided) me through the last three years, particularly Dawn Worrall and Simon Warner for their constant good company.

I acknowledge the SERC and the Scottish Crop Research Institute for financial support.


UNIVERSITY LIBRARY LEICESTER

Publications

Some of the work presented in this thesis has been accepted for publication in the following papers:

M.R. Roberts, A. Kumar, R. Scott, J. Draper (1990) Excision of the maize transposable element *Ac* in flax callus. Plant Cell Reports 9: 406-409.

M.R. Roberts, F. Robson, G.D. Foster, J. Draper, R.J. Scott (1991) A *Brassica napus* mRNA expressed specifically in developing microspores. Plant Molecular Biology 17: 295-299

M.R. Roberts, G.D. Foster, R.P. Blundell, S.W. Robinson, A. Kumar, J. Draper, R. Scott (1992) Gametophytic and sporophytic expression of an anther-specific *Arabidopsis thaliana* gene. The Plant Journal (in press)

Table of Contents

CHAPTER 1

INTRODUCTION

1.1 The Scope of this Thesis

The work described in this thesis concerns the development of transposon tagging strategies, both in terms of specific gene tagging experiments and generally applicable improvements which might be made to existing schemes. An initial aim of the project was to characterize the behaviour of the *Z. mays* (maize) transposable elements *Activator* (*Ac*) and *Dissociation* (*Ds*) in transgenic *Linum usitatissimum* (flax); that is, to ask the questions "does *Ac* excise and reinsert as demonstrated in other plants," and "is *Ds* stable?" For reasons discussed below, *L. usitatissimum* represents a species in which a transposon tagging experiment may prove effective. Transposon tagging has successfully been used with other species as an approach to cloning genes of interest for which the gene product is unknown. The elements *Ac* and *Ds* were chosen because of their ability to transpose in transgenic plants in a manner similar to that in *Z. mays*, their natural host. The degree of success of these initial experiments would influence a decision whether to proceed with a tagging experiment in *L. usitatissimum*, or, whether to concentrate more on an investigation of methods which may be used to increase the efficiency and applicability of *Ac* and *Ds* to generalized transposon tagging strategies, since several drawbacks of existing schemes have been recognized (see section 1.6).

These drawbacks include the lack of clonal mutant characters in plants containing transposed elements because of the prevalence of somatic activity over germinal activity. This means that most first generation plants exhibiting mutant phenotypes cannot easily be used as the source of a tagged gene since sectors containing untransposed elements, and possibly sectors containing different reinserted copies of the transposed element, will exist in addition to the desired mutated cell line. Only when transposition occurs in cells which ultimately form gametes will mutant lines be clonal and not chimaeric, as long as the transposed element does not move a second time. As will be seen, the major work in this thesis concentrates on efforts to clone a microspore-specific promoter which could be used to enhance the frequency of germinal transposition and prevent somatic activity. For this purpose, anther-specific cDNA clones from *Brassica napus* (oilseed rape) were received from Dr. Rod Scott (University of Leicester) and further characterized by northern analysis in order to identify putative microspore-specific clones. These in turn were used as probes to clone homologous *Arabidopsis* genes. The emphasis during this work was on the isolation of a promoter suitable for driving transposition in developing microspores, rather than a detailed investigation of the molecular biology of gametophyte gene expression. However, these two areas are not mutually exclusive, and it was hoped that some light would be thrown onto the control of gene expression in microspores as a result of the necessary analysis of any cloned promoter, especially via the use of reporter gene fusions. The ultimate goal of this thesis project was to enable expression of *Ac* transposase using a microspore-specific promoter, and to determine its ability to drive germinal transposition of a *Ds* element. Such controlled germinal transposition is expected to

generate clonal insertional mutant lines useful for gene isolation experiments.

1.2 *L. usitatissimum* as a Potential Host for Transposon Tagging

L. usitatissimum has a small but increasingly significant commercial value, being an oil seed crop in temperate countries and an important source of natural fibres, which can be woven to form linen. Furthermore, *L. usitatissimum* has several attributes which make it amenable to a gene tagging programme. Firstly, transformation of this species via *Agrobacterium* has been described (Basiran *et al.*, 1987) and is a relatively simple procedure. The availability of haploid plants from twinning embryos (Rajhathy, 1976) and the small genome size of *L. usitatissimum* (n = 0.7 pg) ease constraints on both genetic and molecular analyses. The use of haploid plants would prevent masking of mutated alleles by the expression of co-dominant genes on daughter chromosomes, whilst the small genome size makes the construction of genomic libraries and gene cloning more straightforward than in most other crop species and also increases the chance of tagging a functional gene.

An attractive target for gene cloning experiments is the family of genes conferring resistance to the rust fungus *Melampsora lini*, which is a pathogen of *L. usitatissimum*. Around 30 separate alleles for resistance are known and have been thoroughly described by classical genetic methods (for a review see Ellis *et al.*, 1988). Indeed, it was the investigation of the interactions between *L. usitatissimum* and *M. lini* that originally led Flor to determine the well known gene-for-gene relationship between resistance and avirulence (Flor, 1956). However, a classical gene cloning approach to isolate such genes is not feasible, since the biochemistry of the interaction is not understood, and no function can be attributed to the gene products. Only when a mutation gives rise to a selectable or screenable phenotypic affect can the gene be identified. Furthermore, it is not known whether the resistance genes are inducible by pathogen attack, or if their products are present continually. The more likely situation is the latter, since different pathovars are presumably recognized by specific proteins presented by the plant in order for a resistance response to take place. Hence cloning of cDNAs differentially expressed upon infection by *Melampsora* would not be likely to facilitate the cloning of a resistance gene. However, any one of these alleles might be inactivated by the insertion of a transposable element, leading to the production of a screenable phenotype - sensitivity to infection. The associated resistance gene could then be cloned via its proximity to transposon DNA sequences. With a view to initiating a tagging programme in *L. usitatissimum*, this thesis describes the investigation of the behaviour of the *Z. mays* elements *Ac* and *Ds* in transgenic *L. usitatissimum* tissues.

1.3 Transposable Genetic Elements

It has become increasingly apparent over the last two decades that a large proportion of the genome classified as middle repetitive DNA consists of transposable elements

(Smyth, 1991). Transposable genetic elements are defined DNA sequences, ranging from hundreds to tens of thousands of base pairs in length, frequently in multiple copies, which are capable of replication or excision from their position in the genome and subsequent reinsertion at a random site elsewhere (Freeling, 1984; Doring and Starlinger, 1986). It is now recognized that this mobile DNA is responsible for many of the various kinds of genome changes which are observed, from simple insertion and deletion mutations to gross chromosome rearrangements (Smyth, 1991). Transposable elements exist in many forms, but are recognizable by certain conserved characteristics (Freeling, 1984). The most important of these is the presence of short inverted repeat sequences at the termini of the element. Frequently, subterminal long imperfect direct repeat or inverted repeat sequences also occur, bordering the internal regions of the element, and these account for many of the tandem repeat sequences detected in genomes by Cot analysis (Smyth, 1991). A final characteristic of transposons is that their insertion invariably results in the production of short direct duplications of the target DNA, which forms a "transposon footprint" if this is not repaired upon re-excision (Doring and Starlinger, 1986).

Two major groups exist under the general heading of transposable elements, these being the retroelements and the transposons. The retroelements transpose via RNA intermediates, whilst the transposons transpose via DNA intermediate forms. Transposition of the DNA-intermediate elements is either by a replicative or non-replicative process, depending on the element in question. Non-replicative transposition, common in eukaryotes, involves excision and reinsertion of the element, whereas replicative transposition, the norm in prokaryotes, occurs when a DNA copy of the element is synthesized and inserts in a new location in the genome. Active transposable elements encode one or more proteins which mediate the site-specific recombination required for their transposition, including a "transposase" enzyme, able to recognize and interact with the terminal and subterminal repeat sequences (e.g. see Saedler and Nevers, 1985; Kunze and Starlinger, 1989). Many comparisons have been drawn between the biology of transposable elements and viruses, particularly retroelements and retroviruses. The major difference between *Ty1-copia* group retroelements from yeast and *Drosophila*, homologues of which are found in many other species, including plants, is the presence of an *env* gene in the viruses which enables cell-cell transport. The presence of retrotransposons in primitive eukaryotes and their simpler genetic constitution argues that they may have been the progenitors of retroviruses. Alternatively, the retrotransposons may have evolved from retroviruses by the loss of their *env* gene. (Reviewed by Flavell *et al.*, 1992.)

Transposition of mobile DNA to new locations in the genome can produce insertion mutations, and unstable phenotypic effects are often observed when non-transcribed or non-translated regions of genes are interupted. Indeed, this is classically how many transposable elements were first identified. Reversion rates of mutations caused by

transposon insertions are frequently much higher than those caused by other mechanisms, as most transposons have the ability to excise and restore gene function. Such insertional mutagenesis has proved to be a useful phenomenon since the advent of molecular biology, with the development of a method of gene isolation called "transposon tagging". This is a technique whereby cloned transposons are used as molecular probes to identify mutant alleles which can in turn be used to identify the wild type allele.

1.4 Plant Transposable Elements

The first transposable elements to be identified were found in *Z. mays* by Barbara McClintock and were termed "*Activator* " and "*Dissociation*," (McClintock, 1946). During genetic studies of *Z. mays* kernel pigmentation, she correlated chromosome breakage with developmentally regulated factors which she termed "controlling elements." A stable element, *Dissociation* (*Ds*), could cause chromosome breakage only when activated by the product of a separate genetic locus, *Activator* (*Ac*). In 1948 she reported that *Ds* was capable of transposition (McClintock, 1948). Work following the discovery of insertion sequences in bacteria (Jordan *et al.*, 1969) led to the discovery of transposable genetic elements in examples of all classes of organism. The most well known of these include the drug resistance-carrying transposons (Tn) of *Escherichia coli*, the *Ty* element of *Saccharomyces cerevisiae*, and the *copia* and *P*-elements of *Drosophila melanogaster.* Many plant transposable elements have now been identified, mainly in *Z. mays* and *Antirrhinum majus* (snapdragon) though they have been shown to be present in a wide range of species.

Plant transposable elements appear to fall into a number of groups which can be recognized by similarities in gene organization in the case of the retroelements, or by conservation of terminal repeat sequences. These groups appear to have each derived from different common ancestors and contain members from a surprisingly diverse range of species. One group (see Hehl *et al.*, 1991) includes *Ac* and *Ds* from *Z. mays* and *Tam3* from *A. majus* (Sommer *et al.*, 1985), whilst another (see Doring and Starlinger, 1986) includes *En/Spm* from *Z. mays* (Pereira *et al.*, 1985), *Tam1* and *Tam2* from *A. majus* (Bonas *et al.*, 1984; Upadhyaya *et al.*, 1985), *Tgm1* from *Glycine max* (soybean) (Goldberg *et al.*, 1983) and *Pis1* from *Pisum sativum* (pea) (Shirsat *et al.*, 1988).

One feature unique to plant transposable elements is that most exist as families containing autonomous and non-autonomous members (Doring and Starlinger, 1986). The autonomous element of a family is fully functional and able to transpose independently, whilst the non-autonomous members are usually elements possessing the terminal repeats and *cis* -acting sequences required for transposition, but are incapable of producing the transposase enzyme(s) and/or other gene products necessary for transposition. Non-autonomous elements are active only in the presence of an active autonomous relative which can supply transposase functions in *trans.* Some families of

transposons exist for which there are no known functional members. For example, Voytas *et al.* (1990) have discovered a number of different retroelement groups present in many *Arabidopsis thaliana* ecotypes, which appear to have entered the *A. thaliana* genome before global dispersal of the species, and which have failed to undergo further transposition.

One of the best studied transposon families is the *Ac/Ds* group, originally discovered by McClintock in *Z. mays*. Several *Ac* and *Ds* elements have been cloned and their behaviour investigated in transgenic plants of other species (so called heterologous systems) as well as in *Z. mays*. Transposition of *Ac* and *Ds* in a range of plant species is known to require only one transposon-specific gene product. For these reasons, *Ac* and *Ds* were selected as the elements of choice for the work which will be described in this thesis.

1.5 The *Activator* and *Dissociation* Elements

Once McClintock had deduced that *Ds* could transpose to new locations in the *Z. mays* genome, she discovered that *Ac* and *Ds* were responsible for insertion mutations at other loci. Moreover, she found that *Ac*-induced mutations were inherently unstable, but *Ds*-induced mutations were unstable only in the presence of *Ac*. Further, she observed that *Ac*-type mutations could change directly into *Ds*-type mutations, implying that *Ds* elements may be altered *Ac* elements (McClintock, 1965). It has since been confirmed by molecular studies that *Ds* elements are indeed stable forms of *Ac*, which is a simple transposon. *Ac* was originally cloned from *Z. mays* using a *waxy* cDNA clone to probe a genomic library of the *wx-m9 Z. mays* line, which was known to carry an *Ac*-induced mutation of the *waxy* allele (Fedoroff *et al.*, 1983). Behrens *et al.* (1984) subsequently cloned a structurally indistinguishable *Ac* element from the *wx-m7* mutant allele using a similar method. The DNA sequence of the two elements, as determined by Pohlman *et al.* (1984) and Muller-Neumann *et al.* (1984) was identical. *Ac* is 4.6 kb in length, has 11 bp terminal inverted repeats and creates an 8 bp target duplication. The *Ac* element has been shown to express a single heteronuclear RNA of 4.1 kb which after the removal of four introns leaves an mRNA of 3.5 kb encoding an 807 amino acid polypeptide with a calculated mass of 92 kDa (Kunze *et al.*, 1987). This putative transposase enzyme was purified after over-expression in a baculovirus vector (Hauser *et al.*, 1988) and has been shown to interact with a repeated hexanucleotide motif (AAACGG) located in a 70 bp stretch 30 bp inside the 5' terminus of *Ac* (Kunze and Starlinger, 1989). These findings complement the earlier observations of Coupland *et al.* (1988) that deletions of base pairs 44-92 and 75-181 of the 5' terminus of *Ac* strongly reduces excision when transposase is supplied in *trans*. Further similar experiments have demonstrated that a complex set of interactions between the *Ac* transposase protein and sequences in the termini of *Ac* govern the efficiency of transposition (Coupland *et al.*, 1989; Li and Starlinger, 1990; Zhou *et al.*, 1991).

Despite this fact, *Ds* elements capable of transposition are known which possess the 11 bp terminal repeats of *Ac* but no other homologous sequences (Sutton *et al.*, 1984; Wessler *et al.*, 1986).

Three general classes of *Ds* element have thus far been recognized, the members of each class exhibiting diverse structural characteristics. The groups include those elements which are simple *Ac* deletion and/or substitution derivatives, *Ds* elements with minimal *Ac* homology, and the "double *Ds*" elements. Most *Ac*-derivatives merely lack some or all of the transposase coding region (e.g. Fedoroff *et al.*, 1983; Doring *et al.*, 1984b; Dooner *et al.*, 1985). There are also *Ds* elements which range in size from 0.4 kb elements with little *Ac* homology other than the 11 bp terminal repeats, to large elements with intermittent *Ac* homology (e.g. Sutton *et al.*, 1984; Doring *et al.*, 1984a; Schiefelbein *et al.*, 1985; Merckelbach *et al.*, 1986; Wessler *et al.*, 1986). Double *Ds* elements are those elements which were first identified by McClintock as causing chromosome breakage, and most of the several chromosome dissociating elements cloned contain a specific type of *Ds* structure. The first *Ds* element in which the double *Ds* sequence was identified is a 30 kb insertion in the *Z. mays* sucrose synthase *sh-m5933* mutant allele (Doring *et al.*, 1984b). The double *Ds* structure was cloned from one end of the insert and upon sequencing was found to be 4 kb long, consisting of one *Ds* element inserted into a second identical copy in the opposite orientation. The 2.0 kb unit appears to have arisen by the deletion of the central 2.6 kb of *Ac,* leaving 1.0 kb at either end. Analysis of the insertion in the *sh-m5933* allele implied that a second double *Ds* was located at the opposite end of the 30 kb element. This finding has recently been confirmed, and the products of excision investigated by molecular methods (Doring *et al.*, 1989). When one complete *Ds* element excised from one of the double *Ds* structures, a half *Ds* element was left behind, with the resulting structure found to be responsible for chromosome breakage. A second *Ds* element, related to that described above, was identified as a 45 kb insertion in the *sh-m6258* mutant (Doring *et al.*, 1990). This element is also bordered by double *Ds* structures which undergo a similar aberrant excision process to cause a chromosome breakage.

The insertion of some *Ds* elements into genes have produced various interesting properties apart from loss of function mutations. The *Adh1* mutant allele *adh1-Fm335* retains normal regulation and uses the normal transcription start site despite the presence of a 405 bp *Ds* element in the untranslated leader sequence. S1 nuclease mapping indicates that the *Ds* element is spliced from the *adh1-Fm335* message (Dennis *et al.*, 1988). There has also been a report that *Ds* excision from the *Shrunken* mutant locus *sh-m5933* can cause an alteration in tissue specificity of sucrose synthase isozymes (Chourey *et al.*, 1988).

The mechanism of transposition of *Ac* (and *Ds*) is as yet unclear, but the basic features are known. A non-replicative "cut and paste" mechanism for *Ac* and *Ds* was deduced, based on the observation that transposition can restore mutant phenotypes to the wild

type (Saedler and Nevers, 1985). Genetic investigations have also revealed that *Ac* tends to transpose preferentially to genetically linked sites, and does so during chromosome replication. *Ac* elements excising from the *Z. mays P-vv* and *bz-m2* mutant alleles reinserted at linked sites in 61% and 53% of revertants respectively (Greenblatt, 1984; Dooner and Belachew, 1989). Of these reinsertions, the majority were within 5 centimorgans of the donor locus. This tendency appears to hold true for *Ac* elements in transgenic *Nicotiana tabacum* (tobacco) and *Lycopersicon esculentum* (tomato), as described by Jones *et al.* (1990), Dooner *et al.* (1991) and Osborne *et al.* (1991), and for a *Ds* element in the *Z. mays bz-m4* locus (Dowe *et al.*, 1990). Such findings led to the proposal that *Ac* transposase does not generate a free, extrachromosomal intermediate, but mediates the formation of transposition complexes between the *Ac* donor site and competent receptor sites that are physically nearby (Greenblatt, 1984; Dooner *et al.*, 1991). The analysis of twin pericarp sectors in *Z. mays* led Greenblatt and Brink (1962; 1963) to propose that *Ac* transposes during chromosome replication, into either replicated or unreplicated acceptor sites. Molecular investigations by Chen *et al.* (1987) and Schwartz (1989) demonstrated that in the majority of cases, transposition was from a replicated donor locus to an unreplicated target. The occurrence of target duplications at the site of insertion implies that transposases create staggered nicks during the transposition process which act as substrates for enzymes in the host DNA repair mechanism. This model is corroborated by other known mechanisms of eukaryotic DNA cutting and joining, which always involve the creation of staggered nicks and have recognition sequences based upon a CA dinucleotide motif. It is interesting to note in this context that most transposons sequenced thus far end in CA.

Factors controlling the timing and frequency of transposition are less well understood, but the susceptibility to developmental triggers is well documented. For example, many studies on the genetics of *Ac* rely on different patterns of *Z. mays* kernel pigmentation, resulting from different timing and frequencies of transposition, as influenced by the developmental roles played by particular mutated genes. One phenomenon recognized some time ago by McClintock was the effect of *Ac* "dosage." Increasing numbers of active *Ac* elements produce a delay in the timing and a reduction in the apparent frequency of transposition of both *Ac* and *Ds* elements, irrespective of the number of *Ds* elements present. This is known as negative dosage, but is a characteristic which appears to be reversed in transgenic *N. tabacum* plants carrying single and multiple copies of *Ac* (Jones *et al.*, 1989). A second factor which affects *Ac* transposition is DNA methylation. *Ac* elements can be reversibly inactivated by methylation (e.g. see Schwartz and Dennis, 1986; Chomet *et al.*, 1987) and Kunze *et al.* (1988) correlated inhibition of *Ac* transcription with hypermethylation of *Ac* terminal sequences.

1.6 Transposon Tagging

Mutations with a visible phenotype caused by the insertion of a transposon can be analysed at the molecular level if that transposon has previously been cloned. Many genes from *Z. mays* and *A. majus* have been cloned after mutations in them were established to be transposon-induced. The first example of this was the cloning of the *bronze* locus of *Z. mays* using *Ac* as a tag (Fedoroff *et al.*, 1984). A number of genes from the *Z. mays* anthocyanin biosynthesis pathway were subsequently cloned from loci tagged by a particular transposon, e.g. chalcone synthase (Wienand *et al.*, 1986), *opaque-2* (Schmidt *et al.*, 1987) and the genes from the *c* locus (including chalcone synthase; Paz-Ares *et al.*, 1986). These genes are amongst those which, when mutated, were responsible for the kernel pigmentation phenotypes so well studied genetically. These mutable loci are not essential to the plants' normal development, and thus insertional mutagenesis provides a simple approach to their identification. One shortcoming of transposon tagging is that insertions into some genes may produce a lethal phenotype, thus preventing screening for tagged alleles.

However, a transposon tagging strategy can theoretically be applied to any gene within a plant containing transposable elements, providing that a screenable mutant phenotype may be produced. In *Z. mays* and *A. majus*, where plant transposons are best characterized, many copies of most elements exist within the genome, complicating the correlation of a mutation with an individual transposon. Additionally, in almost all commercially important species, few transposons have been identified, and those which have are poorly characterized (with the obvious exception of *Z. mays*). Since one of the advantages of transposon tagging is the isolation of important genes for which the product is unknown, it would be extremely useful to develop systems where single copy transposons were effective as tags in a wide range of species. One way of achieving this is to introduce foreign transposons into plants by genetic transformation. A number of elements from *Z. mays*, as well as *Tam3* from *A. majus*, have been introduced into heterologous backgrounds with encouraging results. In the first such experiment, Baker *et al.* (1986) introduced *Ac* into *N. tabacum* and demonstrated its excision from the T-DNA and reinsertion into the *N. tabacum* genome. *Ac* is also known to transpose in *Arabidopsis* and *Daucus carota* (carrot) (Van Sluys *et al.*, 1987), *Solanum tuberosum* (potato) (Knapp *et al.*, 1988), *L. esculentum* (Yoder *et al.*, 1988), *G. max* (Zhou and Atherly, 1990) and in *Oryza sativa* (rice) (Murai *et al.*, 1991). Furthermore, *Ds* can be trans-activated by *Ac* in transgenic plants (e.g. Hehl and Baker, 1989; Lassner *et al.*, 1989). The En/Spm system is active in transgenic *N. tabacum* (Masson and Fedoroff, 1989) and in *S. tuberosum* (Frey *et al.*, 1989) whilst *Tam3* has been shown to be active in *N. tabacum* (Martin *et al.*, 1989). The above examples clearly demonstrate the potential for the development of a wide host range transposon tagging system, and are reviewed by Haring *et al.* (1991).

One feature of a transposon tagging system which is desirable for the ease of gene

cloning, is the stabilization of the inserted inactivating element. One approach to facilitate this is the use of genetically-marked non-autonomous elements such as *Ds* as tags which can be segregated away from trans-activating elements in crosses. Another more complex way would be to control the activity of transposons via the expression of transposase functions from inducible or tissue specific promoters.

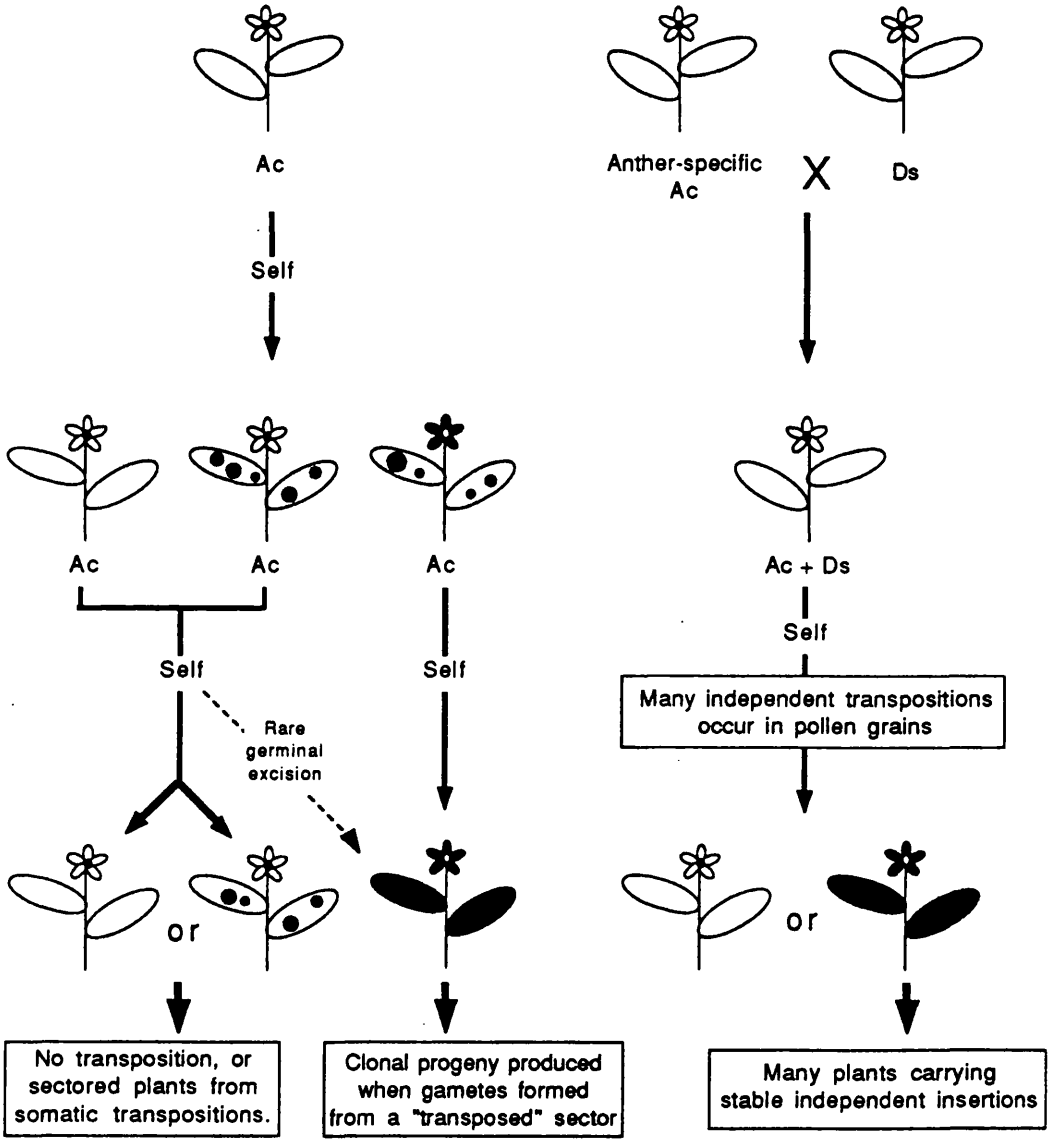
A second problem arising using tagging systems is that most transposon activity is somatic, therefore sectored, or variegated, mutant lines are produced in the first generation (see Fig. 1.1). Germinal transpositions give rise to the required genetic uniformity in plants containing transposed elements, but these occur at low frequencies. The germinal transposition frequency of *Ac* in *Z. mays* is estimated at around 1% - 10% (Brink and Nilan, 1952; McClintock, 1956; Dooner and Belachew, 1989) and is lower in transgenic *N. tabacum* at 1-3% (Jones *et al.*, 1990) and *Arabidopsis* at 0.2% - 0.5% (Schmidt and Willmitzer, 1989). These figures indicate that most plants in the first generation following transposon activation will exhibit sectoring of tissues containing inserted elements, and any showing a desired altered characteristic would require a second generation to produce a non-variegated phenotype, with the requirement that a flowering axis would arise from a sector of interest. This event relies upon the germ cells being derived from a sector containing the transposon insert of interest, so that the gametes will include the mutated gene rather than the wild type. An additional problem arises because variegation of mutants may prevent penetrance of the mutant phenotype, making impossible the identification of phenotypes which are not cell autonomous. Only in plants containing insertion mutations resulting from germinal transpositions would such phenotypes be identified, and furthermore, the mutated genes could be cloned from the same plant if desired. Thus any mechanism which increased the frequency of germinal transposition and reduced somatic transposition would significantly improve the chances of identifying and isolating a tagged gene.

Promoters from genes which have often been considered as constitutively expressed, such as CaMV 35S and octopine synthase, are in fact active at very low levels in the developing gametophyte (e.g. Plegt and Bino, 1989). Thus the use of such promoters to drive transposase expression, whilst increasing somatic transposition activity, is not likely to affect germinal transposition. The approach taken in this thesis was to obtain an anther-specific gene expressed in the male gametophyte during microspore development, and to use its promoter to express *Ac* transposase. Ideally, the promoter of this gene would be active in the interphase before microspore mitosis. This is desirable because of the knowledge that *Ac* and *Ds* tend to transpose mainly during DNA replication. The process of microgametogenesis and the current knowledge of anther-specific gene expression is outlined below.

Figure 1.1.

Schematic representation of transposition patterns following "wild type" and anther-specific transposase expression in transgenic plants.

Black regions indicate sectors derived from a cell in which a transposition has occured.



1.7 Male Gametogenesis in Flowering Plants

The male gametes of angiosperm plants are contained within the pollen grains, which develop in anthers, the male organs of the flower. The pollen grain is the male gametophyte stage of the life cycle, existing as a discrete organism before delivering the sperm cells to receptive ovules, and to withstand the rigours of this free-living period the pollen grain is surrounded by a unique chemically resistant protective wall (Mascarenhas, 1989). The anther is a highly specialized structure which is remarkably similar in all androgynous flowers, and the events leading to the production of the male gametophyte within it are equally well conserved (Bhandari, 1984). In the developing anther, two classes of cells derive from the meristematic primordium, these being the cells comprising the structural body of the anther and the sporogenous cells. The sporogenous cells become enclosed within four locules which form as columns in the connective tissue, and a single cell layer known as the tapetum lines the locules, controlling the transport of nutrients and enzymes to the developing microspores (Carniel, 1963). At a particular stage in development, the sporogenous cells (the pollen mother cells, or meiocytes) after repeated mitotic divisions synchronously enter meiosis. This process, known as microsporogenesis, results in the production of four haploid microspores, which subsequently enter the microspore development phase of gametogenesis (Scott *et al.*, 1991b).

Immediately following microsporogenesis the haploid microspores are contained as tetrads within a "special cell wall" of callose (Heslop-Harrison, 1968), the function of which remains unclear, despite its almost universal occurrence. The deposition of the outer layer of the pollen wall, the exine, begins whilst the microspores remain inside the callose wall. Sporopollenin is the major component of the exine and is synthesized from precursors secreted from the microspores at this time (Echlin and Godwin, 1968), although most is incorporated later when it is secreted by the tapetum. Once exine formation is underway, the uni-nucleate microspores are released from the tetrads by the action of callase (a ß-1,3-glucanase), which is secreted from the tapetum (Mepham and Lane, 1969). Microspore development continues with further exine deposition from the tapetum and the formation of the intine, the inner layer of the pollen wall, consisting of a pectocellulosic matrix including a range of gametophytic proteins (Knox and Heslop-Harrision, 1970; Vithanage and Knox, 1976). During this time, the microspore nucleus completes the extended interphase preceding microspore mitosis. An unequal mitotic division in the microspore produces a large vegetative cell, and a much smaller generative cell. The generative cell nucleus eventually divides to produce the two sperm cells which fertilize the egg cell and endosperm nuclei (Mascarenhas, 1989). In most plants, this second mitotic division takes place in the pollen tube of germinating pollen, thus the mature pollen grain is binucleate at anthesis (Mascarenhas, 1989). However, in some families, including the Brassicaceae, the mature pollen grain is trinucleate, the second mitosis having taken place during pollen

maturation. Pollen maturation is the term given to the phase following microspore development, which ends after the first mitosis. Maturation of the bi- or trinucleate pollen grains sees the completion of pollen wall formation and the storing of reserves for germination, including messenger RNAs as well as proteins and lipid bodies (Mascarenhas, 1990). The tapetum degenerates during pollen maturation and is absent at anthesis, with most of the compounds released from the degenerating tapetum becoming adsorbed onto the surface of the pollen grain, forming the pollen coat (Heslop-Harrison, 1968).

1.8 The Complexity of Gene Expression During Gametogenesis

During the development of the anther, a structure unlike any other found in plants and with a unique function, one would expect to find a large number of specifically expressed genes. Kamalay and Goldberg (1980) determined that around 26,000 different transcripts are present during early anther development in *N. tabacum*, of which around 11,000 are organ-specific. Within the male gametophyte itself, fewer transcripts have been found to be tissue-specific. Willing *et al.* (1984; 1988) have shown that approximately 20,000-24,000 different mRNAs are expressed in mature pollen of *Tradescantia paludosa* and *Zea mays*, of which a maximum of 7200 (30% - 36%) might be pollen-specific, but they judge that this figure may actually be as low as 10% in reality (Mascarenhas, 1989). These results presumably relate to the lower complexity of a pollen grain compared to the whole anther. The difficulty in estimating more precisely the number of pollen/microspore-specific transcripts is likely in part to be due to the expression in pollen of genes encoding different isoforms of proteins present in sporophytic tissues. For example, the α- and ß-tubulins are known to be expressed from different genes in reproductive and vegetative tissues (Ludwig *et al.*, 1988; Hussey *et al.*, 1988).

A second interesting finding came from the investigations of Willing *et al.* (1984; 1988) regarding the levels of expression of genes in pollen compared to shoots. Pollen-expressed genes appear to produce significantly more abundant mRNA levels than their counterparts in shoots. In *Tradescantia* pollen, 75% of total mRNA derives from abundantly expressed genes (of which approximately 40 genes are very highly expressed), whilst only 35% of shoot mRNA consists of highly expressed transcripts. A similar situation was found in *Z. mays*. Additionally, even those transcripts of the low abundance class from pollen were expressed more highly than low abundance messages from shoots. Low abundance mRNAs from *Tradescantia* pollen were present at an average of 100 copies per cell, the equivalent figure in *Z. mays* being 200 copies per cell, contrasting with a figure of 5-10 copies per cell for low abundance messages from shoots. The large concentrations of mRNAs in pollen suggest the requirement for rapid protein synthesis during pollen maturation and/or pollen germination.

No analysis of the complexity of gene expression in anthers or microspores during the middle stages of development have been reported.

1.9 Patterns of Gene Expression During Anther Development

A number of cDNAs and corresponding genes have recently been cloned via the differential screening of cDNA libraries made to mRNA extracted from anthers or mature pollen. The analysis of individual messages with regard to both temporal and spatial expression during anther development has allowed some important generalizations to be made. During early gametogenesis, the large majority of mRNAs thus far examined are expressed specifically in the tapetum (Koltunow *et al.*, 1990; Smith *et al.*, 1990; Scott *et al.*, 1991a). This bias in cDNA cloning ostensibly relates to the transcriptional dominance of the tapetum in immature anthers which results in a high abundance of tapetum-specific clones in libraries (Scott *et al.*, 1991b). The tapetum has for some time been recognized as the most transcriptionally active tissue during early microgametogenesis and this now appears to result, at least in part, from the very high activity of a range of genes, represented by those cloned to date. Despite the estimated low numbers of pollen-specific genes, the ease of collection of mature pollen has enabled cDNA libraries to be constructed which have yielded a variety of pollen-specific clones. Clones isolated in this fashion represent genes which are first expressed after microspore mitosis, mRNA levels being maximal at pollen maturity and anthesis. For some of these clones, message levels decrease markedly during pollen germination (Stinson *et al.*, 1987) and these are therefore believed to belong to the class of stored mRNAs which accumulate during pollen maturation for translation upon germination (reviewed by Mascarenhas, 1990). A different clone isolated in the same study, Zmc13, was characterized more fully by Hanson *et al.* (1989) and the Zmc13 RNA was found not only to accumulate during pollen maturation, but is maintained at a high level during germination and pollen tube growth. In an independent study, McCormick *et al.* (1987) cloned a cDNA LAT52 from a library of mature *L. esculentum* anthers, the peptide sequence of which shows a significant degree of similarity to Zmc13. *In situ* hybridizations showed that the LAT52 mRNA is located in mature pollen grains and in the anther wall (Ursin *et al.*, 1989). The genes corresponding to both of these cDNAs, Zmg13 and LAT52 were subsequently cloned (Hamilton *et al.*, 1989; Twell *et al.*, 1989) and it was demonstrated that their promoters were active with the same temporal and spatial specificities in a number of transgenic plant species (Guerrero *et al.*, 1990, Twell *et al.*, 1990).

These pollen-specific genes are representative of the later of the two classes of genes expressed during microgametogenesis. This classification has been formed on the basis of the expression patterns of individual mRNAs (reviewed by Mascarenhas, 1990) and of the profiles of extracted proteins and *in vitro*-translated proteins from staged microspores and pollen of *Triticum aestivum* (wheat) (Vergne and Dumas, 1988), *Z.*

mays (Bedinger and Edgerton, 1990) and of *Lilium* and *N. tabacum* (Schrauwen *et al.*, 1990). The "late", or "class II" genes, as illustrated by the above examples, are expressed following microspore development and throughout pollen maturation, with levels of mRNA maximal just prior to anthesis. Few "early," or "class I" microspore expressed genes have been identified, but they are represented by genes such as actin and alcohol dehydrogenase, which although not microspore-specific, are expressed in the microspores following tetrad release, and maximally around the time of microspore mitosis, message levels falling rapidly to negligible levels at maturity (Stinson *et al.*, 1987; Stinson and Mascarenhas, 1985). Recently, several microspore-specific clones showing class I - type expression have been isolated from *B. napus* (Albani *et al.*, 1990; 1991; and Scott *et al.*, 1991a). The promoter of the *B. napus* Bp4 gene retains microspore-specific activity in transgenic plants (Albani *et al.*, 1990) implying together with the results of the LAT52 and Zmg13 promoter studies that the factors regulating gametophytic gene expression are conserved in flowering plants.

At the present time, the function of any of the cloned microspore-specific class I genes is unknown, with the exception of the Bp19 gene of *B. napus*, which is speculated to have pectin esterase activity (Albani *et al.*, 1990). Tapetum-specific genes expressed during the period when microsporogenesis and microspore development is occurring include two *N. tabacum* phospholipid transferase proteins implicated in nutrient transfer (Koltunow *et al.*, 1990). Of the later expressed anther- and pollen-specific genes, two *L. esculentum* clones, LAT56 and LAT59, encode proteins with similarities to pectate lyases of *Erwinia* (Wing *et al.*, 1989) and are expressed in the pollen tube. These enzymes may function in the mobilization of reserves for pollen tube cell wall growth, or to aid in penetration of the stylar transmitting tissue. A similar function has been proposed for the P2 gene product from *Oenothera organensis*, which is also present in germinating pollen tubes, and has a high degree of identity (54%) with *L. esculentum* polygalacturonase (Brown and Crouch, 1990). The *N. tabacum* clone TA56 (Koltunow *et al.*, 1990) encodes a thiol peptidase, and has been shown by *in situ* hybridization to be localized in the anther connective and stomium, maximally at the time of anthesis. This gene is thus implicated in the degradation of anther wall tissues leading to dehiscence.

The cloning of tissue-specific cDNAs is thus proving to be a productive method for the elucidation of gene function during male gametogenesis, and provides a method of obtaining anther-, microspore- and pollen-specific promoters for use in transgenic studies. Of the genes cloned to date, only Bp4 and Bp19 (Albani et al., 1990; 1991) provide promoters which are active in the microspores at the time when it is thought that expression of a transposase enzyme would be most useful for driving transposition for the purposes previously outlined. These genes are protected by patent however, and are not available for unrestricted studies of potential commercial importance. This thesis describes the isolation of a novel *Arabidopsis* gene exhibiting the desired anther-specific expression pattern.

1.10 Overview of Thesis Aims

The preliminary aim of the project upon which this thesis is based was to determine the activities of the *Ac* and *Ds* transposable elements in transgenic *L. usitatissimum* tissues. An *in vivo* excision assay for these elements was made available (see Chapter 3) which allowed the rescue of transformed tissue containing excised elements by antibiotic selection. The transformation of *L. usitatissimum* and the behaviour of *Ac* and *Ds* in transgenic callus is discussed.

The second aim of the project, which forms the main body of the thesis, was to develop methods of improving transposon tagging strategies. The isolation of an anther-specific gene which is expressed in developing microspores is described. The controlling sequence from this gene was fused to the *Ac* transposase gene, and the use of this construct to promote germinal transposition in transgenic plants is discussed.

CHAPTER 2

MATERIALS AND METHODS

Recipes and formulae of molecular biological buffers and solutions which are refered to in this chapter are given in Appendix III.

2.1 Sources of Molecular Biological Reagents, Enzymes, and Plant Tissue Culture Chemicals

Molecular biology grade chemicals and reagents were purchased from Sigma Chemical Company Ltd., or from BDH Ltd. Enzymes were obtained from BRL, Boehringer Mannheim, Stratagene, Pharmacia or NBL. Tissue culture salts and hormones were purchased from Flow Laboratories. Agar and agar based media were from Difco Laboratories.

2.2 Bacterial Culture and Storage

2.2.1 Strains and Genotypes

Eschericia coli:

XL1-Blue : recA1, endA1, gyrA96, thi, hsdR17, supE44, relA1, lac, {F', proAB, laclq, ZΔM15, Tn10, (tet^R)}

LE392 : hsdR514, hsdM, supE44, supF58, lacY1 or Δ(lacIZY)6, galK2, galT22, metB1, trpR55.

P2392 : P2 lysogen of LE392.

JM109 : recA1, endA1, gyrA96, thi, hsdR17, supE44, relA1, Δ(lac-proAB), {F' traD36, proAB, $lacI^qZ\Delta M15$}

HB101 : hsdR, hsdM, supE44, ara14, galK2, lacY1, proA2, rspL20, xyl-5, mtl-1, recA13, mcrB.

Agrobacterium tumefaciens:

LBA4404, binary construct host strain.

C58C1 rif (pGV2260), binary construct host strain.

The strains C58C1 rif (pGV3850HPT::pKU2), C58C1 rif (pGV3850HPT::pKU3) and C58C1 rif (pGV3850HPT::pKU4) were constructed by Dr. George Coupland and received as a gift from Prof. Peter Starlinger.

2.2.2 Growth of Bacterial Cultures

Liquid cultures were inoculated from bacterial colonies or from glycerol stocks. Single colonies were picked using a pipette tip or sterile tooth pick into the appropriate volume of NB or NZY medium containing selective antibiotics. (All bacterial growth media are described in Appendix I). From glycerol socks, a small piece of the frozen stock was removed using a flame-sterilized scalpel blade and placed into culture medium. Cultures were grown overnight, or until the required amount of growth had occurred, in a shaking

incubator at 37°C for *Eschericia coli*, or 28°C for *Agrobacterium tumefaciens*. Single colonies were produced by streaking cell suspensions from an overnight culture onto agar-solidified media using a flamed bacteriological loop. Media was cooled to around 50°C before adding antibiotics and pouring into petri dishes. Streaked plates were inverted and incubated at 37°C or 28°C overnight to allow colonies to form. Single colonies were alternatively produced by spreading (with a flamed glass spreader) 100-300 µl of cell suspensions onto selective media. All cultures were handled within the confines of a laminar air flow cabinet to prevent contamination.

2.2.3 Storage of Bacterial Cultures

Cultures were maintained for short periods of time (1-2 months) as single colonies on sealed plates at 4°C in a fridge. For longer periods of time, glycerol stocks of cultures were made from overnight cultures mixed 1:1 with a solution of 60% glycerol in nutrient broth in a 2 ml cryogenic tube. Stocks were flash frozen in liquid nitrogen and stored at -80°C indefinitely.

2.2.4 Antibiotics

The following antibiotics were commonly used for selection of plasmids, episomes, or bacterial strains carrying resistance genes, at the given concentrations. All antibiotics were supplied by Sigma Chemical Co.

ANTIBIOTIC	CONCENTRATION FOR SELECTION (µg/ml)	
	E. coli	*A. tumefaciens*
Ampicillin	100	-
Kanamycin	100	50
Rifampicin	100	50
Tetracycline	12.5	2.5
Cefotaxime	-	250

2.2.5 Conjugation of Plasmids from *E. coli* to *A. tumefaciens*

Plasmids used for plant transformation were based on binary vectors carrying replication origins (RK2 replicons) allowing maintainance in *E. coli* and *A. tumefaciens*. Such plasmids were initially manipulated in *E. coli*, then mobilized into *Agrobacterium* by the introduction of helper functions on a separate plasmid. The plasmid pRK2013 (Ditta *et al.*, 1980) carries the required mobilization (*mob*) and transfer (*tra*) genes. Transfer was achieved in a single tri-parental mating between the donor *E. coli* strain, the recipient *Agrobacterium* host and the helper strain (HB101::pRK2013). Overnight cultures of each strain were mixed (100 µl each) and spread onto a nutrient agar plate. Plates were then incubated overnight at 28°C. Streaks of the resulting lawn of cells were removed on a bacteriological loop and streaked onto agar containing antibiotics selective for the *Agrobacterium* host and the introduced recombinant plasmid. Plates were

incubated for 2 days at 28°C before single colonies were grown in liquid culture and analysed for correct transfer.

2.2.6 Small Scale Isolation of *Agrobacterium* Total Nucleic Acids

This method is as described in Draper *et al.* (1988).

A 5 ml culture of the *Agrobacterium* strain was grown overnight from a single colony. Cells were pelleted from 1.5 mls of culture by microcentrifugation for 5 minutes, and resuspended in 300 µl of TE, to which was then added 100 µl of 5% Sarkosyl. After mixing, 150 µl of pronase was added and the mixture incubated for 1 hour at 37°C. The pronase treated cell suspension was then mixed with 500 µl of phenol/chloroform, by repeatedly passing through a 1 ml pipette tip. After centrifugation, the aqueous solution was submitted to this treatment a further 3 times. DNA was precipitated from the remaining supernatant by the addition of one twentieth volume of 5 M NaCl and 3 volumes of ethanol. The nucleic acid was collected by microcentrifugation for 10 minutes, rinsed in 70% ethanol, dried, and dissolved in 50 µl of sterile distilled water.

2.3 Plant Tissue Culture

N.B. All culture media are described in Appendix II

2.3.1 Transformation of *Linum usitatissimum* var antares (Flax)

L. usitatissimum hypocotyls were transformed using the method described in Basiran *et al.* (1987). *L. usitatissimum* seeds (obtained from Booker Seeds) were sterilized in 10% bleach for 20 minutes before sowing on sterilized vermiculite soaked with Hoaglands solution in glass Kilner jars. Seedlings were grown for 2 days following germination and then harvested for hypocotyl transformation. Hypocotyl sections 1-2 mm in length disected from directly behind the apical meristem were inoculated for 10 minutes by submersion in a 1 in 50 dilution in MSD4x2 medium of an overnight culture of *Agrobacterium*. Excess liquid was removed from the explants before they were placed (20 per 9 cm petri dish) onto solid MSD4x2 medium. Hypocotyls and bacteria were cocultivated for 2 days at 28°C in a constant environment growth room with 16 hour photoperiod before transfer onto MSD4x2 medium containing selective antibiotics (250 µg/ml cefotaxime and 30 µg/ml hygromycin). Plates were returned to the growth room to allow production of transformed callus tissue from the cut ends of the hypocotyls. Hygromycin resistant calli were maintained by subculture onto selective MSD4x2 medium.

2.3.2 Transformation of *Nicotiana tabacum* (Tobacco)

Petite Havana SR1 and Petite Havana *N. tabacum* plants were transformed using a leaf disk method described in Draper *et al.* (1988). Young, just fully expanded leaves were taken from greenhouse grown plants and sterilized in 10% bleach for 15 minutes. The

leaves were then washed in four changes of sterile tap water. Leaf disks of about 0.5-1 cm square were cut (using sterile dissection equipment) avoiding the mid-rid and major leaf veins. Disks were placed onto MSD4x2 solid medium and incubated in the growth room for 2 days to promote callus formation at the cut edges of the leaf. After this time, the explants were inoculated by immersion for 15 minutes in a suspension of *Agrobacterium* (an overnight culture diluted 1 in 50 in MSD4x2 medium). After removal of excess liquid, the leaf disks were replaced on the same MSD4x2 plates and incubated for a further 2-3 days at 28°C for cocultivation. After this time (or when *Agrobacterium* colonies appeared on the plates) explants were transferred (5-6 per 9 cm petri dish) onto MSD4x2 plates containing selective antibiotics (250 µg/ml cefotaxime and 100 µg/ml kanamycin). Plates were incubated in the growth room until antibiotic resistant shoots arose from the callus around the leaf pieces. Once these shoots had 2 or 3 well formed leaves, they were sterilely disected away from callus material and transferred to MSO medium plus 250 µg/ml cefotaxime and 100 µg/ml kanamycin in suitably sized sterile containers. Shoots produced roots after 1-3 weeks and these were allowed to grow until about 2 cm in length. Plantlets were then transferred to soil, sealed in a plastic bag to provide a humid atmosphere for the first one or two days' growth, and then grown as normal in a contained growth room.

2.3.3 Transformation of *Arabidopsis thaliana*

Arabidopsis thaliana (C24 ecotype) cotyledons were transformed using a method developed in Leicester by Kenan Turgut and Mehdi Barghchi (manuscript in preparation). C24 seeds were sterilized in 10% bleach for 10 minutes, washed in sterile tap water and sown on MSO medium. Plates were incubated in the growth room until 4 days after germination. At this time cotyledons were sterilely dissected away from the apical meristem and floated in liquid MSO medium, until approximately 30 or 60 cotyledons had been collected. These were then transferred in a sterile sieve to a 1 in 50 dilution in MSO medium of an overnight *Agrobacterium* culture for 3 minutes. The cotyledons were dried by placing the sieve on a sterile filter paper and transferred (30 per 9 cm petri dish) to MSD4x2 medium. When 300 cotyledons per transformation experiment had been treated in this way, plates were incubated for 2-3 days for cocultivation at 28°C. Explants were then transferred to MSD4x2 medium containing 400 µg/ml augmentin and 50 µg/ml kanamycin.

The cotyledons were subcultured every 10-14 days onto selective MSD4x2 medium, eventually discarding those which did not produce viable green callus. Shoots arising from kanamycin resistant calli were cut away from the callus and placed in MSO medium with the addition of 1 µg/l naphthalene acetic acid plus 400 µg/ml augmentin for 3 days. Shoots were then transferred to a medium containing MS salts, 1.5% sucrose and 200 µg/ml augmentin for rooting.

2.3.4 *In Vitro* Germination of *N. tabacum* Pollen

Pollen was collected into microcentrifuge tubes from dehisced anthers and germinated in a medium containing 15% sucrose and 0.1 mg/ml boric acid. Germination was allowed to procede at 28°C for 2-4 hours before microscopic examination and use in gene expression analysis.

2.3.5 Handling of Transgenic Seed

Transformed plants were grown until they flowered, at which time flowering heads were isolated to ensure self fertilization, and also to aid in the containment of transgenic pollen. In the case of *N. tabacum*, plastic bags were used to cover the inflorescences, whilst *Arabidopsis* plants were maintained individually in closed plastic containers. Seeds were collected after flowering and held in labelled containers. Transgenic seed was germinated sterilely on MSO medium containing antibiotics to select for those plants carrying a T-DNA within their genome. This permitted an analysis of segregation of T-DNAs to establish the number of loci of insertions within the primary transformed plants, as well as selecting for transformed plants for further study.

2.4 Identification and Isolation of Tissue-Specific mRNAs via cDNA Cloning

2.4.1 A *Brassica napus* Anther cDNA Library

A cDNA library constructed from mRNA extracted from *B. napus* anthers of between 2 mm and 4 mm in length was obtained from Dr. Rod Scott (Scott *et al.*, 1991a). In *B. napus*, there is a strong correlation between bud and anther lengths and the developmental stage of the anthers. It has been shown by Scott *et al.* (1991a) that *B. napus* anthers of 2-4 mm in length are in the microspore development phase of gametogenesis. That is, they include microspores that have been released from the meiotic tetrads through to those about to enter the second microspore mitosis. The library was constructed in λZAP (Stratagene) a bacteriophage lambda insertion vector which contains a multiple cloning site within pBluescript SK^- phagemid sequences. The filamentous f1 phage origin of replication in this phagemid allows the rescue (via R408 helper phage co-infection) of cDNA inserts in pBluescript plasmid. This replication origin also permits the production of single stranded plasmid DNA, useful for DNA sequencing, by co-infection with the filamentous helper phage M13K07.

2.4.2 Plating of Lambda Phage

Competent *E. coli*, strain XL1-Blue, were produced by inoculating 10-50 mls of NZY medium containing 0.2% maltose and 10 mM $MgSO_4$ with 100-200 µl of an overnight

culture. When the culture reached an OD_{600} of 0.5-0.8, the cells were pelleted by centrifugation for 7 minutes at 3000rpm, resuspended in 10 mM $MgSO_4$, and stored for up to 3 days at 4°C.

To obtain plaques of lambda clones on a bacterial lawn, 600-800 µl of competent cells were mixed with the desired number of p.f.u. (plaque forming units) from the library and incubated at 37°C for 15 minutes to permit adsorption of phage particles to the cells. To this suspension as added 8 mls of top agar (NZY solidified with 0.7% agarose) at 48°C, and the resulting mixture poured immediately onto solid NZY medium in 14 cm petri dishes. Plates were incubated inverted for 8-16 hours at 37°C. For 9 cm plates, 150-200 µl cells and 3 mls of top agar were used.

2.4.3 cDNA Library Screening

Plates were cooled to 4°C in the fridge to harden the top agar prior to plaque lifts being taken. Circles of "Hybond-N" (Amersham) nylon membrane were cut to the size of the plates in use, and membranes were laid onto plates for 30 seconds, during which time orientation marks were made using a needle and marker pen to enable later realignment of the filter to the plate. Membranes were laid plaque-side-up on filter paper soaked in denaturing solution for 7 minutes, then transferred similarly to paper soaked in neutralizing solution for 7 minutes. The membranes were finally rinsed in 2 X SSC, dried, and DNA bound to the filter by cross-linking on an ultra violet transilluminator (Ultra Violet Products inc.) which had previously been calibrated to show that 90 seconds was an optimum time for linking.

Plaque lifts were pre-hybridized for at least 1 hour in DNA hybridization solution at 42°C to block non-specific binding before the addition of labelled probe, and then hybridized at 42°C overnight. Filters were washed in three changes of wash solution A at 65°C for 20 minutes, and if higher stringency was required, subsequently in two changes of wash solution B at 65°C for 15 minutes. After washing, filters were blotted dry on filter paper, wrapped in "Saran Wrap" plastic film and exposed to X-ray film. Autoradiograms were developed after a period of 2 hours to 2 days (depending on signal) and any hybridizing plaques located on the original plates.

2.4.4 Storage of Isolated Phage Clones

Clones were isolated by coring plaques out of plates using a sterile Pasteur pipette and eluting phage particles into 500 µl suspension medium (SM) containing a drop of chloroform to prevent bacterial growth. These phage suspensions can be stored for several years at 4°C. In cases where cores were taken from densely plated phage, a second round of plating and screening was carried out in order to obtain a single plaque.

2.4.5 Preparation of Radio-Labelled DNA Probes

DNA probes were produced using the random priming method of Feinberg and Vogelstein (1984). DNA was isolated by preparative agarose gel electrophoresis, and used as a low melting point agarose gel slice probe stock, or purified from the gel. Probe stocks were denatured by boiling for 5 minutes before 10 ng of DNA was removed and added to the labelling reaction as outlined below:

Oligolabelling buffer*	-	3.0 µl
10 mg/ml BSA	-	0.6 µl
^{32}P-α-dCTP	-	1.5 µl
DNA polymerase I Klenow fragment	-	0.6 µl
Probe stock	-	x µl
Sterile distilled water	to	15 µl

*See Appendix III.

Synthesis of labelled DNA was allowed to proceed for 1 hour at 37°C before incorporation of ^{32}P-dCTP was measured. Probe was then boiled for 3-5 minutes to denature the DNA and added to the hybridization reaction.

2.4.6 Measurement of Isotope Incorporation into DNA

The labelling reaction was made up to 100 µl by the addition of 85 µl of stop solution and 1 µl transferred to 0.5 mls of ice cold 500 µg/ml herring sperm DNA. This DNA mix was precipitated with 125 µl of 50% trichloroacetic acid (TCA) for 5 minutes on ice. A second 1 µl aliquot of probe solution was spotted onto the centre of a Whatman GF/C glass fibre disk. The precipitated DNA was transferred onto a second GF/C disk on a filter tower, washed twice with 10% TCA, once with methylated spirit and dried. Both disks were then placed in vials for liquid scintillation counting. Incorporation was calculated as a percentage of the precipitated (incorporated) counts against the total counts in the labelling reaction.

2.4.7 Rescue of pBluescript Carrying cDNA Inserts from Lambda ZAP via an *In Vivo* Excision Protocol

200 µl of host *E. coli* cells (XL1-Blue) at OD_{600} = 1.0 grown in the presence of tetracycline were added to 200 µl of phage stock and 1 µl of R408 helper phage (>7.5×10^{10} p.f.u./ml). Co-transfection was carried out at 37°C for 15 minutes before adding 5 mls 2xYT medium (Appendix I) and shaking at 37°C for 3-6 hours. Tubes were then heated to 70°C for 20 minutes to lyse the bacteria and release pBluescript packaged as a filamentous phage-like particle. Cell debris was removed by centrifugation at 4000rpm for 10 minutes and the supernatant phagemid stock stored at 4°C (for up to 2 months). pBluescript was rescued by mixing 10 µl of the above stock and 10 µl of a 1:100 dilution, with 200 µl of XL1-Blue cells at OD_{600} = 1.0, transfecting for 15 minutes at 37°C and spreading 50 µl onto NA (Appendix I) plates containing ampicillin.

Overnight incubation of these plates at 37°C produced colonies of cells containing pBluescript in a plasmid form. cDNA inserts could then be isolated by restriction digests of purified plasmid DNA.

2.5 Analysis of Gene Expression by Northern Blotting

2.5.1 Extraction of Plant RNA

Total RNA from plant tissues was extracted essentially as described in Draper *et al.* (1988). Tissue was flash frozen in liquid nitrogen and ground with sand in a mortar and pestle until a fine powder was produced. To this was gradually added grinding buffer (2 mls/g fresh tissue) whilst grinding continued. The homogenized mixture was transferred to a 50 ml polyallomer centrifuge tube and an equal volume of phenol/chloroform added. This mixture was left to thaw on ice and emulsified by inversion. The protein precipitate was removed by a 5 minute centrifugation at 3500rpm (4°C) and the aqueous supernatant subjected to three further phenol/chloroform extractions. To the resulting solution, 0.05 volumes of 4.0 M sodium acetate pH 6.0 and 2.5 volumes of -20°C ethanol were added. Total nucleic acids were collected by centrifugation at 10,000 rpm in siliconized Corex tubes for 10 minutes. The pellet was resuspended in a minimum volume of sterile distilled water and high molecular weight RNAs (>5S) precipitated by adding 3 volumes of 4.0 M sodium acetate pH 6.0 followed by incubation on ice for 2-3 hours. RNA was collected by centrifugation at 10,000rpm for 10 minutes, resuspended in a small volume of sterile distilled water and the concentration measured by scanning spectrophotometry and adjusted to 10 mgs/ml.

2.5.2 Scanning Spectrophotometry of Nucleic Acids

Nucleic acids absorb light maximally at ultra violet wavelengths, with an absorption peak at 260nm. Measurement of OD_{260} can thus be used to determine concentration. Nucleic acid solutions were diluted into 300 µl of distilled water, placed in a quartz cuvette and scanned across the range 200-300nm. A sharp peak at 260nm and a trough at 220-240nm indicates a pure nuceic acid preparation. An OD_{260} of 1.0 is equivalent to a 37 µg/ml RNA solution, or a 50 µg/ml double stranded DNA solution.

2.5.3 Northern Dot Blotting

1-10µg of total RNA was denatured by heating to 65°C for 5 minutes in 3 volumes of RNA denaturing solution then chilled on ice and 1 volume of cold 20x SSC added. Samples were spotted onto Hybond-N membrane (pre-wetted in 10x SSC) in 2 µl aliquots, dots being allowed to dry between applications. The RNA was fixed to the filter by 90 seconds u.v. cross-linking.

2.5.4 Northern Gel Blotting

Total RNA, denatured as described above, was run on a 1.5% agarose gel including marker RNAs of known size where appropriate. The ethidium bromide stained gel was photographed before equilibrating for 20 minutes in 20x SSC and capillary transfer of RNA onto nylon membrane (see section 2.7.3). After overnight transfer, the filter was dried and u.v. crosslinked.

2.5.5 Hybridization of DNA Probes to RNA Bound to a Nylon Support

Filters were prehybridized in RNA hybridization solution for at least one hour before the addition of radiolabelled probes. Hybridization was carried out overnight at 42°C. Washes were carried out as for DNA:DNA hybridizations (section 2.4.3). Hybridizing mRNAs were then detected by autoradiography.

2.6 Construction and Analysis of a Genomic Library from *Arabidopsis thaliana*

2.6.1 Isolation of High Molecular Weight Plant DNA

DNA was isolated using a CTAB (hexadecyltrimethylammonium bromide) extraction method, essentially as describd by Draper *et al.* (1988). Leaf tissue (avoiding major veins to minimize contaminating carbohydrates) was flash frozen in liquid nitrogen and ground to a fine powder in a mortar and pestle. This powder was transferred to a 50 ml polyallomer centrifuge tube mixed with 2 mls/g of 60°C 2 x CTAB buffer and incubated for 30-60 minutes at 60°C. Proteins were precipitated by gentle emulsification in an equal volume of chloroform : octanol (24:1) and removed by centrifugation. The aqueous phase was added to a tenth volume of 60°C 10% CTAB solution and re-extracted with chloroform : octanol to remove final traces of proteinaceous material. The DNA/CTAB complex was precipitated with an equal volume of cold isopropanol and collected by centrifugation at 4000rpm for 10 minutes. The pellet was washed with 70% ethanol and air dried before resuspension in sterile distilled water. The resulting DNA solution is suitable for gemonic cloning or Southern blot analysis.

2.6.2 Partial Digestion of Genomic DNA

Sau 3AI is a restriction endonuclease with a tetrameric recognition site, common in genomic DNA. Partial cleavage with *Sau* 3AI generates the random overlapping fragments required for construction of a representative genomic library. Partial digests were carried out, and conditions selected to generate fragments in the size range 10-25kb; the size of insert accepted by the cloning vector. To do this digests were set up in a large volume containing a low concentration of enzyme, and aliquots removed at timed intervals.

Reaction:

DNA	-	38 µl (50 µg)
10xReact Buffer	-	100 µl
Sau 3AI Enzyme	-	2 µl (2 Units)
Water	-	860 µl
Total Volume	-	1000 µl

Aliquots of 12 µl and 150 µl were removed at 2 minute intervals from 4 minutes onwards (after the addition of enzyme to pre-warmed reactants at 37°C) and frozen on dry ice. After the reacton was complete, the 12 µl samples were analysed by electrophoresis on a 0.6% agarose gel. The 150 µl aliquots of time points including fragments mainly in the 10-25 kb size range were used for cloning.

2.6.3 Dephosphorylation of Restriction Fragments

In order that genomic fragments could not concatomerize in the ligation reaction, the partially digested DNA was treated with calf intestinal alkaline phosphatase (CIP) to remove 5' phosphate moieties. The partial digest was thawed and ethanol precipitated, resuspended in 10 µl of sterile distilled water, and incubated with 0.05 units of CIP in CIP buffer at 37°C for 30 minutes. The enzyme was inactivated by incubation at 65°C for 45 minutes in the presence of 30 mM trinitrilotriacetic acid. DNA was recovered by ethanol precipitation after phenol/chloroform extraction.

2.6.4 Ligation to Lambda DashII *Bam* HI-Prepared Arms

The genomic library was constructed in the replacement vector Lambda DashII (Stratagene). This was purchased as pre-prepared *Bam* HI digested phage arms, ready for ligation onto genomic DNA fragments.

Ligation Conditions:

(0.5 µg) *Sau* 3AI digested genomic DNA	-	1 µl
(1 µg) Lambda DashII / *Bam* HI arms	-	1 µl
5x Ligation buffer	-	1 µl
2 units/µl T4 DNA ligase	-	1 µl
Distilled water	-	1 µl
Total volume	-	5 µl

The ligation mixture was incubated at room temperature (22°C) for 1 hour and left at 4°C overnight.

2.6.5 Packaging of Lambda DNA

Gigapack II Gold packaging extracts, purchased from Stratagene, were used as specified by the manufacturer. DNA is packaged into phage particles by mixing a ligation mixture with the two packaging extracts supplied and incubation for 2 hours at 22°C. The two packaging extracts contain complementary isolations of the proteins required for phage head assembly and insertion of DNA into phage particles. After packaging, the reaction was suspended in 500 µl SM plus 20 µl chloroform.

Dilutions of 1 µl of this stock were plated on P2392 lysogenic *E. coli* grown to phage competency as described in section 2.4.2 to select for recombinant phage and determine the titre of the library. The Lambda DashII vector employs the *Spi* selection system, whereby only recombinant phage will grow on a lysogenic strain of *E. coli*. This is because the inhibition of secondary transfection afforded by a lysogen depends upon the *red* and *gam* genes of phage Lambda; these are located on the region of Lambda DashII which is replaced by genomic DNA during cloning.

2.6.6 Amplification of the *Arabidopsis* Genomic Library

The entire genomic library of 200,000 p.f.u. was plated on competent P2392 cells on 14 cm petri dishes at a density of around 30,000 p.f.u. per plate. The plates were incubated for 8 hours so that plaques were small and did not become confluent. To amplify the library, the plates were overlayed with 10 mls of SM and left to stand at 4°C overnight. The phage suspension was decanted into 50 ml polyallomer tubes and the plates rinsed with a further 2 mls of SM which was pooled with the rest of the suspension. Bacterial cells were lysed by adding chloroform to 5% and incubating the tubes for 15 minutes at room temperature. Cell debris was removed by centrifugation at 4000rpm for 5 minutes and the supernatant stored in a fresh tube with chloroform at 0.3%. This stock was stored at 4°C with several aliquots taken for long term storage. 70 µl of 100% dimethyl sulphoxide was added per ml of stock before flash freezing the aliquots in liquid nitrogen and storing them at -80°C.

2.6.7 Preparation of Phage Lambda DNA

Bacteriophage lambda DNA for restriction digestion and Southern blot analyses was prepared from lysed liquid cultutres as described in Sambrook *et al.* (1989).

50-100 µl of a single plaque suspension was adsorbed to 400 µl competent LE392 cells for 15 minutes at 37°C, 5 mls of NZY medium was added and cultures grown overnight in a 37°C shaking incubator. These cultures became lysed through phage propagation; remaining cells were lysed by adding 50 µl chloroform and continuing shaking for 15 minutes. Cell debris was removed by centrifugation for 10 minutes at 4000rpm. The supernatant was removed to a new tube and treated with 35 units of DNase I and 5 µg RNase A at 37°C for 30 minutes to digest *E. coli* nucleic acids. Phage particles were precipitated by the addition of 20% polyethylene glycol (PEG) 6000 + 2 M NaCl in SM and incubation on ice for 1 hour. Phage were pelleted by centrifugation at 10,000rpm for 10 minutes in Corex tubes. All traces of supernatant were removed by aspiration before resuspending the pellet in 0.5 mls SM. This suspension was transferred to a microcentrifuge tube and residual solids removed by centrifugation at 13,000rpm for 2 minutes. DNA was then extracted by disrupting the phage by adding 5 µl 10% SDS and 5 µl 0.5 M EDTA and incubating tubes at 65°C for 5-10 minutes. Proteins were then extracted with phenol/chloroform and the DNA precipitated from aqueous solution with

an equal volume of -20°C isopropanol. The nucleic acid was collected by microcentrifugation for 10 minutes, rinsed in 70% ethanol and dried under vacuum. The pellet was resuspended in 50 µl of distilled water, and 5-10 µl used for restriction digestion.

2.7 Restriction Enzyme and Southern Blot Analysis of DNA

2.7.1 Digestion of DNA with Restriction Endonucleases

Restriction enzymes were purchased with an accompanying buffer, and used as per manufacturers recommendations. Up to 2 µg of plasmid or phage DNA, or 10 µg of plant genomic DNA were incubated for 1-16 hours at the appropriate temperature with the enzyme and buffer, and usually 0.05 µg RNase A. Where necessary, enzymes were removed by phenol/chloroform extraction, or denatured by heat.

2.7.2 Agarose Gei Electrophoretic Separation of DNA

Nucleic acids fragments were routinely separated on agarose gel matrices. 2-3 µl of loading buffer was added to the nucleic acid sample to be analysed before pipetting into the well of a gel. Gels were made by dissolving agarose in TAE buffer at 95°C; agarose concentrations varied between 0.6 and 2% depending upon the expected size range and/or required resolution of fragments. Gels were run submerged in TAE buffer at 10-100 Volts, with appropriate molecular weight markers included on the gel. DNA was stained by the inclusion in the gel of 0.1 µg/ml ethidium bromide, which intercalates nucleic acids and can be visualized by it's fluorescence under u.v. light on a transilluminator.

2.7.3 Southern Blotting

Gels were prepared for Southern transfer by soaking in 3 gel volumes of depurinating solution for 7 minutes, denaturing solution for 30-120 minutes and neutralizing solution for 30-120 minutes. The gel was rinsed in water between treatments. Transfer was carried out by capillary blotting as follows. The gel was laid on a piece of Whatman 3MM filter paper on a sponge standing in a tray of 20x SSC. The remaining surface of the paper and sponge was covered with plastic cling film. On top of the gel was placed the nylon membrane, ensuring no air bubbles were trapped between it and the gel. The membrane was marked with pencil to allow orientation and comparison to a photograph of the stained DNA in the gel. Two pieces of 3MM filter paper soaked in 20x SSC were laid on top of the membrane, and a further two pieces of dry 3MM on top of these. A stack of paper towels was then placed over the gel and a 0.5 kg weight put on the top. Transfer of buffer from the tray to the towels through the gel was left to occur overnight. The membrane was finally rinsed in 2 x SSC, dried and the DNA bound to the filter by 90 seconds u.v. cross linking. Hybridization of transferred DNA to probe sequences of interest was carried out using the same procedure as described in section 2.4.3.

2.7.4 Re-use of Blots

Probes were removed from filters by incubation at 45°C for 30 minutes in 100 mls of 0.4 M NaOH, before washing in 100 mls of 0.1 x SSC, 0.1% SDS, 0.2 M Tris-HCl pH 7.5 for a further 30 minutes at the same temperature. Blots were then stored at 4°C before prehybridization and hybridization with a different probe.

2.8 Subcloning of DNA into Plasmid Vectors

Small fragments of DNA (up to 10 kb) were generally manipulated in the plasmids pBluescript SK^- and SK^+. These phagemid vectors have a high copy number in *E. coli*, contain a selectable ampicillin resistance gene, and a multiple cloning site in the *lac* Z gene, permitting a simple colour assay for recombinants. Other vectors used for more specific tasks, such as the binary vectors pBin19 and pBI101 will be discussed below.

2.8.1 Large Scale Preparation of High Purity Plasmid DNA

This method is taken from Sambrook *et al.* (1989).

2.8.1.1 Alkaline Lysis of Cultures

Cultures of cells carrying the required plasmid were grown overnight under selection in 100-500 mls of NB (Appendix I). Cells were pelleted by centrifugation and resuspended in 2.5 mls of lysis solution per 100 mls of culture. To this suspension was added 2 volumes of alkaline SDS and mixed by gentle inversion. 1.5 (original) volumes of 3 M potassium acetate pH 5.2 were then added to precipitate protein and chromosomal DNA. This was spun off in a microcentrifuge for 10 minutes and the plasmid precipitated from the supernatant with 0.6 volumes of cold isopropanol. Nucleic acids were collected by microcentrifugation for 10 minutes and resuspended in 0.5 mls distilled water per 100 mls culture.

2.8.1.2 Purification of Plasmid

High molecular weight RNAs were precipitated by the addition of an equal volume of ice cold 5 M lithium chloride and removed by microcentrifugation for 5 minutes. The remaining nucleic acids were precipitated with an equal volume of cold isopropanol, collected by microcentrifugation for 10 minutes, rinsed in 70% ethanol and the pellet dried under vacuum. This pellet was resuspended in 100 µl of 20 µg/ml RNaseA in 0.1 mM EDTA, 0.1 mM NaCl, and left for 15 minutes at 37°C. The plasmid DNA was then precipitated by adding 100 µl of 1.6 M NaCl / 13% PEG 6000 and collected by microcentrifugaton for 5 minutes. The pellet was dissolved in 40 µl distilled water and the solution extracted once with phenol/chloroform. The DNA was finally precipitated by the addition of an equal volume of 10 M ammonium acetate and 2 volumes of ethanol. The plasmid was collected by microcentrifugation, rinsed in 70% ethanol and dried before dissolving in the required volume of sterile distilled water.

2.8.2 Purification of DNA From Agarose Gels

DNA molecules such as those produced after restriction digestion of genomic clones were cloned after purification from agarose gel slices following electrophoresis. DNA was recovered from agarose by binding to silica glass suspensions in high salt conditions, washing, and eluting the DNA into water. The "Geneclean II" DNA purification kit was purchased for this purpose from Bio 101 Ltd. and used as per manufacturers instructions. Fragments purified in this way were ready for ligation into vectors, or enzyme modification prior to ligation.

2.8.3 Filling-In of Protruding 5-Prime Termini

In order to ligate fragments produced by restriction enzymes which leave a 5' overhang into blunt-ended vectors, the overhang was filled in by synthesizing a second strand with the Klenow fragment of DNA polymerase I. DNA was purified from restriction digests by phenol/chloroform extraction and ethanol precipitation, and resuspended in 10 µl of sterile distilled water. To this was added 3 µl of AB buffer (made by combining oligolabeling soltions A and B (Appendix III) in the ratio of 2:5) 0.6 µl 10 mg/ml BSA, 0.5 µl (1 unit) Klenow enzyme and 1 µl of 2 mM dCTP. The reaction was left at 37°C for 1 hour, the Klenow enzyme heat inactivated for 5 minutes at 70°C and the DNA recovered by ethanol precipitation.

2.8.4 Ligation of DNA Fragments

Purified DNA was mixed in approximately equimolar ratios of vector and insert, and 5x ligation buffer (Appendix III) and 2 units of T_4 DNA ligase were added into as small a volume as possible.

Typical Ligation Reaction:

Vector DNA	- x µl
Insert DNA	- y µl
5x Ligation Buffer	- 3 µl
T4 DNA Ligase	- 1 µl
Water	to 15 µl

Ligations were generally left at 4°C overnight, though in some instances, sticky-ended ligations were left at room temperature for 2-4 hours before transformation into competent *E. coli*.

2.8.5 Transformation of *E. coli* with Plasmid DNA

This method is essentially as described in Sambrook *et al.* (1989).

2.8.5.1 Preparation of Competent Cells

An overnight culture of the required host strain was subcultured into 100 mls of fresh medium, and grown to an OD_{600} of 0.6 before the cells were collected by centrifugation

at 3000 rpm for 10 minutes. The pellet was resuspended in 50 mls of ice cold 50 mM $CaCl_2$ and left on ice for 15 minutes if for immediate use, or 1 hour if the cells were to be frozen and stored. The bacteria were collected by centrifugation at 4°C, resuspended in 10 mls of 50 mM $CaCl_2$ and used immediately, or alternatively resuspended in 10 mls of 50 mM $CaCl_2$ in 20% glycerol, aliquoted into 100 µl batches and flash frozen in liquid nitrogen. Frozen cells were stored at -80°C.

2.8.5.2 Transformation Procedure

To a 100 µl batch of competent cells thawed on ice was added a completed ligation reaction and 10 µl of 10 x TCM buffer. This mixture was incubated on ice for 30 minutes. The cells were then heat shocked by incubation at 37°C for exactly 2 minutes, followed by incubation at room temperature for 10 minutes. A second heat shock was given for 2 minutes at 42°C after the addition of 1 ml of pre-warmed 37°C NB. The cells were then shaken at 37°C for 1 hour, pelleted in a microcentrifuge and resuspended in 100 µl of NB. Transformant colonies were produced by spreading this suspension on selective NA plates and incubating at 37°C overnight.

2.8.6 Identification of Recombinant Plasmids in Transformed Colonies

A colour assay using X-GAL + IPTG is the simplest method of identifying colonies containing recombinant plasmids, but is only posible when using certain vectors. These vectors are those which contain a multiple cloning site within the *lac*Z α gene product. This protein, the α-subunit of ß-galactosidase, complements a mutant gene in the host and facilitates the metabolism of the indigogenic substrate X-GAL (5-bromo-4-chloro-3-indoyl-ß-D-galactopyranoside) when induced by IPTG (isopropyl-ß-thio-galactopyranoside) forming a blue product easily visible in colonies. When this gene is interupted by an insert, recombinant colonies are white and can thus be distinguished from non-recombinants. For colour selection, 20 µl of 40 mg/ml IPTG in water, and 20 µl of 40 mg/ml X-GAL in dimethylformamide were spread onto plates prior to spreading cells.

2.8.7 Colony Hybridization

When colour selection was not available, large numbers of colonies were screened by transfering them onto nylon membrane, denaturing the bacteria and probing with the required DNA in a hybridization reaction. Nylon membrane cut to the size of the culture plate was laid onto the plate for 2.5 minutes, during which time orientation marks were made, then transferred colony side up onto Whatman 3MM filter paper soaked in 2 x SSC + 0.5% SDS and left for a further 2.5 minutes. The filter was microwaved on full power for 2.5 minutes to dry the filter and fix the DNA to the membrane. The filter was then hybridized as for plaque lifts (section 2.4.3). Clones of interest were recovered by replacing the culture plate in the 37°C incubator for about 6 hours to regrow colonies.

Putative recombinants were then characterized further by isolating small quantities of plasmid DNA and submitting it to restriction enzyme digestion.

2.8.8 Small Scale Isolation of Plasmid DNA

This method is essentially as described in Birnboim and Doly (1979).

Recombinant colonies were picked into 5 mls of selective medium and grown overnight. Cells were then pelleted from 1.5 mls of culture in a microcentrifuge, and plasmid prepared by alkaline lysis as described in section 2.8.1. The cell pellet was resuspended in 100 µl of lysis solution and lysed by mixing with 200 µl of alkaline SDS. Chromosomal DNA was precipitated with 150 µl of 3 M potassium acetate pH 5.2 and removed by microcentrifugation for 10 minutes. The remaining solution was extracted with phenol/chloroform and the plasmid DNA precipitated by the addition of 2 volumes of ethanol. The nucleic acid was collected by microcentrifugation for 10 minutes, rinsed in 70% ethanol and dried before dissolving in 50 µl of sterile distilled water. 5 µl of this DNA was used for restriction digestion.

2.8.9 Amplification of Sequences by the Polymerase Chain Reaction

The standard PCR conditions used in this work to amplify DNA from clones in plasmids were as follows:

Reactants :

Target DNA (1ng/µl)	-	1 µl
10xPCR Buffer	-	3 µl
Primers (200ng/µl)	-	1µl+1µl
Taq DNA polymerase I (5U/µl)	-	0.3 µl
Sterile distilled wate	to	30 µl

Cetus DNA thermal cycling machine program :

Denaturing step = 94°C for 1 minute; Annealing step = 50°C for 1 minute; Extension step = 72°C for 2.5 minutes.

Reactions were carried out with 10 cycles and the products analysed by electrophoresis.

2.9 DNA Sequencing

2.9.1 Production of Single Stranded DNA from Phagemids

The cell line harbouring the plasmid containing the DNA to be sequenced was grown in 2xYT (Appendix I) to an OD_{600} of 0.5-0.8. A 2 ml aliquot of this culture was removed and infected with filamentous helper phage M13KO7 at a multiplicity of infection of 10 by shaking at 37°C for 1 hour. 400 µl of these cells were then transferred to 10 mls of medium containing 70 µg/ml kanamycin and grown overnight. The resulting culture produces phagemid DNA packaged as single stranded filamentous phage which is extruded from cells into the medium.

2.9.2 Purification of Single Stranded DNA

The cells from 1.5 mls of culture produced as above were pelleted by microcentrifugation for 5 minutes, and the upper 1 ml of the resulting supernatant removed to a new tube. To this was added 250 µl of 20% PEG 8000 / 2.5 M NaCl to precipitate phage particles. The mixture was left at room temperature for 15 minutes and the phage collected by microcentrifugation for 10 minutes. The pellet was resuspended in 200 µl of 1 mM Tris-Cl, 0.1 mM EDTA, 0.2% ß-mercaptoethanol, after careful removal of all the supernatant, and the phage coats disrupted by the addition of 200 µl of phenol and vortexing. After 1 minute, 50 µl of chloroform was added and the mixture vortexed again. The aqueous DNA solution was then recovered after microcentrifugation for 5 minutes and the nucleic acid precipitated with 10 µl of 5 M NaCl and 500 µl of ethanol. After 10 minutes microcentrifugation, the DNA was rinsed in 70% ethanol, dried, and dissolved in 33 µl of water. 3 µl of this solution was analysed by electrophoresis, and 10 µl used for sequencing.

2.9.3 Sequencing Using T7 DNA Polymerase

Templates were sequenced using a T7 DNA polymerase sequencing kit (purchased from Pharmacia) which employs dideoxynucleotide chain-terminating mixes, as described by Sanger *et al.* (1977). Templates primed with the appropriate oligonucleotide were sequenced with ^{35}S-labelled dATP, and the products separated by denaturing polyacrylamide gel electrophoresis. Gels were made by polymerizing 60 mls of 6% acrylamide (5.7% acrylamide, 0.3% bis-acrylamide in 7 M urea and 1 x TBE) by the addition of 48 µl tetramethylethylenediamine (TEMED) and 360 µl 10% ammonium persulphate. This mixture was poured into sealed Bio-Rad "Sequi-Gen" sequencing plates and left to polymerize for 1 hour. The gel was pre-run to 50°C in 1xTBE buffer before loading pre-heated samples. Gels were dried onto Whatman 3MM filter paper and exposed to autoradiographic film at room temperature overnight without the use of an intensifying screen.

2.9.4 Data Handling

Sequence data was compiled and analysed on the Leicester University Computer Centre VAX mainframe, using the Wisconsin Genetics Computer Group (GCG) program suite (Devereux *et al.*, 1984) and other DNA analysis programs provided by the university molecular biology users group.

2.10 Analysis of Gene Expression in Transgenic Plant Tissues

2.10.1 Dot Assay Technique for Detection of NPT-II Activity

A simple protein dot assay was used to screen large numbers of tissues using a method described by McDonnell *et al.* (1987). Around 0.5 g of tissue was frozen and ground in

liquid nitrogen to a fine powder, and transferred to a microcentrifuge tube. Extraction buffer (DEB) was added at 1 ml/g and mixed by vortexing for 60 seconds. Samples were kept on ice. The plant debris was pelleted by microcentrifugation for 5 minutes at 4°C and the supernatant removed to a new tube. 15 µl of this extract was mixed with 15 µl of assay buffer (DRB) and incubated at 37°C for 30 minutes. During this time, Whatman P81 paper was prepared for dot blotting. The paper was marked into 2 cm squares, soaked in 20 mM ATP + 100 mM sodium pyrophosphate and dried. After the assay was complete, samples were again microcentrifuged for 5 minutes before spotting 20 µl onto the P81 paper. The dried paper was washed for 2 minutes in 80°C 10 mM sodium phosphate buffer pH 7.5, and then in 3-5 changes of room temperature buffer for 10 minutes. Blots were dried and autoradiographed overnight.

2.10.2 Detection of NPT-II in Polyacrylamide Gels

This method is as described by Draper *et al.* (1988). N.B. All manipulations were carried out in a 4°C cold room.

Protein was extracted by grinding 100 mg of callus or 250 mg of plant tissue in 20 µl/100 mg of extraction buffer (GEB). During grinding, 5 µl of 10 mg/ml BSA and 2 µl of PMSF (phenylmethyl sulphonyl fluoride at 7 mg/ml in isopropanol) were added to the extraction. Samples were microcentrifuged for 3 minutes and the supernatant removed to a new tube. Aliquots were then removed for protein quantification at a later stage. The remaining sample was loaded onto a 10% non-denaturing polyacrylamide gel and electrophoresed in the cold room for 16 hours at 100 V. The apparatus was dismantled and the resolving gel trimmed to the exact size of the assay base plate cut from teflon sheet. The gel was rinsed in distilled water and left for 15 minutes to equilibrate in reaction buffer (GRB) at 4°C. The gel was then transferred to the base plate, any gaps being filled with 2% low gelling temperature agarose (LGT) and the reaction gel overlay added. This was pre-prepared as follows: {0.075 x gel area} mls of melted 2% LGT agarose cooled to 37°C was mixed with {0.075 x gel area} mls of 37°C reaction buffer and {0.2 x gel area} µl of 25 mg/ml kanamycin. {0.5 x gel area} µCi of γ-^{32}P-ATP was added immediately before pouring this reaction mixture over the protein gel. The reaction was left to proceed at room temperature for 30 minutes. Whatman P81 paper was placed over the gels and a stack of 12 3MM filter papers laid on top. A 500g weight was placed on the filters and left to blot for 3 hours. The blot was washed in hot (80-90°C) 10 mM Na_2PO_4 pH 7.6 3 or 4 times for 5 minutes, dried, and autoradiographed.

2.10.3 Approximated Protein Quantification via Coomassie Staining

Total protein concentration in the NPT-II extracts was determined by a Coomassie blue "dot binding assay." Samples were dotted neat and at 1:1 and 1:9 dilutions onto Whatman 3MM filter paper. Next to the samples was dotted a range of BSA standards at 1-10

mg/ml. The filter was immersed in Coomassie (0.1% Coomassie blue dye, 25% methanol, 10% acetic acid) for 20-30 minutes at 37°C on a slow shaking platform. The filter was destained for 1 hour in 25% methanol, 10% acetic acid. Protein concentrations were estimated by comparison to the BSA standards of the intensity of staining.

2.10.4 Histochemical Localization Of ß-Glucuronidase (GUS) Enzyme

GUS assays were carried out following the methods described by Draper *et al.* (1988). Fresh tissues and sections were suspended in 50 mM $NaPO_4$ pH 7.0, 0.5 mM potassium ferricyanide, 0.5 mM potassium ferrocyanide, 1 mM EDTA and 1 mg/ml X-Gluc (5-bromo-4-chloro-3-indolyl glucuronide) and incubated at 37°C for 20 minutes to overnight, depending on the intensity of staining. The presence of active GUS enzyme is indicated by a blue precipitate in the tissues.

For more accurate localization of GUS, anther tissues were fixed prior to staining using the method of Koltunow *et al.* (1990). Thick sections were fixed in 0.1 M $NaPO_4$ buffer containing 0.1% formaldehyde, 0.1% Triton-X 100 and 0.1% ß-mercaptoethanol by vacuum infiltration for 15 minutes. The sections were then rinsed extensively in 100 mM $NaPO_4$ pH 7.0 containing 0.1% ß-mercaptoethanol followed by 50 mM $NaPO_4$ pH 7.0 containing 0.1% ß-mercaptoethanol. The staining reaction was carried out in 50 mM $NaPO_4$ pH 7.0, 0.1% Triton-X 100, 0.1% ß-mercaptoethanol and 0.5 mg/ml X-Gluc. The sections were re-fixed in 5% formaldehyde, 5% acetic acid, 20% ethanol for 2 hours and cleared by shaking in 50% ethanol for 2 hours and 100% ethanol overnight.

2.10.5 Quantitation of GUS Activity Using a Fluorimetric Assay

Protein was extraced by grinding plant material in GUS extraction buffer. The supernatant was recovered after microcentrifugation for 5 minutes and used as a crude GUS preparation. Reactions were carried out by adding an appropriate amount (10-50 µl) of protein extract to 500 µl of 37°C pre-warmed assay buffer (extraction buffer supplemented with 1mM methyl umbelliferyl glucuronide; MUG) and continuing the incubation at 37°C. A time course for the reaction was produced by removing 20 µl aliquots from the reaction into wells of an opaque microtitre plate containing 180 µl of 0.2 M Na_2CO_3 stop solution after zero, 10, 20, 30, and 60 minutes. The fluorescence of the GUS-catalysed breakdown product of MUG, 4-methyl umbelliferone (4-MU) was measured using a fluorimeter, a linear increase in fluorescence during the reaction time course indicating GUS activity. Specific enzyme activities were calculated by comparison of fluorescence values to those produced by known concentrations of 4-MU, and standardizing the results relative to the quantity of total protein present in the reaction.

2.10.6 Protein Quantitation using Bradford's Assay

Assays were carried out in a microtitre plate using a method based on that of Bradford (1976). In the wells of the microtitre plate were combined 100 µl distilled water, 100 µl Bradford's solution and 1-10 µl of the protein sample. Solutions of bovine serum albumin in the range zero to 1.0 mg/ml were used as standards to calculate protein concentration. Concentrations were measured directly using a custom-written program on a Dynatech MR5000 microtitre plate reader.

CHAPTER 3

BEHAVIOUR OF THE MAIZE TRANSPOSABLE ELEMENTS *Ac* AND *Ds* IN TRANSGENIC FLAX CALLUS.

3.1 Introduction

Transposon tagging of genes in a wide variety of plants might be possible if an autonomous transposable element which is active in any chromosome background can be identified. The *Z. mays* transposable element *Ac* is increasingly becoming a likely candidate for such a role. Transposition has thus far been demonstrated following the introduction of *Ac* into a wide range of plant species, as discussed in Chapter 1. It was intended that genes encoding resistance to the *L. usitatissimum* rust pathogen, *M. lini* might be targeted for cloning by transposon tagging. To this end, constructs carrying *Ac* and *Ds* elements were used to transform *L. usitatissimum* in order to examine their behaviour in this species.

Several systems have now been developed which permit phenotypic assays for the excision of transposable elements in plants. The first of these was developed by Baker *et al.* (1987) and uses the reconstitution and expression of a neomycin phosphotransferase II (*npt-II*) gene following transposon excision to monitor activity. The *npt-II* gene is initially disrupted by the presence of the transposable element in the untranslated leader sequence. Three such constructs, containing either the *Ac* element (pKU3), the *Ds* element (pKU4), or no element (pKU2) were constructed by Dr. George Coupland and kindly received as a gift from Professor Peter Starlinger for the work with *L. usitatissimum*. The organization of these clones is illustrated in the restriction maps shown in Fig. 3.1. The experimental procedure was to compare the frequencies of generation of kanamycin resistant tissues from *L. usitatissimum* explants transformed with each construct.

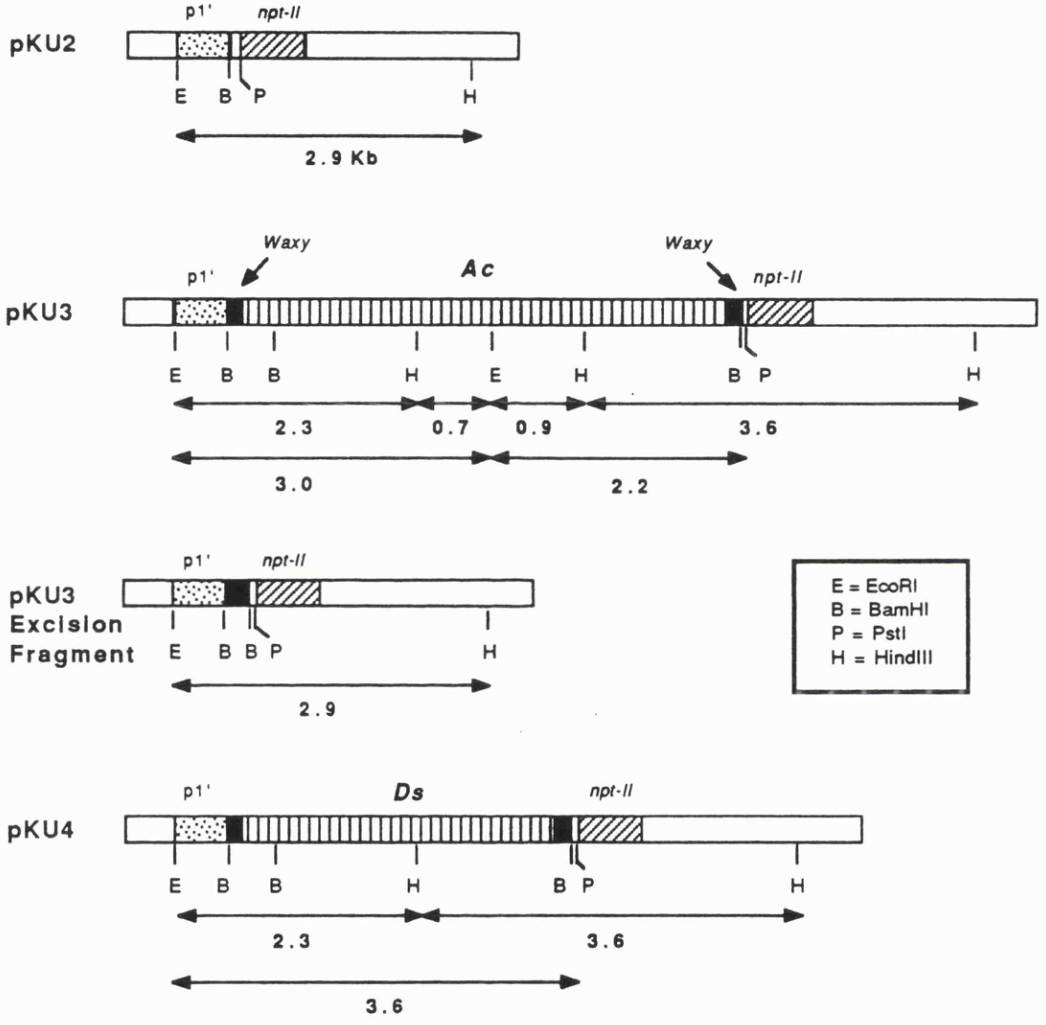
3.2 Results

3.2.1 Transformation and Tissue Culture

Previous work in Leicester had established a transformation system for *L. usitatissimum* which involved regeneration of plants from callus produced from the cut ends of hypocotyl sections from *L. usitatissimum* seedlings (Basiran *et al.*, 1987). This system was chosen as a method for obtaining transformed *L. usitatissimum* tissue carring transposable elements in T-DNA constructs. The pKU series of constructs described above had been recombined into the *Agrobacterium* strain C58C1rif carrying a derivative, pGV3850HPT, of the disarmed Ti plasmid pGV3850 (Zambryski *et al.*, 1983) which carries a hygromycin phosphotransferase (*hpt*) gene on the T-DNA, allowing selection for transformants using the antibiotic hygromycin B. In order to determine the optimal concentration of hygromycin for use in tissue culture, sterile *L. usitatissimum* hypocotyl sections were placed on MSD4x2 medium containing zero, 5, 10, 15, 30 or 50 µg/ml of the antibiotic. After 10 days incubation in a 28°C controlled environment growth room, the extent of callus production was observed, and 30 µg/ml was chosen as the lowest concentration which effectively inhibited callus formation. For

Figure 3.1

Diagram showing the structures of the constructs used to transform *L. usitatissimum*, and the derived structure resulting from *Ac* excision in transformed plant tissues. Restriction fragment sizes are shown in kilobase pairs. The p1' fragment is the 1' promoter from octopine TR-DNA (Velten *et al.*, 1984). The *Waxy* DNA is derived from the *wx* allele of *Z. mays* (encoding starch glucosyl transferase) from which the *Ac* element was isolated (Fedoroff *et al.*, 1983).


pKU2
p1'
npt-II
E
B
P
H
2.9 Kb
pKU3
Waxy
Ac
Waxy
p1'
npt-II
E
B
B
H
E
H
B
P
H
2.3
0.7
0.9
3.6
3.0
2.2
pKU3
Excision
Fragment
p1'
npt-II
E
B
B
P
H
2.9
E = EcoRI
B = BamHI
P = PstI
H = HindIII
pKU4
p1'
Ds
npt-II
E
B
B
H
B
P
H
2.3
3.6
3.6

selection of kanamycin resistant *L. usitatissimum* callus, Basiran *et al.* (1987) had determined that 200 µg/ml was an appropriate concentration for use.

Control uninoculated hypocotyls and hypocotyls inoculated with agrobacteria carrying the pKU2, pKU3 and pKU4 constructs were transferred onto MSD4x2 media containing the various selective agents (combinations of cefotaxime, hygromycin and kanamycin). Callus growth was evident on hypocotyls under no selection, or on cefotaxime only, as early as 7 days after inoculation. Two or three days later, inoculated hypocotyls began to produce callus under hygromycin selection. In the presence of kanamycin, callus production was significantly delayed, and the high degree of swelling observed in other hypocotyls was reduced. Twenty to thirty days post-inoculation, growth of explants had progressed to a stage where the differences in generation of hygromycin and kanamycin resistant callus were apparent. All hypocotyls produced large amounts of shoot-bearing callus on media containing only the bacteriostatic agent cefotaxime. On media containing hygromycin, only hypocotyls inoculated with agrobacteria harbouring pGV3850HPT constructs produced callus. This demonstrates that the *hpt* gene on the T-DNA permitted selection for transformed tissues. It was observed that callus was produced from all inoculated hypocotyls, indicating a high competency for transformation. Similarly, all hypocotyls inoculated with the strain carrying the pKU2 construct produced kanamycin resistant callus. Whilst hypocotyls always produced hygromycin resistant callus when inoculated with the pKU3 and pKU4 strains, no kanamycin resistant callus was ever generated by pKU4-inoculated explants. Around one third of hypocotyls inoculated with the pKU3 strain did, however, form callus on medium containing kanamycin, implying that the *npt-II* gene could sometimes be reconstituted if interrupted by *Ac*, but not if by *Ds*.

A summary of the numbers of hypocotyls transformed, and the antibiotic resistant callus produced, is shown in Table 3.1. The appearance of typical hypocotyls transformed with each construct under kanamycin selection after 30 days incubation is shown in Fig. 3.2. Resistant callus of interest was subcultured on selection media for up to 3 months before bulking up for protein and DNA extractions. Subcultured callus lines grew vigorously in the presence of selective antibiotics. It was found that hygromycin resistant calli, once established, were also resistant to kanamycin at high concentrations (greater than 500 µg/ml). It was thus not possible to select for calli expressing NPT-II amongst randomly chosen hygromycin resistant lines from pKU3 and pKU4 transformation experiments.

3.2.2 NPT-II Activity in Kanamycin Resistant Calli

Twenty eight kanamycin resistant pKU3 calli, derived from independently transformed hypocotyls, were subjected to the rapid dot assay for NPT-II activity (section 2.10.1). All 28 were positive, as were control pKU2-transformed calli, whilst untransformed callus contained no detectable NPT-II activity. A representative selection of these results is shown in Fig. 3.3. A number of these positive lines were then used to perform *in situ*

polyacrylamide gel assays (section 2.10.2) which provide information about the size of the NPT-II protein. Protein extracts from pKU3 lines were run on gels along with a bacterial NPT-II extract and an extract from pKU2-transformed callus. The assays (e.g. Fig. 3.4) demonstrated that pKU3-derived NPT-II enzyme co-migrated with the bacterial and pKU2 enzymes. Approximate specific activities of the pKU3 enzymes, estimated as the intensity of signal on autoradiograms relative to the amount of total protein loaded, were generally lower than the pKU2 reconstructions. NPT-II activity was not detected in hygromycin resistant pKU4-transformed callus in similar assays.

3.2.3 T-DNA Analysis by Southern Blotting

More detailed information on the behaviour of *Ac* and *Ds* in *L. usitatissimum* was obtained from Southern analyses. DNA was extracted from a number of NPT-II positive pKU3 callus lines and some hygromycin resistant pKU4 lines, and digested with enzymes which would generate restriction fragments of sizes characteristic of intact T-DNA or T-DNA from which *Ac* (or *Ds*) had excised. DNA digested with *Eco* RI plus *Hin* dIII generates a fragment containing the p1' promoter of 2.9 kb if the *Ac* element has excised from the T-DNA, but only 2.3 kb if the T-DNA is intact (see Fig. 3.1). Southern blots of such DNA digests, when probed with the p1' promoter fragment, demonstrated the presence of both fragments in all nine pKU3 lines tested (e.g. Fig. 3.5). DNA was also digested with the enzymes *Eco* RI plus *Pst* I, which generate *Ac* -containing fragments of unpredictable sizes from elements reinserted in the genome, and fragments of 2.2 and 3.0 kb from *Ac* elements still in their original positions in the T-DNA. Hybridizations of the *Ac* probe to Southern blots of these digests failed to detect any *Ac* sequences other than the 2.2 and 3.0 kb fragments characteristic of elements in the T-DNA. No T-DNA alterations were detected in the DNA of pKU2- or pKU4-transformed callus lines.

Table 3.1

Table showing transformation data for inoculations with agrobacteria harbouring the constructs shown. For each transformation, shown in the first column are the numbers of hypocotyl explants plated onto the various selection regimes, and in the second column, the numbers of hypocotyls which produced viable callus.

Figure 3.2

Representative hypocotyl explants transformed with pGV3850HPT::pKU plasmids after 30 days in the presence of kanamycin.

a. Untransformed

b. pKU2

c. pKU3

d. pKU4

SELECTION	UNINOC-ULATED		pKU2		pKU3		pKU4	
Cefotaxime	20	20	20	20	20	20	20	20
Cefotaxime + Hygromycin	20	0	40	40	127	127	63	63
Cefotaxime + Kanamycin	40	0	50	50	287	102	85	0

a b c d

Figure 3.3

Typical NPT-II dot blot assay, showing spots of ^{32}P-phosphorylated kanamycin produced by incubation with protein extracts from the following callus lines:

a and **f**; independent pKU2 transformants,

b, **c**, **d**, **e**, **h**, and **i**; independent pKU3 transformants,

g; untransformed callus,

j; pKU4 transformant.

Figure 3.4

Non-denaturing polyacrylamide gel *in situ* NPT-II assay. The major signal in each lane represents ^{32}P-phosphorylated kanamycin. Protein extracts were from the following sources; the amount of protein loaded onto the gel is also indicated:

Lane 1; Bacterial NPT-II extract,

Lanes 2 and 8; pKU2 transformant (0.77 mg each lane)

Lanes 3 - 7; pKU3 callus lines (0.63, 0.47, 0.51, 0.73, and 0.68 mg respectively),

Lane 9; Untransformed callus (0.3 mg).


a b c d e
f g h i j
1 2 3 4 5 6 7 8 9

3.3 Discussion

3.3.1 Generation of Kanamycin Resistant Calius

The behaviour of hypocotyls on media without antibiotics compared to the behaviour of transformed hypocotyls on media including selective agents demonstrates that the selection regimes employed were effective. Only transformed callus was hygromycin resistant, and only hypocotyls transformed with the pKU2 construct, which constitutively expresses NPT-II, consistently produced kanamycin resistant callus.

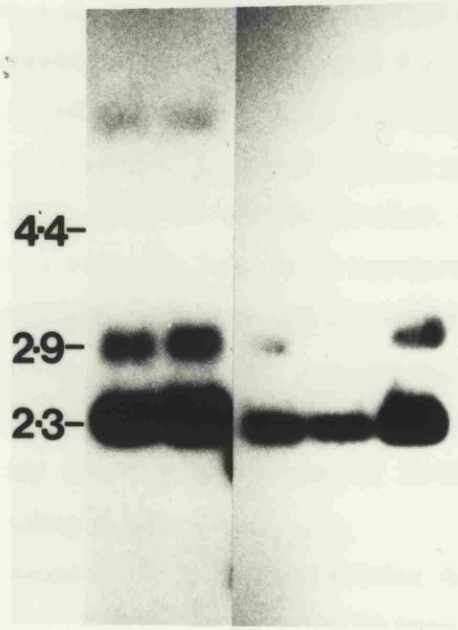
The results in Table 3.1 show that no hypocotyls transformed with pKU4 produced kanamycin resistant callus, although explants from the same inoculations all produced hygromycin resistant callus. This implies that the *Ds* element interrupting the *npt-II* gene is stable, and does not excise from the T-DNA. This was expected, as in the construction of pKU4, the 1.6 kb internal *Hin* dIII fragment of the *Ac* coding region was deleted to produce an artificial *Ds* element which cannot produce an active transposase enzyme. The data also shows that 102 out of the 287 hypocotyls inoculated with *Agrobacterium* harbouring the pKU3 construct and incubated on medium including kanamycin, produced resistant callus. This result implies that in a number of transformed cells, *Ac* has excised from the T-DNA and reconstituted the *npt-II* gene, resulting in the growth of callus. It was noted in the results that kanamycin resistant calli arose around 10 days after the appearance of hygromycin resistant calii from hypocotyls inoculated at the same time. This may infer that the *Ac* element is activated around 10 days after it's integration into the *L. usitatissimum* genome. This activation may be related to the dedifferentiation process going on in the wounded, cut surface of the hypocotyl. Transposon activation has been linked with dedifferentiation in several speculative reports, but Peschke *et al.* (1987) demonstrated a direct correlation between tissue culture-induced dedifferentiation and *Ac* activation in *Z. mays*. However, the observations made in the case of *L. usitatissimum* may simply be due to the response of the expiants to the different antibiotics, though kanamycin resistant calius was produced by pKU2 transformants well before pKU3 transformants.

Previous studies on the behaviour of plant transposable elements have used phenotypic excision assays to measure transposition frequency (see Haring *et al.*, 1991 for review). For example, Baker *et al.* (1987) found that 25% of tobacco protoplasts transformed with pKU3 generated kanamycin resistant calli, whilst Knapp *et al.* (1988) found excision frequencies as high as 50% in shoots derived from *S. tuberosum* leaf disks. In this study, 35% of inoculated *L. usitatissimum* hypocotyls produced kanamycin resistant callus. It is difficult to relate this figure to an absolute frequency of excision however, as one is unable to asses the number of transformed cells which could potentially have undergone excision and restoration of *npt-II*. Despite this complication, in the context of excision frequencies as measured by other workers with similar systems, 35% kanamycin resistance compared to 100% hygromycin resistance would mean that a 35% "excision frequency" in *L. usitatissimum* is a reasonable estimate. An

Figure 3.5

Southern blot of DNA extracted from the five pKU3 callus lines assayed by NPT-II *in situ* gel analysis. DNA was digested with *Eco* RI and *Hin* dIII and the blot probed with the p1' promoter. Fragment sizes are indicated in kilobase pairs.

(All lanes contained a hybridizing band at 2.9 kb, but a faint band in the fourth lane was lost during photographic reproduction.)


4·4–
2·9–
2·3–

attempt was made to test a number of hygromycin resistant lines transformed with pKU3 for kanamycin resistance in order that a more accurate estimate of absolute excision frequency could be made. This proved unsuccessful however, as it was found that the growth of established callus was not inhibited even by high concentrations of kanamycin. This included pKU4-transformed lines which could not express NPT-II. It would have been possible to perform NPT-II assays on a large selection of hygromycin resistant lines, but this was not deemed worthwhile.

3.3.2 Kanamycin Resistance was due to *npt-II* Expression

Inoculated hypocotyls were placed immediately onto selective media, and any callus produced was thus presumed to be expressing NPT-II. This assumption was substantiated by the generation of callus in the predicted fashion from explants inoculated with the three pKU constructs. However, in order to verify that the *npt-II* gene was expressed in kanamycin resistant pKU3 calli (as a result of *Ac* excision) the NPT-II enzyme assays were carried out. The results demonstrated that resistant calli did indeed contain active NPT-II, whilst sensitive callus lines did not. Furthermore, the gel assay demonstrated that active NPT-II enzymes extracted from pKU3-transformed callus lines co-migrated with the native bacterial protein and active NPT-II from pKU2-transformed callus. This shows that the NPT-II activity expressed from reconstituted genes in pKU3 have similar charge and size characteristics to the native protein. It is thus unlikely that the enzymes assayed had suffered any amino acid additions or deletions as a result of T-DNA rearrangements or incomplete excisions.

The *p1'-npt-II* gene produced in pKU3 following precise *Ac* excision is not identical to that in the control construct, pKU2, since there are 60 base pairs of *waxy* DNA flanking the cloned *Ac* element which remain in pKU3 after excision. However, these 60 bp are situated in the untranslated leader of the chimaeric gene and contain no stop codons, thus normal transcription and translation should not be affected.

3.3.3 Southern Data Confirms *Ac* Excision in *L. usitatissimum*

The Southern analyses were designed to show whether *Ac* had left the T-DNA in kanamycin resistant tissues as a complete 4.6 kb element, and to attempt to detect any reinserted elements within the genomes of the same callus lines. DNAs probed with the p1' promoter sequence showed bands characteristic of both complete T-DNAs and those from which *Ac* had excised completely in each line tested. The intensities of signals implied that there were several more copies of intact T-DNA than those from which *Ac* had excised in each case. This situation is similar to that reported by Knapp *et al.* (1988) when using the pKU3 construct in *S. tuberosum*. It is unlikely that this finding results from callus lines being non-clonal, since if this were the case, the Southern data would imply mixed populations of some resistant cells, but mainly sensitive cells within each callus (since there are more intact T-DNAs than excision products). This is

improbable for two reasons. Firstly, kanamycin selection was imposed immediately following inoculation, and secondly, sensitive cells cannot be cross-detoxified by NPT-II enzyme from resistant cells. It is more likely that calli were clonal, but contained multiple inserts of T-DNA; something which is not unusual for transformations using engineered Ti plasmids. In this case, *Ac* excision from only one of several T-DNA copies would be sufficient to restore NPT-II expression and allow the rescue of kanamycin resistant callus. These calli would then show the presence of both types of T-DNA when subjected to Southern analysis. This explanation may also be supported by the NPT-II assays, which showed higher levels of activity in pKU2 extracts than pKU3 extracts. This result would be expected if several copies of the *npt-II* gene were being expressed in pKU2-transformed cells. The fact that only the two predicted bands appeared in Southern blots demonstrates that gross T-DNA rearrangements did not occur, and confirms the implication from the NPT-II assays that in kanamycin resistant pKU3-transformed calli, *Ac* had excised as a complete 4.6 kb element.

Southern blots using the entire *Ac* element as a probe failed to detect any hybridizing bands other than those produced from intact T-DNA. This was unexpected, as in all other species tested, reinsertion was associated with a significant number of excision events (e.g. Baker *et al.*, 1987; Van Sluys *et al.*, 1987; Yoder *et al.*, 1988; Knapp *et al.*, 1988). It may be that in the procedures employed during the course of this work did not produce hybridizations of sufficient sensitivity to detect single copies of *Ac* within the DNAs analysed, or simply not enough *L. usitatissimum* lines were examined to detect reinsertion. Alternatively, *L. usitatissimum* may indeed be unusual in respect to the lack of reinsertion of excised *Ac* elements.

3.4 Summary

An initial long term aim of this work was to develop a suitable system for transposon tagging in *L. usitatissimum* in order to isolate genes such as resistance genes to the rust fungus *Melampsora lini* (see Chapter 1). The work described here demonstrates the excision of *Ac* in transgenic *L. usitatissimum* tissues, but not transposition *per se*, as no elements were detected which reintegrated into the *L. usitatissimum* genome. During the course of these investigations, it also became apparent that transposon tagging strategies in general could be made more feasible if introduced transposons were under greater external control. With this view in mind, it was decided to attempt to use inducible or specifically regulated plant gene promoters to control the transposition of *Ac* and *Ds* in model species such as tobacco and *Arabidopsis*.

CHAPTER 4

THE IDENTIFICATION OF GENES SPECIFICALLY EXPRESSED DURING MICROSPORE DEVELOPMENT.

4.1 Introduction

A requirement of transposon tagging strategies is to produce large numbers of plants, generated after germinal transpositions in the parent plants, with subsequent stabilization of the transposed element. Germinal transposition frequencies of unmodified transposons in most species are low, as discussed in Chapter 1. The use of a gene promoter which could direct specific expression of *Ac* transposase within the gametophyte may be a way of achieving both an increase in germinal transposition of a *Ds* element and the stabilization of the element in vegetative tissues. A line of work was thus undertaken to isolate and characterize putative gametophyte-specific cDNAs from *Brassica napus*, from among known anther-specific cDNAs, and to use these to attempt to isolate the corresponding genes from *Arabidopsis thaliana*, which has a diploid genome with a small DNA content. cDNAs of interest would be identified by a simple differential screen (comparing the hybridization of cDNA clones to RNAs isolated from flower buds of male fertile and male sterile plants) and characterized by more detailed northern and sequence analysis.

The production of mature pollen grains from pollen mother cells is characterized by two distinct phases of gene activity, as described in Chapter 1. The first phase includes meiosis and the end of microspore mitosis, and a later one encompasses the maturation of microspores into pollen grains. Most previous studies on pollen development have used cDNA libraries constructed from mature pollen or mature anther RNA and pollen-specific cDNAs thus isolated are representative of the late phase of gene expression. The current knowledge pertaining to these cDNAs and the few messages expressed during early gametogenesis was discussed in Section 1.9. The work described in this chapter details how the screening of anther-specific clones isolated from a cDNA library representing mRNAs expressed during microspore development (constructed as described in Section 2.4.1 and by Scott *et al.*, 1991a) has led to the identification of putative microspore-specific messenger RNAs that are believed to be of the early phase described above, the period of development representing the most effective time to drive transposition (see discussion in Section 4.3.3). The use of these cDNA clones to screen a genomic library of *Arabidopsis thaliana* resulted in the successful isolation of a putative anther-specific gene homologous to a cDNA termed I3. This cDNA was found to represent a novel putatively microspore-specific oleosin gene.

4.2 Results

4.2.1 The Identification of Microspore-Specific cDNA Clones

A number of anther-specific cDNA clones from *B. napus* were isolated by Scott *et al.* (1991a) as a result of standard differential screening techniques, using labelled first strand cDNA synthesized from anther RNA versus similarly labelled cDNA from seedling RNA. A range of these cDNAs were used to screen northern dot blots of RNAs extracted

Figure 4.1

Diagrammatic representation of important events during in *B. napus* anther development shown in relation to bud length.

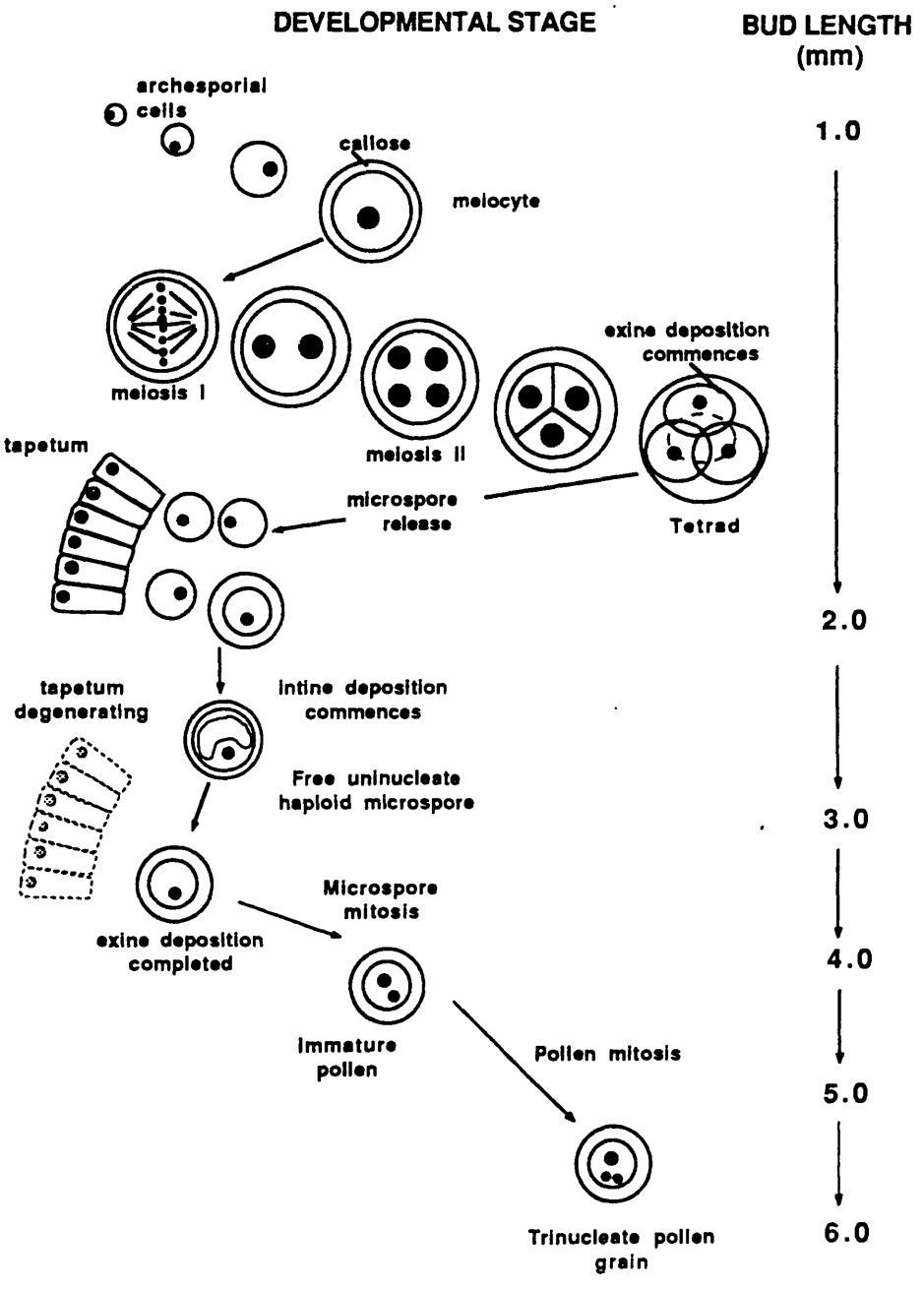

DEVELOPMENTAL STAGE
BUD LENGTH (mm)
archesporial cells
callose
meiocyte
meiosis I
meiosis II
exine deposition commences
tapetum
microspore release
Tetrad
tapetum degenerating
intine deposition commences
Free uninucleate haploid microspore
exine deposition completed
Microspore mitosis
Immature pollen
Pollen mitosis
Trinucleate pollen grain
1.0
2.0
3.0
4.0
5.0
6.0

Figure 4.2

Northern dot blots of the initial screening of putative microspore-specific cDNAs from amongst anther-specific cDNAs.

cDNA probes #17, G2, #14 and I3 were used to challenge 10 µg of RNA from the following *B. napus* tissues: anthers (A), buds (B), microspores and pollen (P), seedlings (S), leaves (L), roots (R).

17

A B P L R S

G2

A B L P R S

14

A L B R P S

I3

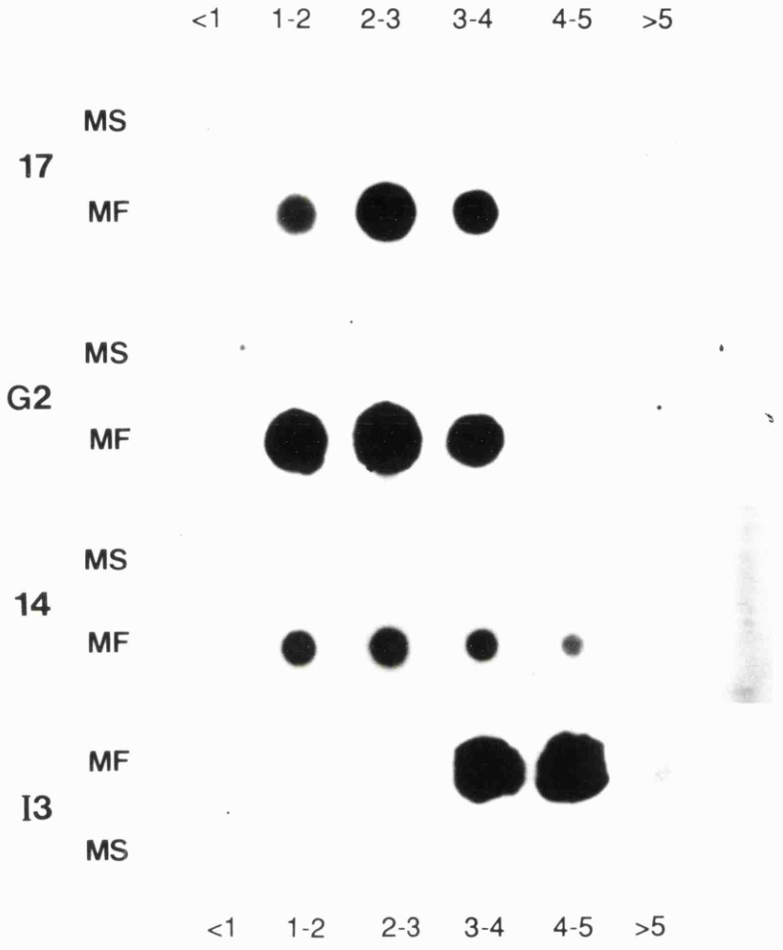
R S A P

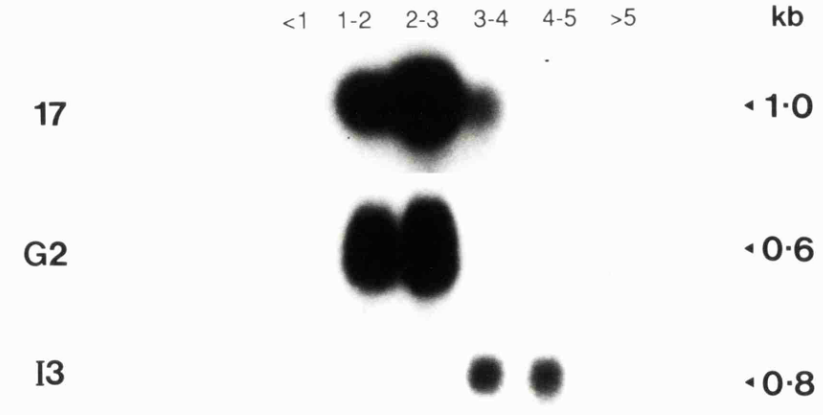
Figure 4.3

Temporal northern dot blots performed on RNA from staged buds of male fertile (MF) and cytoplasmic male sterile (MS) *B. napus*. Bud lengths are indicated in mm. RNA series were probed with the #17, G2, #14 and I3 cDNAs.

Figure 4.4

Northern gel blots of staged RNA from fertile *B. napus* buds, with estimated sizes of the #17, G2 and I3 mRNAs indicated. RNAs were extracted from staged buds of the lengths shown in mm.


<1
1-2
2-3
3-4
4-5
>5
MS
17
MF
MS
G2
MF
MS
14
MF
MF
I3
MS
<1
1-2
2-3
3-4
4-5
>5


<1
1-2
2-3
3-4
4-5
>5
kb
17
1·0
G2
0·6
I3
0·8

from vegetative plant parts (seedlings, leaves, roots) and reproductive tissues (buds, anthers and a mixture of microspores and immature pollen extruded from anthers). A proportion of these showed hybridization to the pollen RNA as well as to anther and/or bud RNA (Fig. 4.2) and were selected for further screening. Northern dot blots were carried out to investigate the temporal expression of the mRNAs represented by these cDNAs in buds from a male fertile variety of *B. napus*, Topaz, and a cytoplasmic male sterile variety, Ogura (Ogura, 1968). It had been shown by Scott *et al.* (1991a) that the cytological stage of gametogenesis within the anthers of a bud of *B. napus* was directly related to the size of that bud. Separating buds into size categories thus provided a simple means of isolating RNAs from different stages of development. Buds were sorted into 1 mm size categories, ranging from less than 1 mm long to greater than 5 mm: the development of the anther and microspores relative to increasing bud size is shown in Fig. 4.1. RNA was extracted from sorted buds from the Topaz and Ogura varieties, blotted onto membrane, and hybridized with labelled cDNA probes. Most of the cDNAs which had shown strong hybridization to microspore RNA, hybridized to temporally expressed mRNAs from the fertile buds, but not to any mRNAs present (at a detectable level) in RNA from the male sterile buds (Fig. 4.3). The major phenotypic characteristic of the Ogura cms variety is the retardation of microspore development and the eventual abortion of pollen grain maturation. It was reasoned, therefore, that anther-specific cDNAs failing to hybridize to RNA from Ogura buds represented putative microspore-specific genes. The cDNAs termed I3, #14, #17, and G2 were included in this class and were characterized further as described below. Northern gel analyses were carried out to confirm temporal specificities and to permit the estimation of the size(s) of mRNA(s) hybridizing to probes from #17, G2 and I3 cDNAs. Hybridizations gave only one distinguishable band in each northern blot, representing mRNAs of approximately 1 kb for #17, 0.6 kb for G2 and 0.8 kb for I3 (Fig. 4.4).

4.2.2 Isolation of a Microspore-Specific Gene

In order to obtain chromosomal DNA fragments containing genes encoding microspore-specific products, a genomic library of *Arabidopsis thaliana* (ecotype Landsberg *erecta*) was constructed in Lambda Dash II, as described in section 2.6. This library was found to contain approximately 200,000 recombinant clones prior to amplification, and was calculated to represent over 99% of the *Arabidopsis* genome using the formula:

$$N = \frac{\ln(1-P)}{\ln(1-f)}$$

where: N = number of recombinant clones,
P = probability of any piece of DNA being included in the library
and f = fraction of the genome per recombinant.

The amplified library was screened by probing plaque lifts containing a total of around 150,000 clones at a time with radio-labelled cDNA probes. This resulted in the isolation of several independent genomic clones hybridizing to each of G2, (λG2.1-λG2.6), #14, (λ14.1-λ14.4), and I3, (λAPG6 and λAPG10). DNA was extracted from these phage clones, and restriction digests and Southern blots carried out in order to map the clones and identify the regions containing the genes of interest.

Bam HI-digested DNA from all λG2 clones possessed a 1 kb restriction fragment which hybridized uniquely to labelled G2 cDNA probe (Fig. 4.5A). This *Bam* HI fragment was thus presumed to include the coding region of the *Arabidopsis* G2 gene and was subcloned. The restriction sites around the G2 gene in *Arabidopsis* which are present in the genomic clones investigated are shown in Fig. 4.5B. The isolated G2 fragment was radiolabelled and used to probe northern dot blots containing RNA from *Arabidopsis* seedlings, buds and flowers, and from *B. napus* seedlings, and buds of 2-3 mm and >5 mm. The probe hybridized to all *Arabidopsis* RNA, most strongly to the seedling RNA. Hybridization to the *B. napus* seedling RNA was also detected, but not as strongly as to *Arabidopsis* seedling RNA, or to *B. napus* bud RNA (Fig. 4.5C).

All four genomic clones isolated using cDNA #14 as a probe had different restriction patterns, and Southern analysis demonstrated that different sized #14-homologous fragments were generated by each restriction enzyme tested (Fig. 4.6A). For each clone, the *Hin* dIII fragment homologous to the #14 cDNA was isolated from an agarose gel and used to probe northern dot blots of RNA from staged *B. napus* buds. Each probe hybridized to RNA present in all bud stages (Fig. 4.6B), not only to the temporally regulated #14 mRNA.

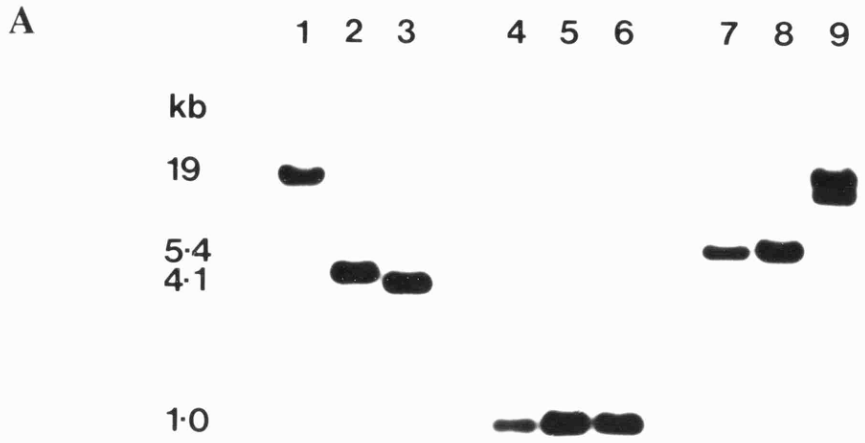
The two genomic clones homologous to the I3 cDNA contained different DNA restriction fragments hybridizing to cDNA probe (Fig. 4.7A). Hybridization was reduced significantly by stringent washing of Southern blots. Restriction mapping determined that the I3-homologous gene was located on the terminal region of the cloned plant DNA in each phage clone, *Eco* RI and *Xba* I restriction sites in the λDash II polylinker being responsible for variance in the sizes of restriction fragments containing the I3 homologue. The restriction maps shown in Fig. 4.7B demonstrate that the λAPG6 clone contained around 2.3 kb more DNA 5' to the I3-homologous gene than λAPG10. The 600 bp *Eco* RI/*Xba* I fragment of λAPG10 (indicated in Fig. 4.7B) hybridized most strongly to I3 and was presumed to include part of the coding region of the gene. This fragment was subcloned and isolated for use as a probe in northern dot blots of RNAs from both *Arabidopsis* and *B. napus.* The results (Fig. 4.7C) showed that this probe hybridized only to RNA from immature flower buds from both species. No hybridization was detected to RNA from *Arabidopsis* open flowers, siliques or whole immature plants. The genomic clones λAPG6 and λAPG10 were thus presumed to contain the *Arabidopsis* gene homologous to the *B. napus* I3 gene. Furthermore, this gene appeared to be regulated in

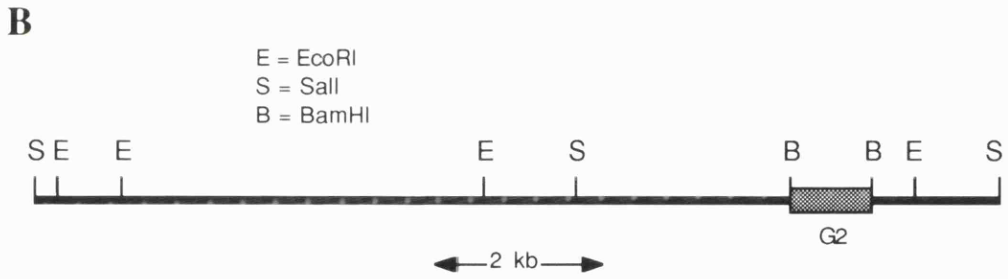
Figure 4.5

A. Southern blot of three G2-homologous *Arabidopsis* genomic clones. Phage DNAs from clones λG2.1, λG2.5 and λG2.6 respectively were digested with *Sal* I + *Eco* RI (Lanes 1-3), *Bam* HI (Lanes 4-6) and *Eco*RI (Lanes 7-9) and the blot probed with G2 cDNA.

B. Restriction map of the region of *Arabidopsis* DNA which includes the G2 gene.

C. Northern dot blot of RNAs probed with the *Arabidopsis* G2 gene (1 kb *Bam* HI fragment).
1; *Arabidopsis* seedling RNA, 2; *Arabidopsis* bud RNA, 3; *Arabidopsis* flower RNA.
4; *B. napus* seedling RNA, 5; *B. napus* 2-3 mm bud RNA, 6; *B. napus* >5 mm bud RNA.


A
1 2 3 4 5 6 7 8 9
kb
19
5·4
4·1
1·0


B
E = EcoRI
S = SalI
B = BamHI
S E E E S B B E S
G2
2 kb

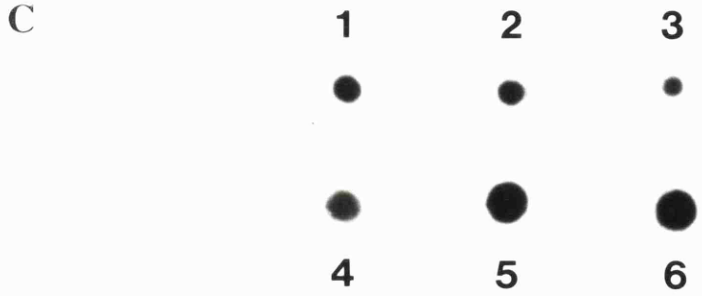
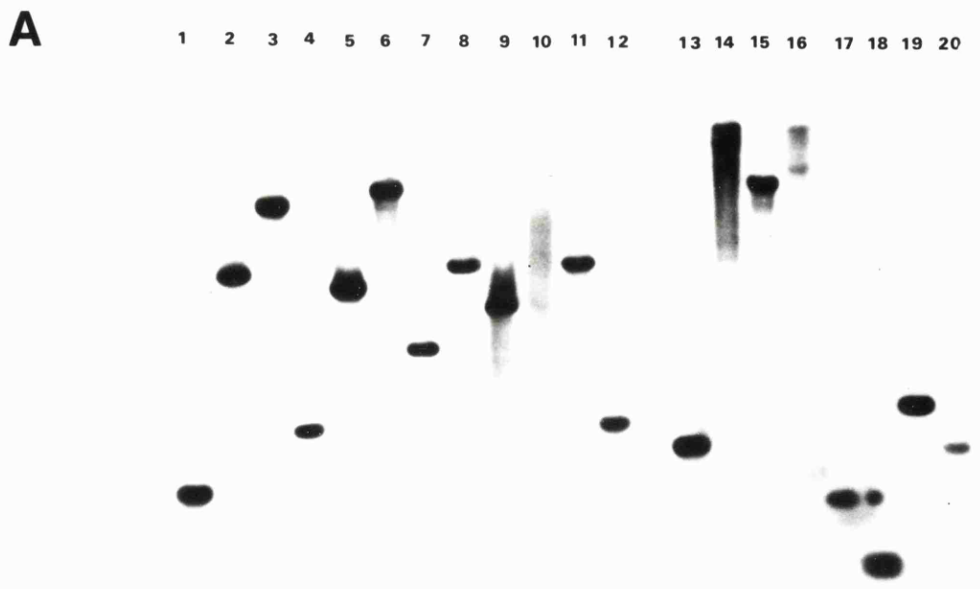

C
1 2 3
4 5 6

Figure 4.6

A. Southern blot of the four #14-homologous *Arabidopsis* genomic clones probed with #14 cDNA.
Lanes 1-4; *Sal* I digests, lanes 5-8; *Eco*RI digests, lanes 9-12, *Xba*I digests, lanes 13-16; *Bam*HI digests, lanes 17-20; *Hin*dIII digests.
Each group of four digests contains phage DNA from genomic clones λ14.1, λ14.2, λ14.3, and λ14.4 respectively.

B. Northern dot blots of RNA from temporally staged *B. napus* buds (lengths indicated in mm).
Blots were probed with #14 cDNA-homologous *Hin*dIII restriction fragments from λ14.1 to λ14.4 respectively.


A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

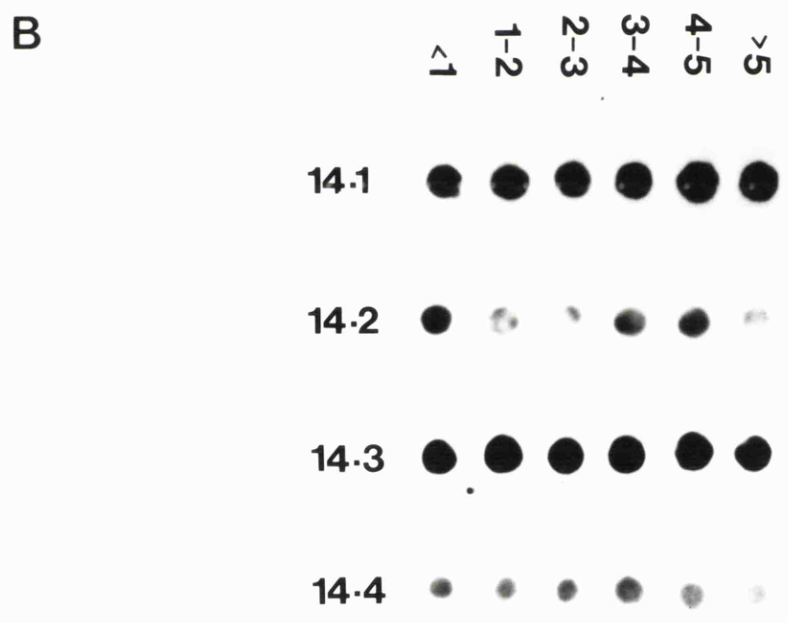
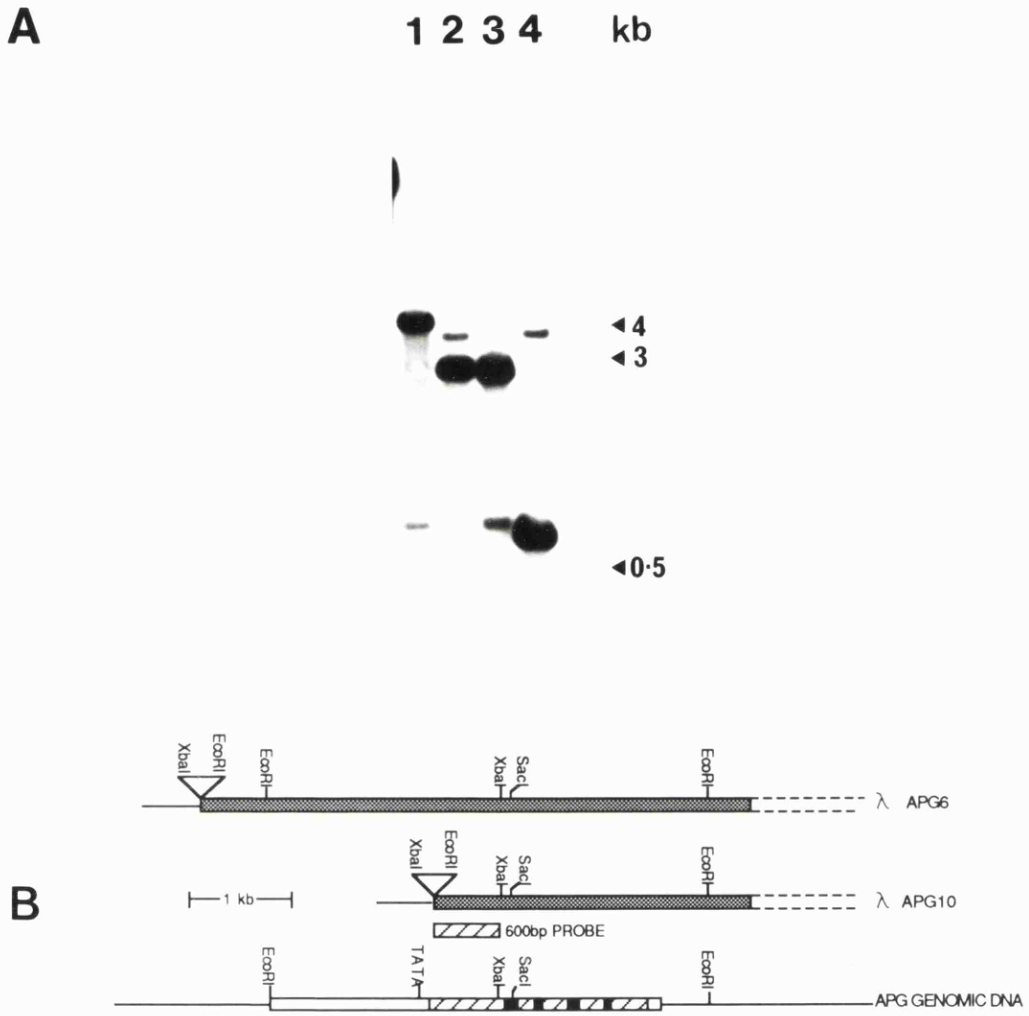

B
<1
1-2
2-3
3-4
4-5
>5
14·1
14·2
14·3
14·4

Figure 4.7

A. Southern blot of the two I3-homologous *Arabidopsis* genomic clones, λAPG6 and λAPG10, probed with I3 cDNA.
Lane 1; λAPG6 x *Eco*RI, lane 2; λAPG6 x *Xba*I,
lane 3; λAPG10 x *Eco*RI, lane 4; λAPG10 x *Xba*I.

B. Restriction map of the genomic clones in the region containing the *Arabidopsis* I3 gene. In the diagram of the APG genomic DNA, the boxed DNA was sequenced (see Chapter 5) allowing the identification of coding (hatched) and inervening (filled) regions.

C. Northern dot blot probed with the 600 bp *Eco*RI/*Xba*I fragment of λAPG10.
Dots contain 10 μg *B. napus* 3-4 mm bud RNA and 1 μg immature *Arabidopsis* plant RNA (I.P), 1 μg *Arabidopsis* bud RNA, 1 μg *Arabidopsis* flower RNA (Flo) and 1 μg *Arabidopsis* silique RNA (Sil).


A
1 2 3 4
kb
4
3
0·5
B
XbaI
EcoRI
EcoRI
XbaI
SacI
EcoRI
λ APG6
1 kb
600bp PROBE
λ APG10
TATA
APG GENOMIC DNA
C
B. napus
3-4
Arabidopsis
I.P.
Bud
Flo
Sil

the same fashion in both species. These clones were therefore chosen for further analysis as a potential source of a microspore-specific promoter.

4.2.3 Determination of the Nucleotide Sequence of the I3 cDNA, and Prediction and Analysis of the Encoded Polypeptide.

Since genomic clones homologous to the I3 cDNA were chosen for analysis, further investigations into the nature of the I3 clone were undertaken. Sequence data can provide information on the possible function of a gene product (by homology to known sequences) as well as aiding the analysis of the organization of genomic clones. The I3 cDNA was sequenced to completion after subcloning and production of single stranded DNA templates. The sequence obtained is shown in Fig. 4.8. I3 was found to contain a 0.75 kb cDNA insert including a poly(A) tail of around 30 nucleotides in length at the 3' end of the clone. The length of the clone is slightly less than the estimated length of the mRNA (0.81 kb) as calculated by comparison to RNA standards of known sizes on northern gel blots. Assuming that the mRNA *in vivo* would possess a longer poly(A) tail and since a full open reading frame was identified, the cDNA was assumed to be near to full length, lacking only the untranslated leader sequence. Computer analysis identified a major potential open reading frame (ORF) the predicted amino acid sequence of which is shown beneath the nucleotide sequence in Fig. 4.8. This ORF, starting at an ATG in position 19, and terminating with a TGA stop codon in position 487, encodes a polypeptide of 156 amino acid residues with a calculated molecular mass of 15.1 kDa. The start methionine of the predicted ORF is in an optimal context for translation, the sequence of TAAACAATG being identical to that reported as a consensus for plant genes by Joshi, (1987). This reading frame is thus thought likely to be that translated from the mRNA *in vivo*. The 3' untranslated region of the I3 mRNA contains a putative polyadenylation signal, AAUAAA (AATAAA in the cDNA sequence) which is typical of eukaryotic mRNAs (Nevins, 1983).

The predicted amino acid sequence of I3 includes several interesting features. The most striking of these is an alanine plus proline-rich C-terminus. This is organized into 8 repeating motifs of two or three alanine (A) residues followed by a proline (P) residue, interspersed only by two repeats of APKP (K=lysine). These repeats are indicated by underlining in the figure. Upon further analysis, a large hydrophobic domain was recognized near the N-terminus of the protein. This extends from amino acids 16-77 inclusive, where all but two residues are significantly hydrophobic. This region can be seen in a hydropathy plot (created using the GCG program "pepplot") as shown in Fig. 4.9 which also indicates that the N-terminus of the predicted I3 protein is highly hydrophilic. Searches of nucleotide sequence databases revealed no similarities between the sequence of I3 and known genes, but when the predicted amino acid sequence of I3 was used to search the NBRF protein database, significant identity with two *Z. mays* oleosins was found. These identities occurred between the central hydrophobic domain of the I3

```
  1  CATTCAAACTAGTAAACAATGGGGATACTCAGGAAGAAAAAACACGAGCGAAACGCGTCG 60
  1                    M  G  I  L  R  K  K  K  H  E  R  N  A  S  14

 61  TTTAAGAGTGTTTTAACCTCAATATTAGCTACACAAGCCGCAACATTCCTCTTGTTGATC 120
 15   F  K  S  V  L  T  S  I  L  A  T  Q  A  A  T  F  L  L  L  I  34

121  TCCGGTGTATCCCTCGCCGGCACAGCCGCCGCATTTATCGCTACCATGCCACTATTCGTA 180
 35   S  G  V  S  L  A  G  T  A  A  A  F  I  A  T  M  P  L  F  V  54

181  GTATTCAGTCCGATTCTCGTACCAGCTGGTATTACCACTGGTTTACTGACTACGGGTTTA 240
 55   V  F  S  P  I  L  V  P  A  G  I  T  T  G  L  L  T  T  G  L  74

241  GCAGCCGCCGGTGGCCGGTGCGACTGCTGTCACCATCATCCTGTGGCTCTACAAGCAAGC 300
 75   A  A  A  G  G  R  C  D  C  C  H  H  H  P  V  A  L  Q  A  S  94

301  AACGGGCAAGGAGCCGCCAGCAGTCCTGTCAAAGTCTTGAAAAAGATAATACCAGGTGCT 360
 95   N  G  Q  G  A  A  S  S  P  V  K  V  L  K  K  I  I  P  G  A  114

361  GCAGCCGCGCCACGAGCCGCACCAGCAGCCGCTCCAGCAGCCGCACCAGCTGCCGCGCCA 420
115   A  A  A  P  R  A  A  P  A  A  A  P  A  A  A  P  A  A  A  P  134

421  GCAGCCGCACCAGCACCTAAGCCCGCAGCCGCACCAGCACCTAAGCCAGCAGCCCCACCG 480
135   A  A  A  P  A  P  K  P  A  A  A  P  A  P  K  P  A  A  P  P  154

481  GCACTATGAAAAGAAGTGGTGGGCATGAGTAAAGGCTGATATGGAAAATTTGATACATGG 540
155   A  L  *

541  AAACAAAAGATTAATCCAACTTTTTAAAAATAAATAACAACTTCACGTGGGGATAGAAAA 600

601  AATTATTAATTCCTACCTAATGGATGTCGTGGTACAAATGTAAGCATATTTATTGTATAG 660

661  CTTTTAAGCGAAGCTGTGTATCGTTGATATTAGTCCTTATTCAAGTTATTCAGGTTCCTC 720

721  TTTA(n)
```

Figure 4.8

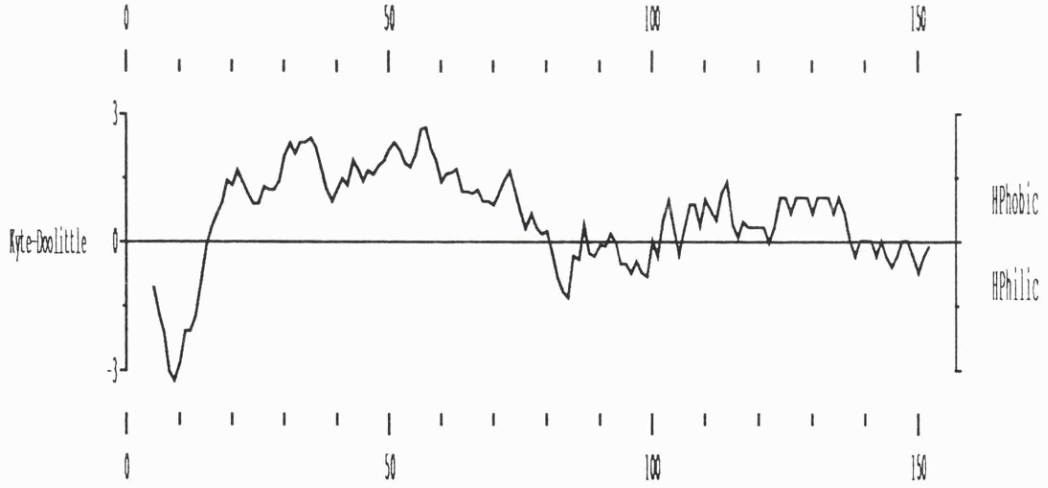
The nucleotide and deduced amino acid sequence of the I3 cDNA. The repeated regions in the C-terminus of the polypeptide which are discussed in the text are underlined, as is the putative polyadenylation signal in the nucleotide sequence.

Figure 4.9

Hydropathy plot of the predicted I3 protein, calculated using the algorithm of Kyte and Doolittle (1982) with a window of 9 amino acid residues. Hydrophobic regions are plotted above the horizontal axis and hydrophilic regions below.

PEPPLOT of: I3.Pep ck: 3059, 1 to 156 March 3, 1992 10:18

Hydropathy Plot of the Predicted I3 Polypeptide.



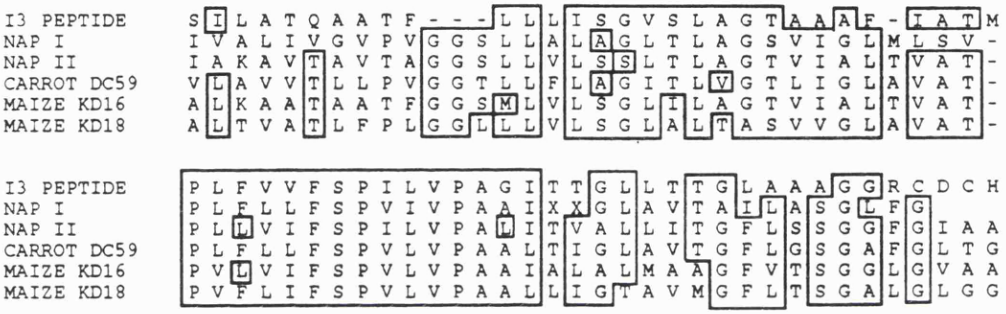
polypeptide and the central regions of the oleosin sequences. Further searches of the literature revealed that at least four other oleosin proteins had been fully or partially sequenced, and alignments (presented in Fig. 4.10) demonstrated homology between all the oleosin sequences and that of I3 in their central domains.

The oleosins are 16 - 26 kD proteins, found in seeds, which occur in the outer layer of lipid storage organelles, the oil bodies. The seeds of a wide range of both monocotyledonous and dicotyledonous plant species contain at least two major oleosin isoforms, which are termed the low and high molecular weight isoforms. These two classes are immunologically distinct, but antibodies raised to either isoform class recognize oleosins of the same class from diverse sources (Tzen *et al.*, 1990). All isolated oleosins exhibit structural characteristics which point to a conserved mechanism of association with the oil body and it's surrounding half unit membrane (Murphy *et al.*, 1991). The proteins comprise an alpha-helical hydrophilic N-terminal domain, which is thought to be associated with the cytoplasm, a central hydrophobic beta-sheet domain, assumed to penetrate the lipid body, and an amphipathic C-terminal alpha-helix which is postulated to interact with both the cytoplasmic and lipid compartments. Computer predictions of secondary structure of the I3 protein show that it conforms to all of these rules, except that it was not possible to predict the presence or absence of amphipathic structures. Thus it appears likely that the I3 gene encodes a novel low molecular weight oleosin protein, which presumably plays a role in the production of lipid storage bodies in the developing pollen grain.

Figure 4.10

Alignment of the amino acid sequences of the conserved central domain of known oleosins with the predicted I3 polypeptide (amino acids 22-85). Regions of conserved amino acid residues are boxed.

Sequences shown are the hydrophobic domains of each of the *B. napus* Nap I and Nap II (a.a. 55-121) oleosins (Murphy *et al.*, 1991), the *D. carota* DC 59 (a.a. 40-106) oleosin (Hatzopoulos *et al.*, 1990) and the *Z. mays* 16 kD (Vance and Huang, 1987) and 18 kD (a.a. 48-114) (Qu and Huang, 1990) oleosins.



HOMOLOGIES:

I3 : NAP I	49%	IDENTICAL,	60%	FUNCTIONALLY	CONSERVED
I3 : NAP II	40%	"	54%	"	"
I3 : DC59	40%	"	57%	"	"
I3 : MAIZE 16 KD	48%	"	62%	"	"
I3 : MAIZE 18 KD	30%	"	54%	"	"

4.3 Discussion

4.3.1 Temporal and Spatial Expression of Anther-Specific cDNAs

Of a number of *B. napus* anther-specific cDNAs isolated by differential screening, four were identified by northern analysis as being expressed in immature pollen but not cytoplasmic male sterile buds. The Ogura cms phenotype was characterized cytologically (Ogura 1968) and sterility was found to be the result of the abortion of pollen grain production during the later phases of development. It was reasoned that genes expressed in the microspores of fertile anthers, but not in cms anthers, were amongst those inactivated as a result of the lesion in Ogura pollen. The cDNAs #14, #17, G2 and I3 were consequently assumed to represent putative microspore-specific genes, since they hybridized to anther and pollen RNA only from fertile flower buds.

The temporal expression of these four genes fails into two main patterns. The first group of cDNAs, including #14, #17 and G2, hybridized maximally to RNA extracted from fertile flower buds of 2-3 mm in length. Significant leveis of hybridization were also observed to RNA from buds of 1-2 and 3-4 mm in length. Buds of these slzes include anthers at stages of gametogenesis ranging from microsporogenesis (meiosis) to microspores undergoing their first mitotic division (see Fig. 4.1). 2-3 mm buds contain free microspores in the interphase before this mitosis. The I3 cDNA demonstrated a different pattern, hybridization being maximal in buds of 3-4 mm, with high levels also being observed in 4-5 mm buds. Low levels of hybridization were sometimes detected in 2-3 mm and >5 mm buds. Buds of 3-5 mm include microspores in the interphase of the microspore mitosis through to those about to enter the polien mitosis. Whilst different hybridization maxima are important characteristics in revealing peaks of gene expression, it cannot be similarly assumed where low levels of hybridization occur that this relates to low levels of expression at that particular developmental stage. This is because size-sorted bud populations are arbitrarily categorized and will obviously contain anthers at overlapping stages of development. For example, based on the high level of expression in 3-5 mm buds, it is probable that the low level of mRNA detectable in >5 mm buds with I3 probe is derived from pollen grains which have not yet undergone their second mitotic division rather than mature pollen grains, as would initially be implied.

4.3.2 Cloning a Microspore-Specific Gene from *Arabidopsis*

B. napus is an amphidiploid species, being a hybrid between *B. oleracea*, (cabbage) and *B. campestris*, (turnip). The *B. napus* genome is therefore likely to include more than one copy of genes which are conserved within the Brassicaceae. The diploid species *Arabidopsis thaliana*, also a member of the Brassicaceae, was chosen as the species from which to cione genes, in order to avoid problems arising from gene multiplicity in *B. napus*. In addition, *Arabidopsis thaliana* has the advantages that it has a small genome which contains little repetitive DNA, is relatively easy to transform, and is amenable to

genetic studies; especially because of a rapid generation time (seed to seed in 6-8 weeks). For these reasons *Arabidopsis* is being developed as a model system for plant biology (see Meyerowitz, 1989, for a review).

The aim of cloning *Arabidopsis* genes homologous to the cDNAs characterized was to obtain regulatory regions providing microspore-specific expression. These would then be used to drive microspore-specific transposition of *Ac* and *Ds* elements. It is believed that the majority of transposition events occur during DNA replication (see Chapter 1) and promoters which were active prior to, or during, DNA replication would thus be most likely to facilitate maximum germinal transposition. All four cDNA clones examined demonstrated maximal expression prior to, (#14, #17, G2), or during, (I3), the microspore mitosis, and the genes encoding these mRNAs were considered suitable targets for the proposed strategy.

Genomic library screening resulted in the isoiation of clones hybridizing to cDNAs #14, G2 and I3. Of these clones, those homologous to the G2 cDNA represented a single gene, the presumed coding region of which was located on a 1 kb *Bam* HI restriction fragment. When used to probe RNA from *B. napus* and *Arabidopsis*, this fragment either did not hybridize to the same message as G2, or additionally hybridized to other messages present in vegetative tissues. The four clones hybridizing to #14 all appeared to inciude different genes, which did not hybridize uniquely to temporally expressed mRNAs in *B. napus* buds. It was therefore supposed that the #14 cDNA hybridized to a gene family in *Arabidopsis*, the members of which either hybridized to one or more mRNAs not expressed in the same way as the #14 gene in *B. napus*, or did not produce a microspore-specific transcript that was easily distinguishable from those expressed from other members of the family.

In contrast, the clones λAPG6 and λAPG10 were found to contain an I3-homologous sequence which did appear to mirror the hybridization characteristics of the I3 cDNA in *B. napus*. The genomic DNA used as a probe was found to hybridize only to *Arabidopsis* mRNA extracted from unopen flower buds; not to RNA from open flowers, implying that the homologous mRNA was present only during the earlier stages of flower development. The probe also hybridized strongly to 3-4 mm bud RNA from *B. napus*, the stage when I3 is maximally expressed in this species. The mRNA to which the I3-homologous probe hybridized was therefore expressed at a similar stage of development in both species. The cloned gene was presumed to be homologous to the I3 cDNA, both in terms of nucleotide sequence and regulation of expression. The gene, termed APG, was therefore selected for further expression and sequence analysis, with the aim of identifying the regions of the gene conferring tissue and temporal specificity.

4.3.3 Analysis of the Sequence of the I3 cDNA

Sequencing of the I3 cDNA demonstrated that the I3 gene encodes a 15.1 kDa protein with an unusual repeated sequence at the C-terminus of the polypeptide. Comparison of the

deduced protein sequence of I3 with sequences present in the NBRF protein sequence database indicated that the gene encodes a novel type of oleosin. This initial belief, based on amino acid sequence identity with other isolated oleosins, was strengthened by comparisons of likely secondary structure with those of the known proteins. The I3 protein includes an alpha-helical hydrophilic N-terminal domain and a central hydrophobic beta-sheet domain, features which are both typical of oleosins. Similarities were less clear, however, in the third, C-terminal domain.

Seed lipid bodies are of value economically as a source of important oils, and their composition in *B. napus* has been studied in comparison to polien lipids, with a view to enabling the selection of varieties with desirable fatty acid compositions (e.g. Evans *et al.*, 1987; 1990a). Such studies show that lipids comprise 22% of the dry weight of the internal pollen compartment (Evans *et al.*, 1987). It has been suggested that different biochemical pathways contribute to the accumulation of storage lipids, which are synthesized by the gametophyte, and structural (membrane and pollen wall) lipids which are synthesized mainly by the sporophyte (Evans *et al.*, 1990a). One might then expect that different oleosin genes may be expressed in pollen and seeds, as is suggested by the expression patterns of both I3 and the known seed oleosin genes, which are active only during embryogenesis. Additionally, I3 expression is maximal at the time during *B. napus* anther development when pollen lipids are beginning to accumulate (Evans *et al.*, 1990b) and this correlation gives further support to the contention that the I3 protein is an oleosin specific to pollen oil bodies.

4.4 Summary

Analysis of a group of anther-specific cDNAs demonstrated that four of these were expressed in the immature pollen, or microspores, of fertile *B. napus*, but not in the anthers of male sterile plants. Two different patterns of expression were observed for these cDNAs when analysed in more detail, three being expressed at an earlier developmental stage than the fourth, though all four clones represented mRNAs which show maximal expression around the time of the microspore mitosis. One of the cDNAs, I3, encodes a product which appears to be a novel pollen-specific oleosin, a class of proteins which occur in oil body membranes. An *A. thaliana* genomic library yielded a number of clones hybridizing to the putative microspore-specific cDNAs, though only sequences homologous to the I3 cDNA exhibited the same patterns of hybridization to RNA populations from vegetative and reproductive tissues. The gene included in these sequences was named APG, and was presumed to be a microspore-specific gene in *Arabidopsis*. The characterization of this gene and applications arising from these studies will be discussed in the following chapters.

CHAPTER 5

STRUCTURE AND EXPRESSION OF THE APG GENE OF *ARABIDOPSIS THALIANA* AND HOMOLOGOUS mRNAs FROM *BRASSICA NAPUS*.

5.1 Introduction

The *Arabidopsis* APG gene was isolated by homology to the I3 cDNA of *B. napus*. The I3 message encodes an oleosin, and is strongly expressed in microspores around the time of the microspore mitosis, but not in mature pollen. The I3 gene is thus representative of the first of the two recognized classes of microspore-expressed genes discussed in Chapter 1. Activity of the I3 gene is localized to anthers undergoing microspore development and is believed to be specific to the microspore itself. The cloned APG gene has been shown to be homologous to a transcript found only in immature *Arabidopsis* flower buds, which contain developing microspores. The same transcript is not detected in open flowers, which contain mature pollen. APG is thus also believed to be a member of the microspore development-specific class of genes, and further work will be described which confirms this hypothesis. Controlling regions responsive to transcription factors present during this phase of development have been implicated as good candidates to act as sequences promoting high germinal excision frequencies when linked to a transposase gene (see Section 4.3.3). A major aim of the characterization of APG clones was to identify, by nucleotide sequence analysis, the upstream controlling region of the gene and subsequently to test the promoter activities of this region by reporter gene fusions in transgenic plants.

This chapter also describes the analysis of a number of other cDNAs of *B. napus* which show a high degree of homology to APG.

5.2 Results

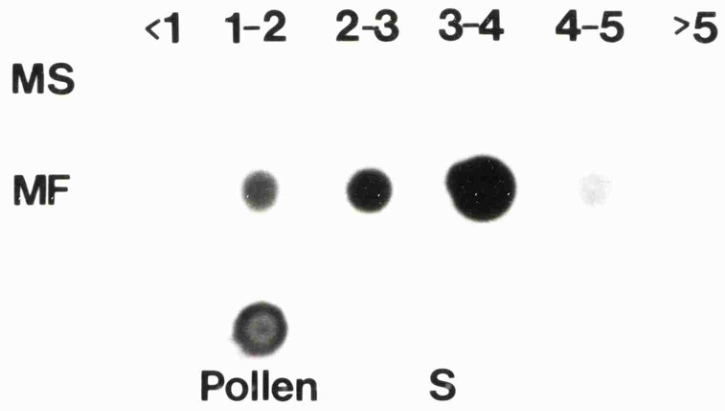
5.2.1 Analysis of APG Expression

The fragment which most strongly hybridized to the I3 cDNA in Southern blots was the 600 bp *Eco* RI / *Xba* I fragment of λAPG10, and this was therefore assumed to contain part of the APG gene coding region. This fragment had previously been shown to hybridize to an mRNA expressed specifically in immature buds of *Arabidopsis* (4.2.3) and was therefore used as a probe in further northern blot experiments carried out to analyse expression of this gene in more detail. Staged bud RNA samples from fertile and cms *B. napus* were probed in order to compare the temporal expression of any hybridizing mRNA to the expression of the I3 mRNA which had already been determined (Section 4.2.1). The results presented in Fig. 5.1 show that the 600 bp probe hybridized strongly to RNA from fertile buds between 2 mm and 4 mm in length, maximally in the 3 mm - 4 mm category. Lower levels of hybridization were detected to 1 mm - 2 mm and 4 mm - 5 mm bud RNAs, though no signal was detected in seedling RNA or in bud RNAs of the Ogura cms variety. Additionally, the probe also hybridized to *B. napus* immature pollen RNA.

These data implied that like the I3 cDNA, the APG genomic fragment hybridizes to a transcript expressed only in fertile flower buds which is also present in immature

Figure 5.1

Northern dot blot of staged *B. napus* bud RNAs from male fertile (MF) and cytoplasmic male sterile (MS) plants probed with the λAPG10 *Eco*RI/*Xba*I 600 bp fragment. Bud lengths in mm are indicated. Also included are *B. napus* extruded immature pollen and seedling (S) RNAs. All dots contain 10 μg of total RNA.


<1
1-2
2-3
3-4
4-5
>5
MS
MF
Pollen
S

pollen. However, the temporal patterns displayed by the messages homologous to the I3 and APG probes were not identical. The I3 mRNA is expressed strongly in 3 mm - 5 mm buds, whereas the APG-hybridizing mRNA appeared to be most strongly expressed in 2 mm - 4 mm buds. Furthermore, I3 mRNA is detectable in buds over 5 mm whilst the APG homologue is not. Conversely APG message it is present in 1 mm - 2 mm buds whilst I3 message levels are negligible before buds reach 3 mm. This evidence indicated that the APG and I3 probes were hybridizing to different mRNAs, and the APG gene may not be as closely related to I3 as it is to another *B. napus* microspore-specific gene. Further analyses were then carried out to assess the similarity between APG, I3, and other *B. napus* messages.

5.2.2 Sequence Analysis of the APG Gene

Two initial subclones in pSK- from the λDash II genomic clones were obtained, these being the 2.9 kb *Xba* I fragment from the 5' end of λAPG6, and the 2.8 kb *Eco* RI fragment from the 5' end of λAPG10 (see Fig. 4.7). Using single stranded templates from these subclones and further subclones subsequently produced, the APG gene (including the region used in northern analyses) was sequenced to completion (Fig. 5.2). A major open reading frame in the nucleotide sequence of the APG genomic clones was identified, beginning at an ATG in position 1 in Fig. 5.2. The coding region of APG is interrupted by four small introns of 164 bp, 89 bp, 115 bp and 88 bp respectively, before a termination codon in position 2064, followed by a putative polyadenylation signal in position 2095. The presence of these introns was confirmed by the sequencing of homologous cDNAs, as described below. The ATG shown as the start site in Fig. 5.2 is in near optimal context for translation, compared to the consensus for plant genes determined by Joshi (1987) of TAAACAATG, whilst the presence of a putative TATA box in position -85 to -78 lends further support to the assumption that the deduced start of translation is that used *in vivo*. The APG gene encodes a predicted polypeptide of 534 amino acid residues, with a calculated molecular mass of 57.9 kDa. One important feature of this protein is a substantial proline-rich region, containing 55% proline in 156 amino acids. Besides proline, the amino acids lysine, alanine and serine are also predominant within this region. The region contains 4 repeated sequences of 13, 15, 11 and 5 amino acids, which are shown in Fig. 5.3. A hydropathy plot of the APG peptide sequence (Fig. 5.4) revealed a hydrophobic N-terminal region, representing a potential signal sequence of the type involved in targeting the ribosome/propeptide complex to the endoplasmic reticulum. This sequence obeys the rules of Von Heijne (1983) for the typical amino acid composition of a signal. Proteins are secreted from the cell via this pathway, or may be further modified before transport to intra- or intercellular compartments.

Only a small region of 106 bases (between nucleotides 393 and 498 on the sequence

Figure 5.2

Nucleotide and deduced amino acid sequence of the *Arabidopsis* APG gene. Underlined are the presumptive TATA box and polyadenylation signal, and the intron donor and acceptor sites.

```
-1501  GAATTCTTGTATGACATAGATACATATACCTCTTTTATGGCTAGCTCTGCTGCAAGTTTTGCTATGGTACGTGAAACCAA
-1421 TCCTCCGATTTAGTTGTTTATAATATGAACTAGTATATCAAACCTACTGAATATGAGTTTAATTTTGGCAGCAACTATATG
-1340 AATCAGGAGCAAAAAAAAATAGGATTTATTGGTGTTTCTCCGATCGGGTGTATACCGATACAGAGAACCACAAGAGGAGGAC
-1259 TTAAAAGAAAATGTGCTGATGAACTCAACTTTGCAGCTCAGCTTTTTAATTCCAAACTCTCCACAAGTTTGAACGAATTGG
-1178 CTAAAACCATGAAAAATACCACATTGGTGTATATTGACATCTACTCTTCTTTCAATGATATGATTCAAACCCAAAAATATG
-1097 GTAACGCTCAACTACTCACTCATCATGGTGTTTTGTGAGAATAAGTTGAGATTTTTTTAAAATTCTTTTTTGGTTTTTTGC
-1016 AGGATTTGATGAGATTGATAGAGGATGTTGCGGGACGGGGCTACTCGAATTAGGTCCGCTTTGCAACAAATATACATCACT
 -935 TCTTTGCAAGAATGTGTCTTCTTTTATGTTTTGGGACTCTTACCATCCAACGGAAAGACTTATAAAATCTTAAGCCAAAAG
 -854 TTTGTAGAGAATGACATGGGCCCTTTCTATGACAATTAAATGTAAAATTTAAGATGACACATTTGAGATGAAATTTGGCTA
 -773 AATAAAAGCTGTATCTCATTTGGATTAGTGTCTTCAAAAACATGTTTATTGTATTATACATATAAATTGACATGGAGCGAT
 -692 AAATTGAGTTGTTTTTCTAAAGATTTGATATTATTGTGCTCTGACAATAATCGGACAGACAAAAGTGGACGACTAAAACCA
 -611 ATAATAATAAAAACAGAGTGTAAAATGCTATTATATTATTGTGAATTTAGCCGAATTGACACACCTTCTTCTTGTTTCACTT
 -530 GATCGAATCCATCTCATTCCAACAAAATATTAACTAAATATAATTTAAATCATGGCCATTAACGAAAGTTTAAAAAGAAGG
 -449 CCAATCCGCTAGAAACTTCTCACCCTCAGACACAGTCGCACAGGCATCAGCTTTGAAATCGCGTAACAAAATTTTGTGGGA
 -368 AGCTATTACATAGACATAAACCCTAACTCCAAATCCAACTGTAGTTACTTTGATAAACTTTTCTGGTTAAGTCAAAGATAG
 -287 GGAAGATTTAATAGAGTTTATTCCTCATGCGTTTCTATGGATTGAGAAGTCACGAAAGATGCTAGAAAGAGTACTTAAGTT
 -206 CCTACTTGAAACTGGAGACCGAATATAAGGGAAGTCTTAAAACGTGTGATGGAAGAAGCCTTGAGACACAGGTAACTTGAA
 -125 CCAGGCCACATTAATTGCATGTCCTAATCTACGTAATGAATATATATACGTACCAAAGCAGATATTGACATGACAGTGGCA
  -44 AACACTTGAATTAAAACTGTTTTTTTACTACTTGTAGTAAAACCATGAAGCGATCATCCTTGGTCGATTCTTGTTCATATT
    1                                                  M  K  R  S  S  L  V  D  S  C  S  Y
   38 CGAGAATATTTCGCAGTATTTTCTGTCTCTTATCCTTTTGTATCTTCTTCTTGACAACGACCAATGCTCAGGTAATGCATA
   13 S  R  I  F  R  S  I  F  C  L  L  S  F  C  I  F  F  L  T  T  T  N  A  Q  V  M  H
  119 GACGGCTCTGGCCTTGGCCATTGTGGCCACGGCCGTATCCACAGCCGTGGCCAATGAATCCACCAACACCTGACCCATCAC
   40 R  R  L  W  P  W  P  L  W  P  R  P  Y  P  Q  P  W  P  M  N  P  P  T  P  D  P  S
  200 CAAAGCCTGTCGCACCACCTGGCCCATCACCAAAACCTGTCGCACCACCAGGCCCATCTCCATGCCCGTCACCACCACCAA
   67 P  K  P  V  A  P  P  G  P  S  P  K  P  V  A  P  P  G  P  S  P  C  P  S  P  P  P
  281 AGCCTCAACCCAAGCCCCCACCAGCACCTAGCCCATCTCCATGCCCATCACCACCACCTAAGCCTCAACCAAAGCCAGTGC
   94 K  P  Q  P  K  P  P  P  A  P  S  P  S  P  C  P  S  P  P  P  K  P  Q  P  K  P  V
  362 CACCACCAGCATGCCCACCAACACCTCCTAAACCTCAACCGAAGCCAGCACCACCGCCAGCACCAAAGCCAGCGCCACCAC
  121 P  P  P  A  C  P  P  T  P  P  K  P  Q  P  K  P  A  P  P  P  A  P  K  P  A  P  P
  443 CAGCACCAAAGCCAGTCCCATGTCCATCACCACCTAAGCCACCAGCCCCAACTCCAAAGCCTGTCCCACCACATGGTCCAC
  148 P  A  P  K  P  V  P  C  P  S  P  P  K  P  P  A  P  T  P  K  P  V  P  P  H  G  P
  524 CTCCAAAACCTGCTCCAGCACCAACGCCTGCCCCTTCCCCAAAACCTGCCCCATCACCACCTAAGCCAGAGAACAAAACTA
  175 P  P  K  P  A  P  A  P  T  P  A  P  S  P  K  P  A  P  S  P  P  K  P  E  N  K  T
  605 TACCGGCCGTGTTTTTCTTTGGCGATTCAGTCTTTGACACAGGAAATAACAATAATCTAGAAACAAAGATAAAAAGTAATT
  202 I  P  A  V  F  F  F  G  D  S  V  F  D  T  G  N  N  N  N  L  E  T  K  I  K  S  N
  686 ATCGTCCCTATGGTATGGACTTTAAATTCAGAGTTGCAACTGGTAGATTCAGCAATGGAATGGTTGCTTCCGATTATTTAG
  229 Y  R  P  Y  G  M  D  F  K  F  R  V  A  T  G  R  F  S  N  G  M  V  A  S  D  Y  L
  767 GTACGTACTCTCTCTCCCTCTCTCCATTCTTTTCTTTTGAGCTCGACCTTGGCCCAATAGATTGACATTGACAATTAATTA
  848 CAAATTTTCAACATACTGAAAAAAAAGGAAAACCAATTACAAATTCTTCTAATTATATTTTTGTTTGGTATCCTAATTCCAA
  929 TAGCCAAATACATGGGAGTAAAAGAAATTGTACCGGCATATTTAGACCCGAAAATACAACCAAACGATCTTCTTACGGGAG
  256    A  K  Y  M  G  V  K  E  I  V  P  A  Y  L  D  P  K  I  Q  P  N  D  L  L  T  G
 1010 TATCTTTTGCATCGGGTGGTGCTGGCTACAATCCTACTACATCCGAAGCAGCGGTAAGTTCGAATCACTTTTATCATCATT
  282 V  S  F  A  S  G  G  A  G  Y  N  P  T  T  S  E  A  A
 1091 ATTCTTATATGGAGTAATTTATATTTGCAATCAGAATCATCATTTGATTTATATAACTCAGAATGCAATCCCAATGCTGGA
  300                                                                 N  A  I  P  M  L  D
 1172 CCAACTGACATATTTCCAAGACTATATAGAGAAAGTAAATAGGTTAGTAAGACAACACAAGAGCCAATATAAGTTGGCAGG
  307    Q  L  T  Y  F  Q  D  Y  I  E  K  V  N  R  L  V  R  Q  H  K  S  Q  Y  K  L  A  G
 1253 TTTAGAGAAGACCAATCAGCTAATATCCAAAGGAGTGGCAATTGTCGTCGGGGGAAGCAATGATCTGATCATTACATATTT
  334    L  E  K  T  N  Q  L  I  S  K  G  V  A  I  V  V  G  G  S  N  D  L  I  I  T  Y  F
 1334 TGGAAGTGGTGCTCAACGACTCAAAAATGACATCGACTCTTATACCACTATCATTGCTGATTCTGCCGCCAGTTTCGTTTT
  361    G  S  G  A  Q  R  L  K  N  D  I  D  S  Y  T  T  I  I  A  D  S  A  A  S  F  V  L
 1415 GGTACGTGATTATTATATATAGCTTTTAAAACCTCCTTTTTTTTTTTTTTGCTTTTCCAAATTCCACAAACATGATTGAATT
 1496 TCCTTAAATTAATGTTAATGTAAATGTATAAATTGGGCAGCAATTGTATGGATATGGAGCAAGACGTATAGGAGTGATCGG
  388                                                  Q  L  Y  G  Y  G  A  R  R  I  G  V  I  G
 1577 AACACCACCACTTGGATGTGTACCATCACAAAGACTAAAGAAAAAGAAAATATGCAATGAAGAGCTAAACTATGCTTCTCA
  402    T  P  P  L  G  C  V  P  S  Q  R  L  K  K  K  K  I  C  N  E  E  L  N  Y  A  S  Q
 1658 GCTTTTCAATTCTAAGCTCCTCCTTATTTTGGGTCAATTGTCAAAAACCCTACCAAATTCTACATTCGTTTATATGGACAT
  429    L  F  N  S  K  L  L  L  I  L  G  Q  L  S  K  T  L  P  N  S  T  F  V  Y  M  D  I
 1739 CTACACTATCATCAGTCAGATGCTAGAAACTCCTGCGGCTTATGGTAAATTATTTATTGTGACAATTCATATAACACTTTG
  456    Y  T  I  I  S  Q  M  L  E  T  P  A  A  Y
 1820 TTATTTATAATATCTGTTCATGCATTTGAATTTTAAAATATTTGCATGCAGGATTTGAGGAGACAAAGAAACCGTGTTGCA
  470                                                        G  F  E  E  T  K  K  P  C  C
 1901 AAACAGGATTATTAAGCGCAGGTGCCCTCTGCAAGAAGTCTACATCGAAAATCTGTCCCAATACATCGTCTTACTTGTTTT
  480 K  T  G  L  L  S  A  G  A  L  C  K  K  S  T  S  K  I  C  P  N  T  S  S  Y  L  F
 1982 GGGACGGTGTGCATCCTACTCAAAGAGCTTATAAAACTATAAACAAAGTACTTATAAAAGAATACCTTCATGTTTTATCTA
  507 W  D  G  V  H  P  T  Q  R  A  Y  K  T  I  N  K  V  L  I  K  E  Y  L  H  V  L  S
 2063 AATAATAGTATTACATTATTTATCTCATGCGAAAATAAAGACTTTTGCTATGTACTAATGATGTAGAAGATGTGGGCCCGT
  534 K  *
 2144 TTCTGTGACTAAAAATGTAGTTAAGCAAACAACATTTTGAGATAGTAATGTAAGATTTTTCTTTTGTGTCAAACGAATCCC
 2225 CATCGAGTCATGTTTGTATTAGTTTATGATCATGTTTGTGGTGTATAATGGCTAAATAAATTTGATCTGTACGTAATTTCT
 2306 TAATTCTTGATATTTGTATGAGAAATTCAGTTACTTTTCTAGGATTTATATTTGATGTGTTAAGAAAAACTAAAGCGGTGA
 2387 GTTTTAGACTATAATCTCGTTTTATTATGATAGATGAGTGCGAC
```

Figure 5.3

Diagram showing the repeated elements within the proline-rich region of the APG protein.

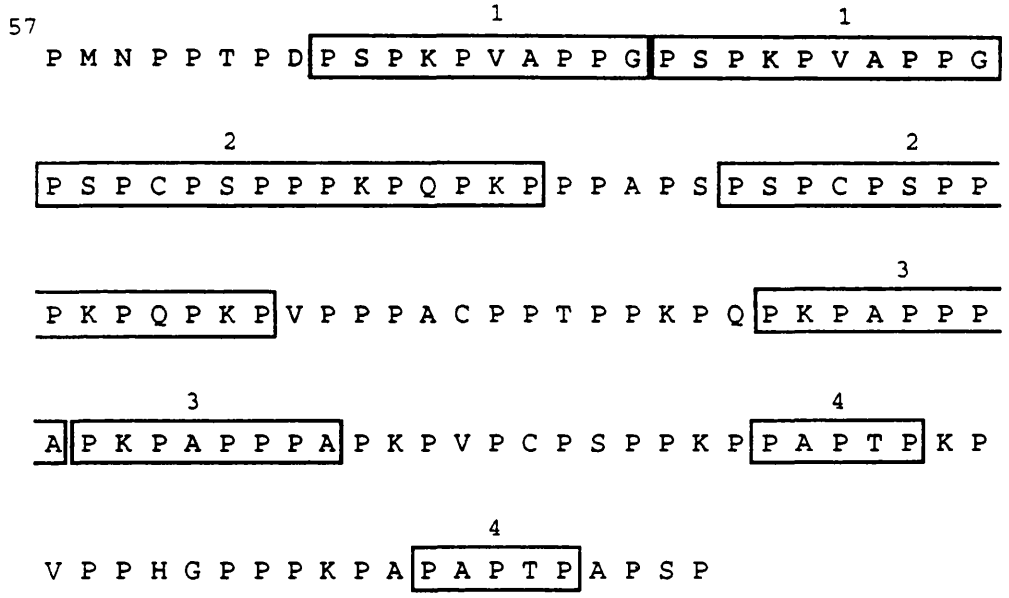
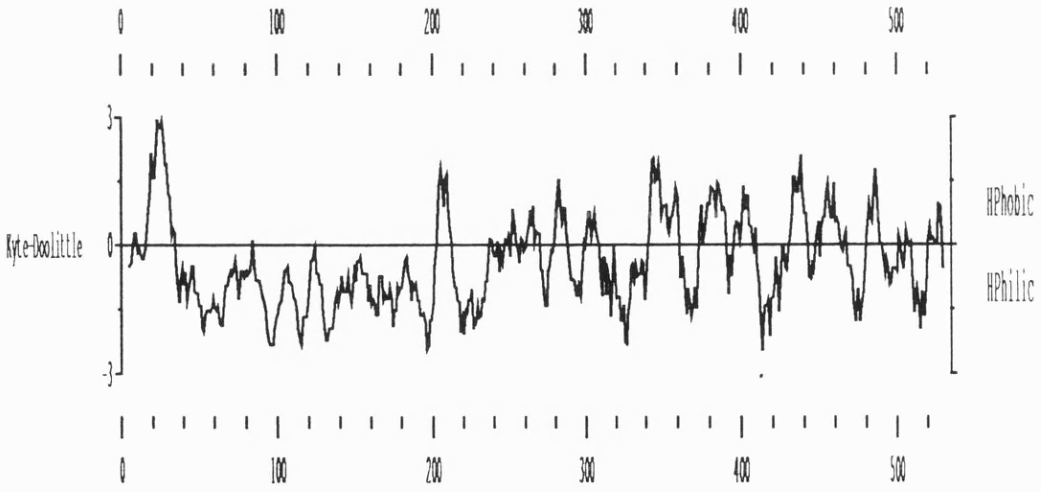

57
1
1
P M N P P T P D P S P K P V A P P G P S P K P V A P P G
2
2
P S P C P S P P P K P Q P K P P P A P S P S P C P S P P
3
P K P Q P K P V P P P A C P P T P P K P Q P K P A P P P
3
4
A P K P A P P P A P K P V P C P S P P K P P A P T P K P
4
V P P H G P P P K P A P A P T P A P S P

Figure 5.4

Hydropathy plot of the predicted APG polypeptide, calculated using the algorithm of Kyte and Doolittle (1982) with a window of 9 amino acids. Hydrophobic regions are plotted above the horizontal axis, and hydrophilic regions below.

PEPPLOT of: Apg.Pep ck: 9743, 1 to 534 March 3, 1992 10:25

Hydropathy of the Predicted Polypeptide of APG.



shown) demonstrated any degree of nucleotide similarity to the I3 cDNA. This homology (70% identity) occurs between the alanine/proline-rich C-terminus of the predicted polypeptide of I3 and part of the proline-rich region of the APG protein, and explains the hybridization between the I3 cDNA and the apparently unrelated APG sequence. No homologies between APG and the oleosins was found, and it would seem that there is no relationship between the APG and I3 genes other than a coincidental region of nucleotide similarity.

5.2.3 Isolation of cDNAs from *B. napus* related to the APG Gene

Together with the differences in temporal expression patterns of I3 and APG-homologous mRNAs, the fact that neither the nucleotide or predicted amino acid sequences of the genomic clone and the I3 cDNA showed any other similarity than that mentioned above, it was concluded that APG did not in fact represent the *Arabidopsis* homologue of I3. The *B. napus* 2 mm - 4 mm anther cDNA library from which I3 was originally isolated was therefore rescreened using the 600 bp coding region fragment of APG as a probe. This resulted in the isolation of three strongly hybridizing cDNAs, termed CEX1, CEX2 and CEX6, which hybridized only weakly to the I3 cDNA. Mapping and sequencing indicated that the CEX1 and CEX6 cDNAs overlapped in a region of 700 bp of identity, the CEX1 cDNA apparently having been formed after the internal priming of the CEX transcript by oligo-dT in an A-rich stretch of mRNA during cDNA synthesis (see Fig. 5.5A). Transcripts from which the cDNA sequences of CEX1 and CEX6 derived shall from now on be referred to as CEX transcript, expressed from the *B. napus* gene CEX, which differs from the CEX2 gene and transcripts. The complete nucleotide sequence of the cDNAs was determined (Fig. 5.5), the contiguous CEX1 and CEX6 sequences being combined into one sequence. The combined CEX sequence (Fig. 5.5A) comprises 1383 base pairs, encoding 449 amino acids. The putative CEX polypeptide, when aligned to the putative APG polypeptide, is 81% similar and 71% identical over it's entire length (Fig. 5.6) but is truncated by approximately 60 amino acids at the N-terminal end, since the cDNA sequence is not full length. The 138 base pair CEX2 cDNA sequence (Fig. 5.5B) encodes a 46 amino acid peptide which is homologous to internal regions of CEX and APG. CEX2 thus appears to represent a short incomplete cDNA sequence derived from a transcript closely related to the CEX transcript. The alignment shown in Fig. 5.6 shows where the CEX2 peptide matches APG and CEX sequences; CEX2 is 89% similar to CEX and 72% similar to APG at the amino acid level.

When used to probe northern dot blots containing the same RNAs as for the APG analysis, the CEX1 and CEX2 cDNAs gave an identical hybridization pattern to each other and to that of the APG probe, implying that the same messages were being detected with each probe as would be expected from their close sequence homologies (Fig. 5.7). No transcripts were detected in vegetative tissues or in buds from the cms line. These results show that at least two related genes which exhibit a high degree of homology to APG are expressed

CEX

```
   1  CCACCTAAGCCTCAACCAAAGCCACCACCAAAGCCGCAACCAAAGCCACCACCGGCACCT
   1   P  P  K  P  Q  P  K  P  P  P  K  P  Q  P  K  P  P  P  A  P
  61  ACCCCGTCTCCATGCCCACCGCAACCACCAAAGCCGCAACCAAAGCCACCACCGGCACCT
  21   T  P  S  P  C  P  P  Q  P  P  K  P  Q  P  K  P  P  P  A  P
 121  ACCCCGTCTCCATGCCCACCGCAACCACCAAAACCGCAACCAAAGCCACCACCAGCACCT
  41   T  P  S  P  C  P  P  Q  P  P  K  P  Q  P  K  P  P  P  A  P
 181  GGTCCATCACCAAAGCCAGGTCCATCACCATCACCACCTAAGCCACCACCTAGTCCAGCT
  61   G  P  S  P  K  P  G  P  S  P  S  P  P  K  P  P  P  S  P  A
 241  CCGAAACCTGTTCCACCACCTTCTCCATCGCCAAAACCTAGCCCACCCAAGCCACCTGCA
  81   P  K  P  V  P  P  P  S  P  S  P  K  P  S  P  P  K  P  P  A
 301  CCATCGCCAAAACCTAGCCCACCCAAGCCACCTGCCCCATCACCACCTAAGCCACAGAAC
 101   P  S  P  K  P  S  P  P  K  P  P  A  P  S  P  P  K  P  Q  N
 361  AAAACCATACCGGCTGTGTTTTTCTTTGGTGATTCAATCTTTGACACAGGAAATAACAAT
 121   K  T  I  P  A  V  F  F  F  G  D  S  I  F  D  T  G  N  N  N
 421  AACCTAGACACAAAGCTAAAATGTAATTATCGTCCCTATGGTATGGATTTTCCTATGGGA
 141   N  L  D  T  K  L  K  C  N  Y  R  P  Y  G  M  D  F  P  M  G
 481  GTCGCCACCGGTAGATTCAGCAACGGAAGGGTCGCTTCCGATTATATATCCAAATATTTG
 161   V  A  T  G  R  F  S  N  G  R  V  A  S  D  Y  I  S  K  Y  L
 541  GGAGTAAAAGAAATTGTACCGGCATATGTAGACAAGAAACTACAACAAAACAACGAACTA
 181   G  V  K  E  I  V  P  A  Y  V  D  K  K  L  Q  Q  N  N  E  L
 601  CAACAAAGCGATCTACTTACGGGCGTATCTTTTGCTTCGGGTGGTGCTGGCTACCTTCCT
 201   Q  Q  S  D  L  L  T  G  V  S  F  A  S  G  G  A  G  Y  L  P
 661  CAAACATCTGAATCATGGAAAGTAACAACAATGTTGGACCAGTTGACATATTTCCAAGAC
 221   Q  T  S  E  S  W  K  V  T  T  M  L  D  Q  L  T  Y  F  Q  D
 721  TACAAGAAAAGAATGAAGAAGCTAGTAGGAAAAAAGAAGACTAAAAAGATTGTGAGCAAA
 241   Y  K  K  R  M  K  K  L  V  G  K  K  K  T  K  K  I  V  S  K
 781  GGAGCAGCCATTGTGGTCGCCGGAAGCAATGATCTGATTTATACATATTTTGGAAATGGT
 261   G  A  A  I  V  V  A  G  S  N  D  L  I  Y  T  Y  F  G  N  G
 841  GCTCAACACCTCAAAAATGACGTCGACTCTTTTACCACTATGATGGCTGATTCTGCTGCC
 281   A  Q  H  L  K  N  D  V  D  S  F  T  T  M  M  A  D  S  A  A
 901  AGTTTCGTCTTGCAACTGTATGGATATGGAGCAAGACGTATAGGAGTGATCGGAACACCA
 301   S  F  V  L  Q  L  Y  G  Y  G  A  R  R  I  G  V  I  G  T  P
 961  CCGATTGGATGTACACCTTCACAGAGAGTTAAAAAGAAAAAAATATGTAATGAAGATCTA
 321   P  I  G  C  T  P  S  Q  R  V  K  K  K  K  I  C  N  E  D  L
1021  AATTATGCTGCTCAGCTTTTCAATTCTAAGCTCGTAATTATTTTGGGTCAACTGTCCAAA
 341   N  Y  A  A  Q  L  F  N  S  K  L  V  I  I  L  G  Q  L  S  K
1081  ACCCTACCAAATTCTACAATTGTTTACGGGGACATCTACTCTATCTTCAGTAAGATGCTA
 361   T  L  P  N  S  T  I  V  Y  G  D  I  Y  S  I  F  S  K  M  L
1141  GAAAGCCCGGAAGACTATGGGTTTGAGGAGATAAAAAAACCGTGTTGCAAAATAGGATTG
 381   E  S  P  E  D  Y  G  F  E  E  I  K  K  P  C  C  K  I  G  L
1201  ACGAAAGGAGGTGTATTCTGCAAAGAGAGGACACTCAAAAATATGTCCAATGCATCGTCT
 401   T  K  G  G  V  F  C  K  E  R  T  L  K  N  M  S  N  A  S  S
1261  TATCTGTTTTGGGACGGTCTCCATCCCAGTCAGAGAGCTTATGAAATATCAAACAGAAAA
 421   Y  L  F  W  D  G  L  H  P  S  Q  R  A  Y  E  I  S  N  R  K
1321  CTTGTAAAGAAATACATCCATTTTATCTGATTAATAAGTACCTTTATTATATTATTATTA
 441   L  V  K  K  Y  I  H  F  I
1381  AAGGAATTC
```

CEX2

```
   1  CAGCCACCACCAGCACCTGCTCCATCACCAAAGCCAGGCCCATCACCGCCACCACCTAAA
   1   Q  P  P  P  A  P  A  P  S  P  K  P  G  P  S  P  P  P  P  K
  67  CCACCTAGTCCAGTTCCGAAACCTGTACCACCACCTGCCCCGTCACCAAAACCTAGCCCA
  23   P  P  S  P  V  P  K  P  V  P  P  P  A  P  S  P  K  P  S  P
 133  CCTGCTCCATCGCCAAAA
  45   P  A  P  S  P  K
```

Figure 5.5

Nucleotide and deduced amino acid sequences of the CEX and CEX2 cDNAs.

The CEX sequence is derived from the contiguous overlapping sequences of the CEX1 and CEX6 cDNAs, CEX1 having been primed in the underlined A-rich region.

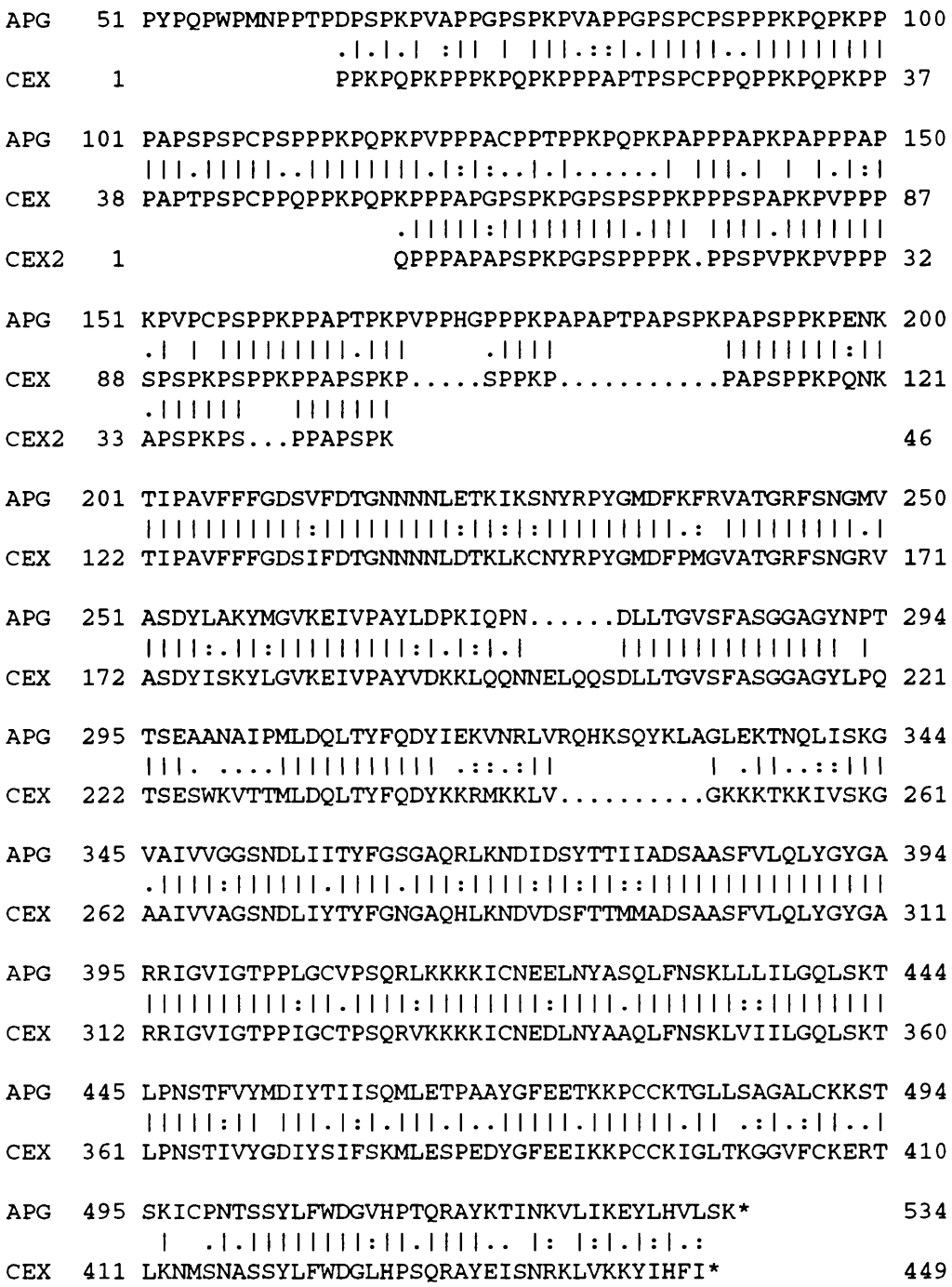


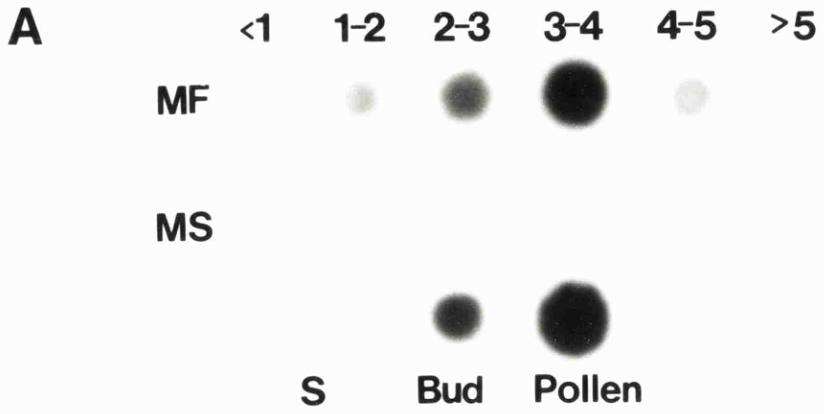
Figure 5.6

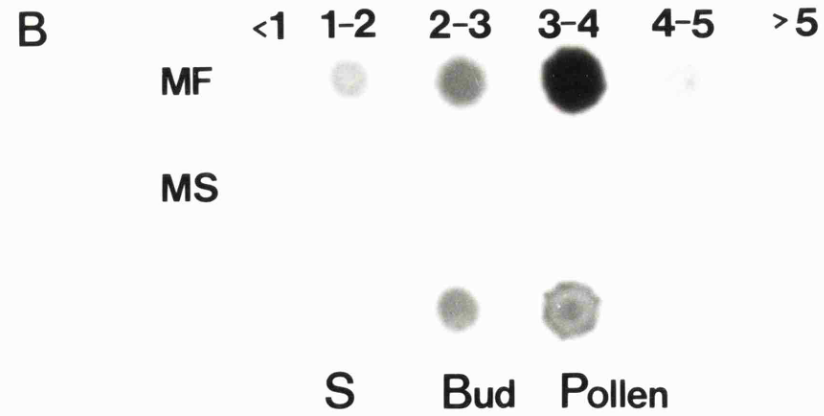
Alignment of the amino acid sequences of the APG, CEX and CEX2 polypeptides. Gaps (.) were introduced to produce maximum alignment.

Figure 5.7

A. Northern dot blot of staged *B. napus* bud RNAs from male fertile (MF) and cytoplasmic male sterile (MS) plants probed with the CEX1 cDNA. Bud lengths, in mm, are indicated. Also included are *B. napus* extruded immature pollen, bud and seedling (S) RNAs. All dots contain 10 µg of total RNA.

B. An identical panel of RNA dots probed with the CEX2 cDNA.


A
<1
1-2
2-3
3-4
4-5
>5
MF
MS
S
Bud
Pollen


B
<1
1-2
2-3
3-4
4-5
>5
MF
MS
S
Bud
Pollen

in fertile flower buds of *B. napus* in a temporally regulated fashion. Furthermore, like the I3 oleosin gene, these genes are not expressed in Ogura cms flower buds, implying that the transcripts may be confined to the developing microspore.

5.2.4 APG and CEX Represent a Conserved Gene in *A. thaliana* and *B. napus*

The degree of similarity between the CEX cDNA and the APG coding region infers that the CEX mRNA is the product of a *B. napus* gene homologous to the *Arabidopsis* APG gene. When used as a probe in a Southern blot of restricted genomic DNA, the CEX1 cDNA hybridized to four fragments in *B. napus* DNA, but only one fragment in *Arabidopsis* DNA (Fig. 5.8A). The 4.5 kb CEX1-homologous *Eco* RI fragment from *Arabidopsis* was the same length as that present in clone λAPG10, implying that CEX1 probe was hybridizing uniquely to the APG gene. This was confirmed by reprobing the filter with cloned APG sequences (Fig. 5.8B). A northern gel blot of *B. napus* and *Arabidopsis* bud RNA, together with *Arabidopsis* flower and leaf RNA was also performed using the CEX1 cDNA as a probe (Fig. 5.9). Two hybridizing mRNAs were observed in *B. napus* buds of 3-4 mm in length, but only one transcript was detected in *Arabidopsis* buds, the mRNA being approximately 2 kb, the same as the major hybridizing mRNA species in *B. napus*. The length of this RNA is of the size which would be predicted on the basis that the deduced translated region of APG is correct. No hybridization to RNA from open *Arabidopsis* flowers or from leaves was detected.

5.3 Discussion

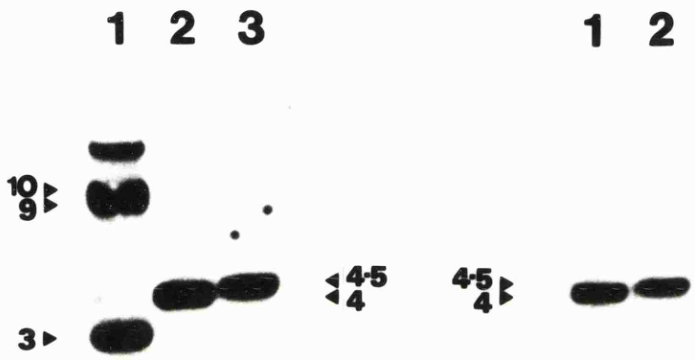
5.3.1 Relationship Between the APG and CEX Genes

The original aim of these investigations was to obtain a genomic clone of the *Arabidopsis thaliana* pollen-specific oleosin "I3". Clearly this aim was not achieved. However, by virtue of a region of similarity between the I3 cDNA and the cloned APG sequence, we have identified a different, but similarly regulated gene present in both *Arabidopsis thaliana* and *B. napus*. The evidence suggests that the *Arabidopsis* APG and *B. napus* CEX (and probably CEX2) genes encode equivalent proteins in both plant species. That a messenger RNA of the same molecular size appears specifically in developing buds of both species strongly supports this contention. Sequence comparison indicates that at least as much identity between CEX and APG lies in regions outside the proline-rich domain as within it, and thus CEX1 cDNA probe is likely to hybridize specifically to CEX, CEX2 and APG sequences in northern and Southern blot analyses, rather than to other transcripts encoding proline-rich gene products. The second mRNA detected in *B. napus* buds is probably expressed from the related gene CEX2, and the Southern blot analysis shows that CEX and CEX2 are members of a small gene family in *B. napus*. Different members of this gene family may well have arisen from the nuclear parents of *B.napus*, as this species is an amphidiploid hybrid between *B. campestris* and *B. oleracea*

Figure 5.8

A. Southern blot of genomic DNA from *B. napus* and *Arabidopsis* probed with the CEX1 cDNA.
Lane 1; *B. napus* DNA x *Eco*RI, lane 2; *A. thaliana* DNA x *Eco*RI, lane 3; *A. thaliana* DNA x *Xba* I.
B. *Arabidopsis* DNA cut with 1; *Eco* RI, 2; *Xba* I. Probed with the 5-prime 2.1 kb *Eco*RI/*Xba* I fragment of λAPG6.

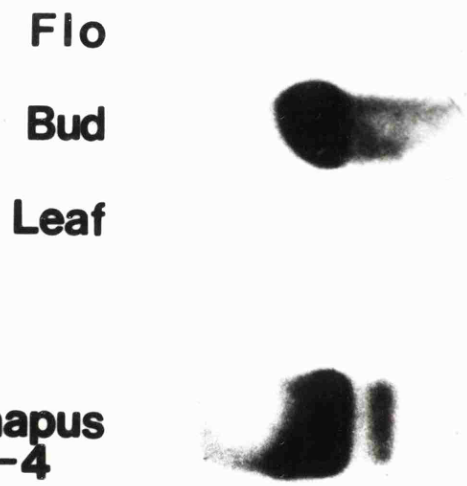
Sizes of restriction fragments are indicated in kilobase pairs.



A B

Figure 5.9

Northern blot including 10 μg RNA from *B. napus* 3-4 mm buds (lane 1), *Arabidopsis* leaves, *Arabidopsis* buds, and *Arabidopsis* open flowers (Flo) probed with the CEX1 cDNA. The major hybridizing band is at approximately 2.0 kb.


Flo
Bud
Leaf
B. napus
3-4

(Stace, 1991). It is thus probable that the CEX and CEX2 genes are derived from the APG genes of these two parental species, and will henceforth be refered to as the APG1 and APG2 genes of *B. napus*.

5.3.2 Expression of APG Transcript

The lack of hybridization of both APG cDNA and genomic fragments to RNA from the male sterile *B. napus* buds indicates that the APG transcript detectable in fertile buds may be confined to the microspore, as the major phenotypic difference in the cms variety is in pollen formation. The possibility of expression in the tapetum cannot completely be ruled out (the tapetum is an important contributer to microspore development, and so lesions in tapetal function may be the cause of sterility), although it has been demonstrated by Scott *et al.* (1991a) that unlike APG1 and APG2, three tapetum-specific genes, A3, A8 and A9 are transcribed in *B. napus* Ogura anthers.
Using the correlation of bud length to the developmental stage of *B. napus* anthers established in the above study, the developmental regulation of the APG mRNAs can be predicted from their hybridization in staged bud RNA populations to cDNA and *Arabidopsis* genomic probes. Northern analysis shows that expression increases through microspore development (2 to 3 mm buds) peaking before the microspore mitosis (3 to 4 mm buds). Transcript levels then decrease dramatically during the interphase preceding the pollen grain mitosis (4 to 5 mm buds) and the mRNA is undetectable in maturing pollen (buds over 5 mm). Although similar experiments were not attempted in *Arabidopsis* because of the small size of the flowers, expression of the APG gene was demonstrated in developing buds but not in open flowers. In evolutionary terms, *B. napus* and *Arabidopsis* are closely related (both are members of the Brassicaceae) and it is therefore presumed that the *Arabidopsis* APG gene and the *B. napus* APG1 and APG2 genes are regulated in a similar spatial and temporal fashion.

5.3.3 Possible Functions of the APG Gene Product

Compelling evidence that APG and CEX represent a conserved gene comes from sequence data. Both open reading frames in the APG and CEX sequences show a high degree of amino acid homology (81% similarity, 71% identity) consistent with the proposal that the two encoded proteins have a common function. Likewise, the CEX2 cDNA shows a comparative degree of similarity to APG over the limited sequence obtained, and presumably represents a further related gene. The isolation of the APG gene by hybridization to the I3 cDNA appears to have been due only to a small region of sequence similarity, and not because of any functional relationship between the two genes. The predicted amino acid sequences of the APG genes contain a large proline-rich region, which may give an indication as to their possible function. Many proline-rich proteins, in the form of hydroxyproline-rich glycoproteins (HRGPs) are found in the plant cell wall (Cassab and Varner, 1988) and although computer comparisons did not recognize any significant

homology between APG and previously characterized proteins, it is possible that they may have a similar architectural function. The sequences encoding several novel proline-rich proteins have recently been cloned (e.g. Sauer *et al.*, 1990; Raines *et al.*, 1991) including three expressed specifically in the epidermis of sunflower anthers (Evrard *et al.*, 1991) and many of these have been ascribed structural functions. The predicted protein sequence of APG includes a typical hydrophobic N-terminal signal peptide region, which is pre-requisite for proteins export to the cell wall. Furthermore, APG also has a high lysine content, and lysine residues are believed to be involved in the peroxidase-mediated cross-linking of plant cell wall proteins (Cassab and Varner, 1988). Additionally, the APG protein includes four potential glycosylation sites, inferring further likeness between it and HRGPs, as well as containing repeated peptide sequences, which are also common in structural proteins. It is known (Vithanage and Knox, 1976) that the microspore contributes exclusively to the deposition of protein to the inner layer of the pollen wall, the intine, during the period when maximal expression of APG is observed, that is, during late microspore development. It is thus possible that these genes encode cell wall or membrane-bound proteins which may play some role in the complex biology of the pollen wall.

5.4 Summary

The flower bud-specific *Arabidopsis* APG gene, which was cloned by homology to the I3 oleosin cDNA, encodes a novel proline-rich polypeptide, but shows only limited similarity to I3 at the nucleotide sequence level, and no significant amino acid sequence similarity. Two *B. napus* transcripts, CEX and CEX2, were identified by the isolation of three cDNAs, CEX1, CEX2 and CEX6, which show a much greater degree of sequence homology to the APG gene. These transcripts were also found to be anther-specific in fertile *B. napus*, but not expressed in the Ogura cms variety, implying that like the oleosin I3, they are probably microspore-specific. Under conditions of high stringency, APG coding sequences hybridize to the same messages and genomic DNA fragments as CEX cDNA probes in northern and Southern analyses. These data, together with sequence comparisons, imply that the APG, CEX and CEX2 coding regions are conserved, and probably derive from a common ancestral sequence. Thus *B. napus* contains at least two APG genes (APG1 and APG2) which produced the transcripts corresponding to the cDNAs CEX and CEX2. Determination of the nucleotide sequence of regions flanking the APG coding region identified the 5' upstream sequence, which is likely to include the elements regulating the transcription of the gene. The construction of reporter gene fusions to analyse the characteristics of any such regulatory elements will be discussed in the following chapter.

CHAPTER 6

ACTIVITY OF THE APG PROMOTER IN TRANSGENIC PLANTS

6.1 Introduction

Before commencing studies on the control of transposition using the regulatory regions of the APG gene, it was important to investigate more thoroughly the developmental and tissue-specific patterns of transcription which would be directed by the promoter. The bacterial gene ß-glucuronidase (*gus*) has been developed as a reporter gene (Jefferson *et al.*, 1987) and is now widely used to monitor promoter activities in transgenic plants. The GUS enzyme can be assayed both quantitatively and qualitatively via it's action on a number of substrates, both *in situ* and *in vitro*, and the activity of promoter-*gus* gene fusions can be easily localized visually at the cellular level by histochemical staining. This chapter describes the behaviour in transgenic plants of fusions of upstream regions of the APG gene to *gus*, and compares this to the expression of the native gene as determined by northern blotting. The implications of these results for experiments using the APG promoter to drive transposition and the construction of APG/*Ac* transposase fusions will be discussed.

6.2 Results

6.2.1 Construction of Reporter Gene Fusions

Sequence analysis of the APG gene and estimation of the length of the APG mRNA, had previously identified the probable start of translation of the APG gene (see 5.2.2). A *Sau* 3AI restriction site was situated at position +8 relative to the ATG, just inside the coding region of APG, and this site was chosen as the point at which to make a fusion with the coding region of the *gus* gene. Such fusions would be translational, in that translation of the transcribed mRNA would initiate at the ATG of the APG sequence, and the resulting GUS protein would include extra amino acids encoded by APG and plasmid sequence. After cloning as described below, the fusions contained 11 and 18 extra amino acids, respectively. This was considered unlikely to affect the behaviour of the enzyme as the GUS enzyme has been shown to be unaffected by quite long N-terminal extensions. However, it was important that the chimaeric *gus* gene was translated in the correct reading frame. Three binary vectors were available (pBI101.1, pBI101.2 and pBI101.3) based on the original pBI101 plasmid described by Jefferson *et al.* (1987) providing the means to create fusions in any of the three possible reading frames.

A 541 base pair *Sau* 3AI fragment containing the -530 to +11 base pair region of the APG gene was subcloned into the *Bam* HI site of pBluescript SK-, forming a plasmid called pPROM7.5 (Fig. 6.1A). A second promoter fragment was obtained by PCR from a plasmid containing the 2.9 kb *Xba*I fragment of λAPG6. An oligonucleotide primer was synthesized with the sequence 5'-AATCGACCAAGGATGATCGC-3', complementary to base pairs +6 to +25, and used to prime the PCR reaction in conjunction with the standard M13 reverse primer, which binds in the polycloning site of pBluescript. The amplified

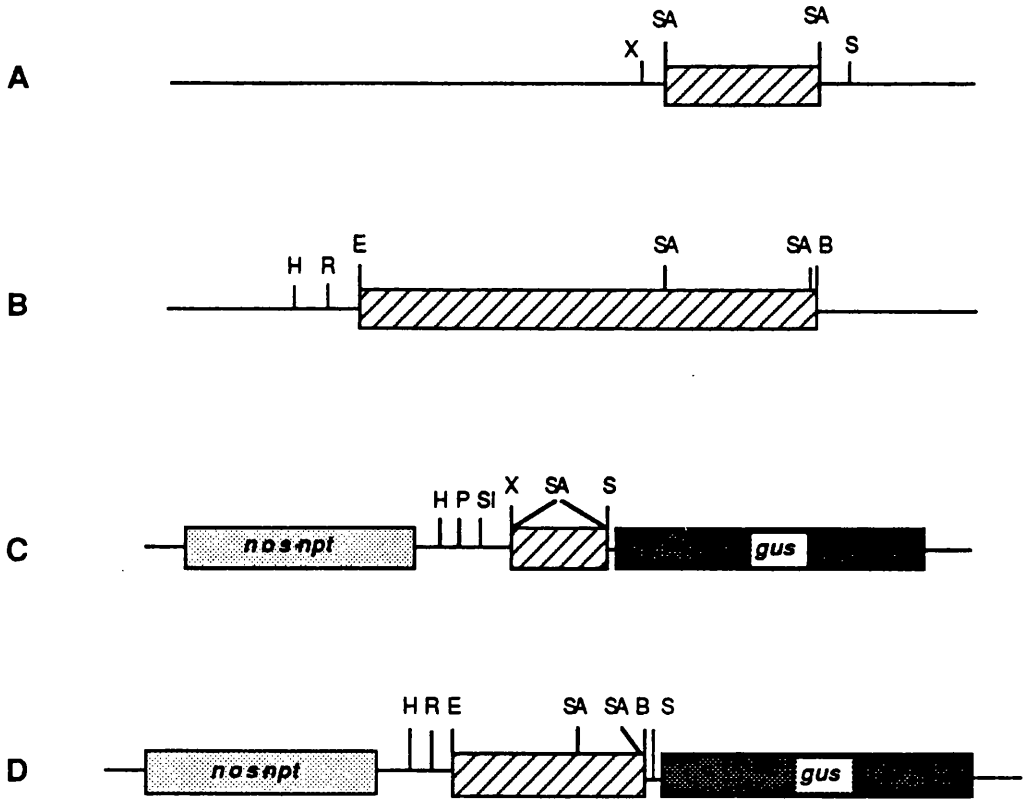
Figure 6.1. Diagrams showing the plasmids involved in constructing the APG-*Ac* fusions.

A. pPROM7-5, containing the 541 bp *Sau* 3AI promoter fragment in the *Bam* HI site of pBluescript SK-.

B. pPROM9, containing the PCR amplified 1.5 kb APG upstream region, cloned as an *Eco* RI / *Bam* HI fragment from pPROM8 in pBluescript SK-.

C. The T-DNA region of pPSG2, containing the pPROM7-5 *Xba* I / *Sma* I promoter fragment in pBI101.3.

D. The T-DNA region of pPSG3, containing the pPROM9 *Hin* dIII / *Bam* HI promoter fragment in pBI101.3.



RESTRICTION SITES:

B = BamHI
E = EcoRI
H = HindIII
R = EcoRV
S = SmaI
Sl = SalI
SA = Sau3AI
X = XbaI
P = PstI

GENES:

nos-*npt* :
kanamycin resistance
(selectable marker)

gus = ß-Glucuronidase

2.2 kb fragment was purified and digested with *Sal* I, after repair of overhanging ends with Klenow, generating a *Sal* I / blunt ended fragment which was cloned into the *Sal* I and *Sma* I sites of pBluescript pSK-, producing pPROM8. This plasmid was cleaved at the *Eco* RI site in position -1501 of the APG gene, plus *Bam* HI, producing a 1.5 kb fragment which was cloned into the *Eco* RI and *Bam* HI sites of pBluescript pSK- to form pPROM9 (Fig. 6.1B). Sequencing of an isolate of this clone demonstrated that no PCR-induced mutations had occured within the 3-prime 350 base pairs, and it was therefore considered suitable for constructing a *gus* fusion. An *Xba* I/*Sma* I fragment of pPROM7.5, containing the APG promoter in the correct orientation, was cloned into the *Xba* I and *Sma* I sites of pBI101.3, to give the plasmid pPSG2, whilst a *Hin* dIII / *Bam* HI promoter fragment of pPROM9 was cloned into the *Hin* dIII and *Bam* HI sites of pBI101.3 to produce pPSG3 (Figs. 6.1C and D). The binary plasmids pPSG2 and pPSG3 were mobilized into *Agrobacterium tumefaciens* strain LBA4404, and the resulting clones used to transform *N. tabacum* (SR1) and, in the case of pPSG3, *Arabidopsis thaliana* (C24).

6.2.2 Quantitative Analysis of GUS Activity in Transgenic Plant Tissues

ß-Glucuronidase activity can be detected in different organs, tissues and cells of transformed plants using a sensitive fluorimetric assay. This procedure was used to assess the activity of the APG-*gus* fusions in plants transformed with the PSG2 and PSG3 constructs described above. Eight SR1 *N. tabacum* plants transformed with the PSG2 construct and seven transformed with PSG3 were analysed. All primary transformants were allowed to selfed-pollinate, and the resulting seed plated on media containing kanamycin to select for the inheritance of a T-DNA copy. In all cases, a high ratio of segregation of resistance to sensitivity was observed (greater than 98:1), demonstrating that T-DNAs were present at multiple loci in the primary transformants.

The expression of the APG genes had been characterized by northern analysis of mRNA levels in *Arabidopsis* and *B. napus*, and the activity of the APG-*gus* gene in *N. tabacum* was expected to mirror the expression patterns of the native genes. This prediction was found to hold true. GUS activity was detected in developing anthers at levels of 200-124,000 pmole 4-MU/min/mg protein, depending on the size of the anther, but not in any other floral organs (assays on buds from which anthers had been removed), nor in developing seed pods, roots, leaves, petioles, or whole seedlings (including cotyledons, developing first true leaves, and roots) of T2 plants when compared against the background in wild type *N. tabacum* (0-20 pmole 4-MU/min/mg protein).
Since the level of specific activity of GUS varied widely between anthers, some form of developmental regulation was assumed to occur, as observed for APG in *B. napus.* A study of temporal regulation in transgenic *N. tabacum* anthers required a preliminary survey of microgametogenesis with respect to anther and bud lengths in SR1 plants. Similarly to

the findings of Schrauwen *et al.* (1990) and Koltunow *et al.* (1990) it was established that the stage of development of the gametophyte could be predicted from the length of *N. tabacum* buds and/or anthers (Table 6.1). In SR1 *N. tabacum* buds, microsporogenesis is complete when buds are 8 - 10 mm long, when they contain anthers of 3 mm in length. Microspore development occurs in buds of 10 -16 mm in length, and for the purpose of this study, this phase of development is arbitrarily divided into early, mid and late stages, which are defined by anther lengths of 3.5, 4.0 and 4.5 mm respectively. Fluorescence microscopy using the nuclear stain DAPI (6-diamidno-2-phenolindole dihydrochloride) showed that microspore mitosis occurs in buds of 16-20 mm in length, after which the binucleate pollen grains undergo maturation in anthers of 4.5-5 mm in length, until dehiscence, when the flower is around 45 mm long.

BUD SIZE (mm)	ANTHER SIZE (mm)	DEVELOPMENTAL STAGE	CHARACTERISTICS OF MICROSPORES.
5 - 8	1.5 - 3.0	Microsporogenesis	Sporogenous cells surrounded by callose.
9 - 10	2.5 - 3.5	Microspore Release	Callose Dissolution; Ghost tetrads present.
10 - 12	3.0 - 3.5	Early Microspore Development	Uni-nucleate Microspores; Spores oval shaped.
11 - 15	3.5 - 4.0	Mid Microspore Development	Uni-nucleate microspores; Spores Rounded.
13 - 16	4.0 - 4.5	Late Microspore Development	Uni-nucleate microspores
16 - 20	4.5 - 5.0	Microspore Mitosis	Dividing nucleus
21 - 30	4.5 - 5.0	Early Pollen Maturation	Binucleate spores enlarging
31 - 40	4.5 - 5.0	Late Pollen Maturation	Fully expanded, round pollen grains, heavy sculpturing.
41 - 45	5.0	Maturity	Pre-Anthesis; Mature Pollen grains.

Table 6.1

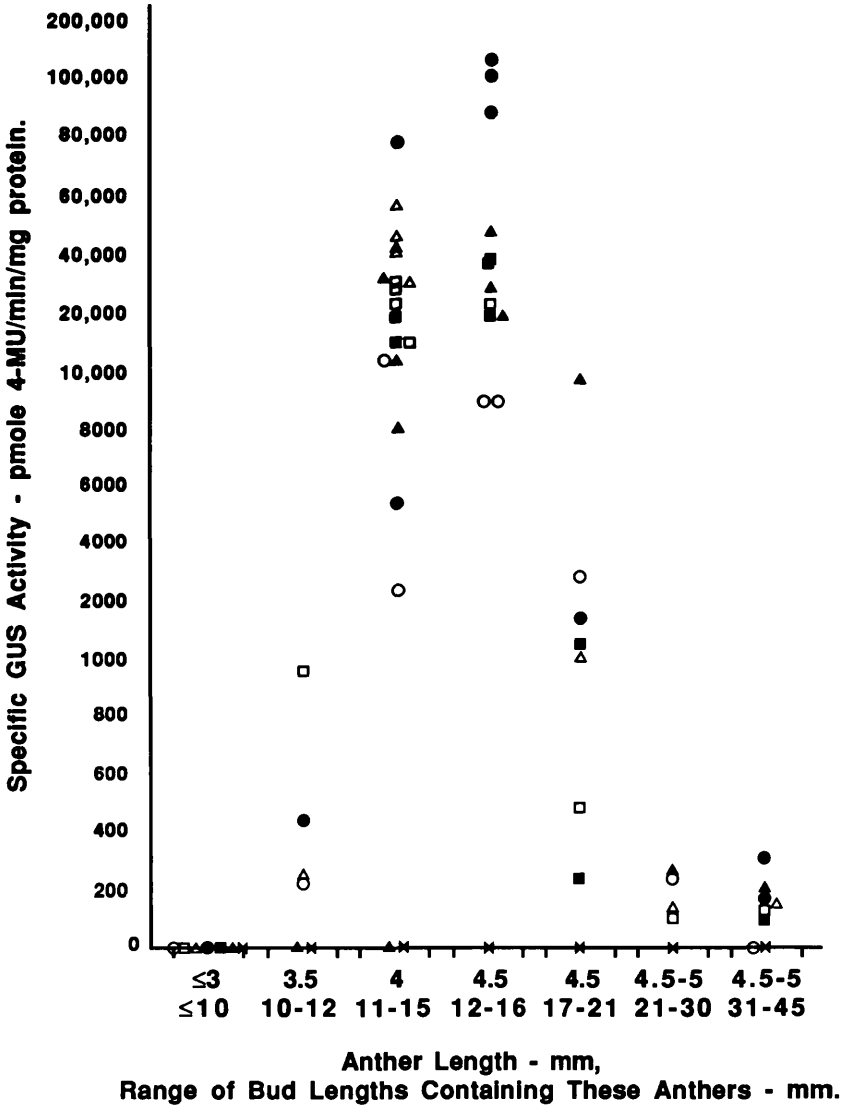
Relationship between *N. tabacum* bud and anther lengths, and developmental stage of the gametophyte.

Three independent transformants for each of the PSG2 and PSG3 constructs were assayed in detail for the level of specific GUS activity in protein extracts from anthers at a range of developmental stages. The resulting data, presented in Fig. 6.2, show that a similar pattern of activity was found for all six transformants, highest levels of activity being confined to anthers of 4-4.5 mm from buds of 11-16 mm in length, corresponding to mid to late microspore development. No activity could be detected during microsporogenesis, and levels of activity during early microspore development were low. Following microspore development, GUS activity decreases during the period of microspore mitosis to a low level which was maintained during pollen maturation.

6.2.3. Histochemical Localization of GUS Activity in Anthers

To investigate the spatial regulation of APG, GUS activity was histochemically localized in a large number of fixed anther sections from many independent *N. tabacum* transformants harbouring both the PSG2 and PSG3 fusions. The reader is referred to Fig. 4.1 for an illustration of the major events during anther development. Assays established that both fusions behaved in a similar manner, and were activated not only in the microspores as had been expected, but additionally in several different sporophytic cell types within the anther. This activity was found to be regulated in a complex developmental sequence, and a summary of this information is presented in Table 6.2. Histochemical staining reactions showed that the APG-*gus* genes were first active in the tapetum during early microspore development (Fig. 6.3a). During mid and late microspore development, when maximum GUS activity is observed, GUS was detected in three sporophytic anther tissues, the tapetum, the stomium (the line of weakness which ruptures to bring about dehiscence) and the anther wall (Fig. 6.3b) and additionally in a proportion of microspores. Staining in the tapetum and wall was significantly reduced in anthers from buds around 17-20 mm (Fig. 6.3c), the time of microspore mitosis, and was not detected in buds over 25 mm. Activity in the stomium remained at a high level until the early pollen maturation stage (Fig. 6.3c). In the gametophyte, staining became more rapid and intense in binucleate pollen grains, reaching an apparent peak in activity (though differences in staining may be affected by differing permeability of the developing pollen wall) during early maturation before decreasing in buds over 25-30 mm until dehiscence. This pattern of gametophytic expression was confirmed by histochemical assays of isolated microspores and pollen in the presence of ferri- and ferrocyanide, oxidizing agents which prevent diffusion of the intermediate product of the staining reaction, eliminating the possibility of contamination by GUS enzyme from other tissues. Light micrographs of spores from a single PSG3 transformant subjected to such an assay are shown in Fig. 6.3. This PSG3 plant was shown to contain multiple loci for T-DNA insertions by kanamycin resistance segregation analysis of selfed progeny. Therefore one would expect the majority of microspores to inherit one or more copies of the APG-*gus* gene following meiosis. Upon staining for GUS activity, none was detected in

Figure 6.2. GUS activity in transgenic *N. tabacum* anthers.
Scatter plot showing specific GUS activity in protein extracts from anthers of staged buds. Size classes represent the developmental stages of microsporogenesis, early, mid and late microspore development, microspore mitosis, and early and late pollen maturation respectively. The open and closed symbols represent three independent transformants with the PSG2 and PSG3 constructs respectively. Wild type activity is shown by the symbol 'x'.


Specific GUS Activity - pmole 4-MU/min/mg protein.
200,000
100,000
80,000
60,000
40,000
20,000
10,000
8000
6000
4000
2000
1000
800
600
400
200
0
≤3 ≤10
3.5 10-12
4 11-15
4.5 12-16
4.5 17-21
4.5-5 21-30
4.5-5 31-45
Anther Length - mm,
Range of Bud Lengths Containing These Anthers - mm.

Developmental Stage	Spores	Tapetum	Stomium	Wall
MSG	-	-	-	-
MR	-	-	-	-
EMD	-	+	-	-
MMD	+/-	+++	+++	++
LMD	+/-	+++	+++	++
MM	++	+	+++	+
EPM	++	-	-	-
LPM	+/-	-	-	-
M	+/-	-	-	-

MSG = Microsporogenesis
MR = Microspore Release
EMD = Early Microspore Development
MMD = Mid Microspore Development
LMD = Late Microspore Development
MM = Microspore Mitosis
EPM = Early Pollen Maturation
LPM = Late Pollen Maturation
M = Maturity

\- = No Staining
+/- = Staining in some spores
\+ = Weak Staining
++ = Moderate Staining
+++ = Strong Staining

Table 6.2

Table showing relative staining (rate of reaction and intensity of staining), following incubation with X-GLUC, in *N. tabacum* anthers from plants transformed with the PSG2 and PSG3 constructs. Developmental stages are as shown in Table 6.1

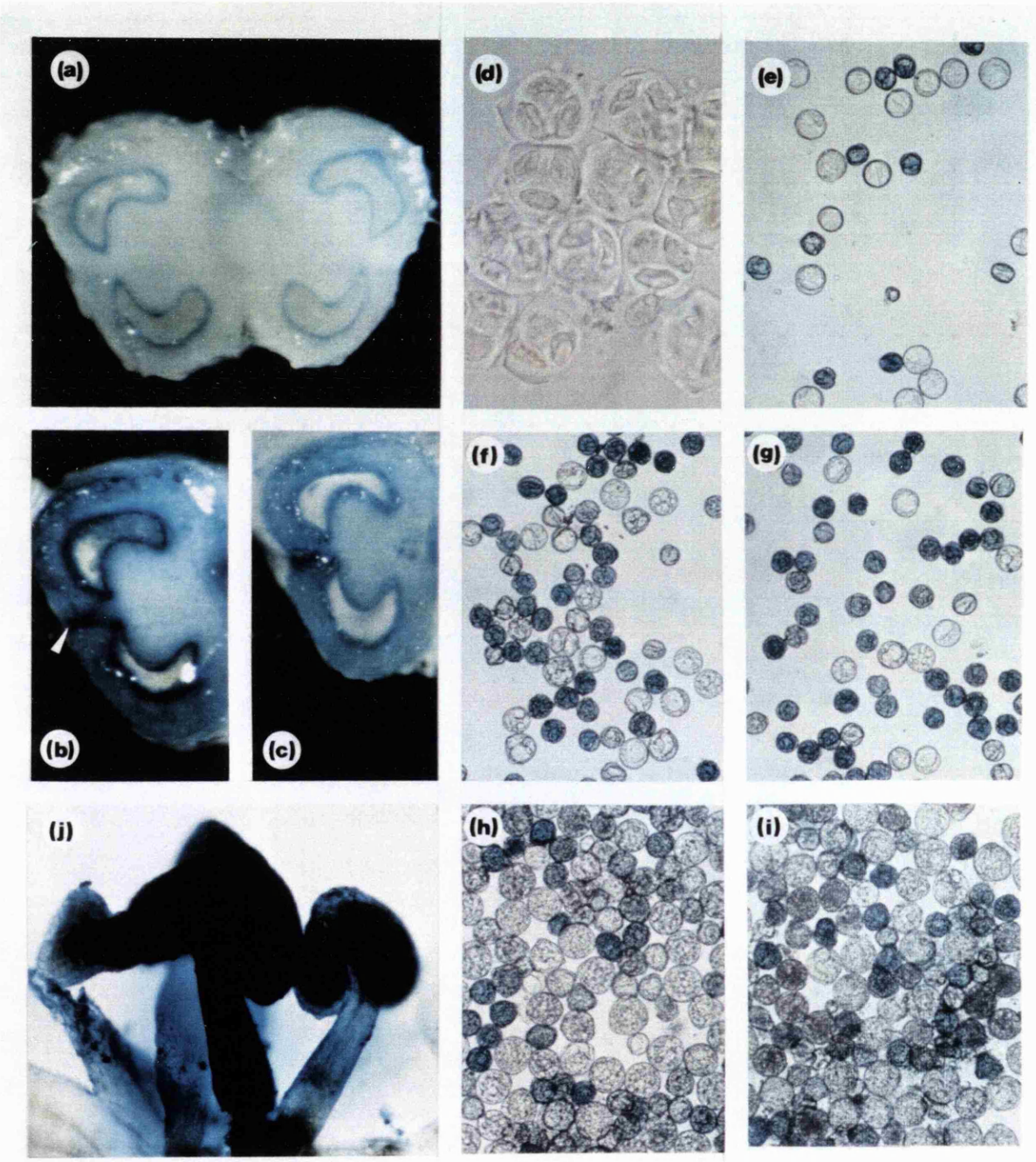
Figure 6.3. Localization of APG-*gus* Expression.

(a)-(c) Histochemical localization of GUS activity in transverse sections of fixed *N. tabacum* anthers.

(a) Anther from a PSG3 transformant showing staining only in the tapetum, bud length = 11 mm, anther length = 3.5 mm. **(b)** Anther from a PSG2 transformant showing staining in the tapetum, wall and stomium (arrowed), bud length = 12 mm, anther length = 4 mm. **(c)** Anther from a PSG2 transformant showing strong staining only in the stomium, bud length = 17 mm, anther length = 4.5 mm.

(d)-(i) Histochemical localization of GUS activity in spores extracted from anthers of a PSG3 transformant. **(d)** Tetrads of microspores showing no GUS activity. **(e)** Microspores from a late microspore development stage anther (bud = 15 mm). **(f)** Spores from an early pollen maturation stage anther (bud = 20 mm). **(g)** Spores from an early pollen maturation stage anther (bud = 25 mm). **(h)** Spores from a late pollen maturation stage anther (bud = 30 mm). **(i)** Mature pollen grains (bud = 40 mm). **(d)**: magnification 12.5x100; **(e)-(i)**: magnification 12.5x16.

(j) Squash of an *Arabidopsis* flower from a PSG3 transformant showing high GUS activity in the anthers (This tissue was not fixed and the filament is stained because of diffusion of the reaction product.)


(a)
(d)
(e)
(b)
(c)
(f)
(g)
(j)
(h)
(i)

tetrads (Fig. 6.3d) or immediately following microspore release, but during late microspore development approximately one third of microspores stained GUS-positive (Fig. 6.3e). Highest activity was detected during early pollen maturation (Fig. 6.3f and 6.3g) before a reduction in expression after pollen grain expansion in the transition to late maturation (Fig. 6.3h and 6.3i). It is interesting to note that the non-staining spores in Figs. 6.3e, 3f and 3g are swollen and vacuolate, a form classically known as the "signet ring" stage, which is characteristic of mid microspore development. The large vacuole is reduced during late microspore development as the microspore synthesizes new cytoplasm in preparation for mitosis. Thus, although the GUS-positive spores are smaller than signet ring stage microspores, they may be more mature, since they become more prevalent as development procedes. Because most microspores were expected to carry an APG-*gus* gene, the difference in expression in microspores from the same bud are probably attributable to this lack of synchrony of development, which may be due to non-ideal growth conditions for example. Despite the lack of synchrony, DAPI staining confirmed that most GUS-positive microspores from late microspore development anthers were uninucleate. Maximum expression was detected after mitosis, but Figs. 6.3g and 6.3h demonstrate a subsequent reduction in GUS activity which is concurrent with further maturation of the pollen grains in buds over 30 mm in length.

The PSG3 construct was also used to produce transformed *Arabidopsis* plants, and three independent transformants were regenerated which flowered. When entire floral whorls from these plants were assayed, only anthers gave positive staining (Fig. 6.3j). Stems, leaves and other floral organs showed no activity, even after overnight staining. GUS activity was limited to a small proportion of the buds in a flowering whorl, these being buds of an intermediate size, presumably containing spores midway through development, whilst in slightly larger buds, only the gametophyte appeared GUS-positive.

6.2.4 APG is not Reactivated Following Pollen Germination

The APG and CEX mRNAs were not detected in RNA from mature pollen of *Arabidopsis* or *B. napus*, which produce trinucleate pollen, as opposed to *N. tabacum*, which produces binucleate pollen. It was inferred that expression of CEX in *B. napus* ceases by the end of the interphase preceding the second mitosis (section 5.3.2); in *N. tabacum* the cell cycle is arrested in this interphase in mature pollen, and this difference in gametogenesis prevents an accurate extrapolation of GUS activity in *N. tabacum* to APG expression in *Arabidopsis* at this stage of development. It appeared however, that APG may still be active at a low level in mature pollen. Furthermore, many previously identified gametophyte-expressed mRNAs are present in germinating pollen, e.g. the LAT52, LAT56 and LAT59 transcripts of tomato (Ursin *et al.*, 1989) and the Zmc13 message in maize (Hanson *et al.*, 1989). To test whether the APG promoter was active during pollen

germination, GUS activity in dehisced and *in vitro*-germinated transgenic *N. tabacum* pollen was assayed fluorimetrically. Fig. 6.4 shows a comparison of specific activity in dehisced and germinated pollen from a range of transformants. No clear trend to either an increase or decrease in activity can be seen following germination, levels of activity remaining low, suggesting APG is not reactivated upon germination.

6.2.5 Construction of APG-*Ac* Gene Fusions

To create an *Ac* transposase gene capable of specific expression in the anthers, and especially the microspores of transgenic plants, the APG promoter was fused to the transposase coding region. Although expression from both 0.54 kb and 1.5 kb fragments was indistinguishable at the level of detection of the GUS investigations, both were used to produce transposase fusions. Unlike the *gus* gene fusions however, the transposase fusions needed to be transcriptional, in order that a native enzyme be synthesized. Therefore, the same promoter clones which were produced during the *gus* fusion work could not be used. An oligonucleotide primer was therefore designed to enable PCR amplification of promoter regions containing the entire untranslated region of APG, with the sequence 5'-GGTTTTACTACAAGTAGT-3' complementary to bases -18 to -1 of the gene. This primer was used to synthesize promoter fragments from plasmids pPROM7 and pPROM8 which were cloned into pBluescript and their 3-prime regions sequenced. Because the experiments that we wished to undertake involved a two-element system, in which a transposase produced from one T-DNA locus would *trans* -activate a *Ds* element at a second locus, the engineered *Ac* element should ideally be stabilized. This can be achieved by removal of either or both terminal repeat sequences. In the construction of stabilized *Ac* elements under the control of other promoters, several workers have replaced the entire 5' region of the element with the desired promoter. One such construct is a CaMV35S-*Ac* element produced by Dr. Steve Schofield (Sainsbury Laboratories, IPSR, Norwich) in the *E. coli* plasmid pSLJ1101. This plasmid was kindly made available for the construction of APG transposase fusions. Plasmids termed pAPGAc0.5 and pAPGAc1.5 (shown in Fig. 6.5) were constructed by the substitution of the CaMV35S promoter in pSLJ1101 for the APG promoters. These constructs also include the *N. tabacum* mosaic virus "omega" translational enhancer (Gallie *et al.*, 1987) which may increase levels of transposase enzyme upon expression of the gene. The fusions were ligated into the binary vector pBin19, mobilized to *A. tumefaciens* strains LBA4404 and pGV2260 and the resulting strains used to transform *N. tabacum* var. Petite Havana, and *Arabidopsis thaliana* Landsberg *erecta*. At the time of writing, transformed shoots were being selected in tissue culture.

Figure 6.4. APG activity following pollen germination.
Comparison of specific GUS activity in protein extracts from dehisced and germinating pollen from a single flower of a number of transformants. 2-1-2, 2-1-3 and 2-1-4 carry the PSG2 construct, and 3-1, 3-3, 3-5 and 3-9, the PSG3 construct.

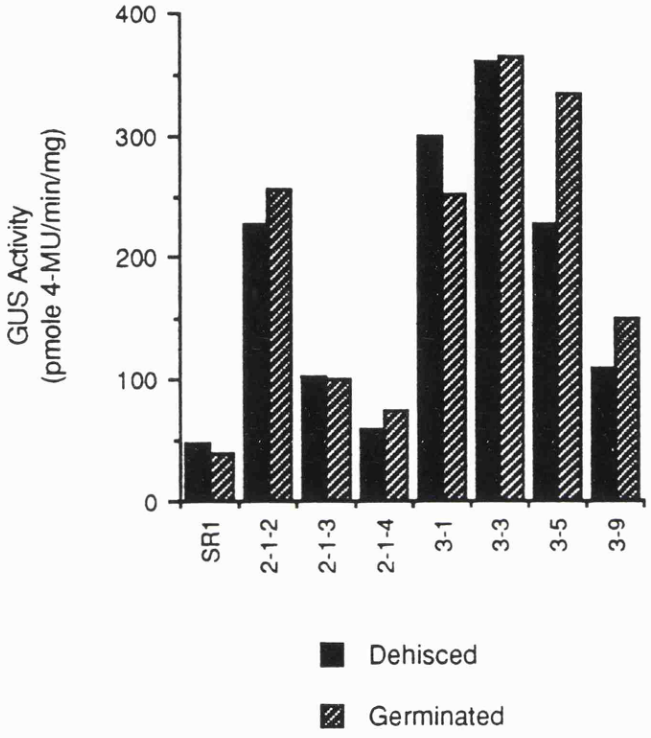
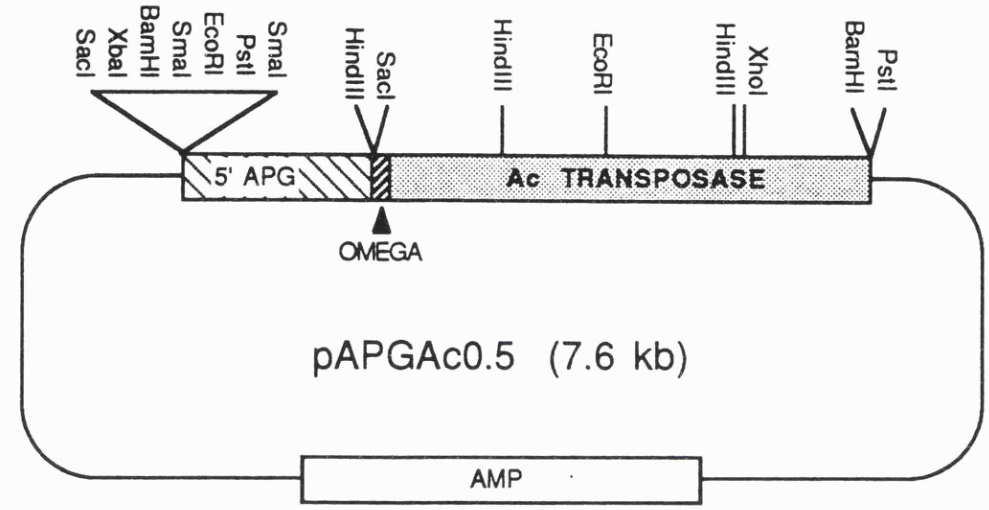
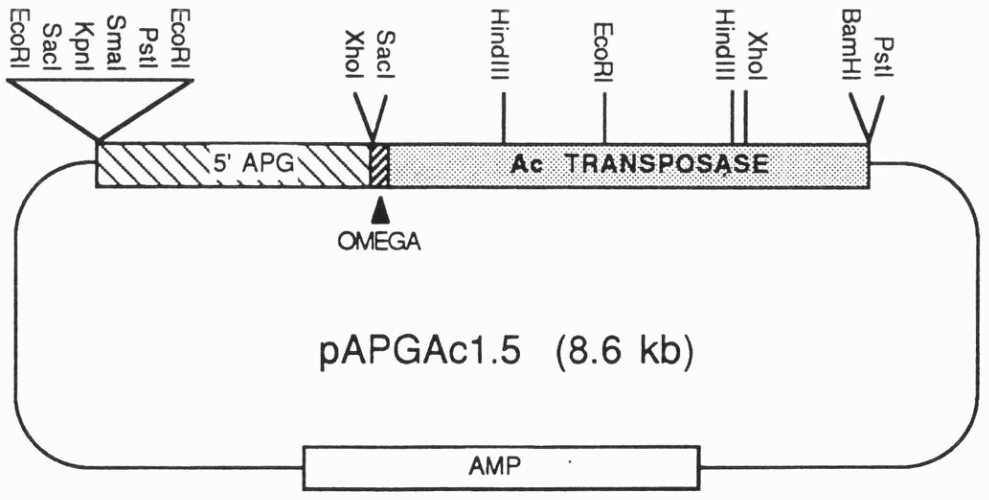

GUS Activity
(pmole 4-MU/min/mg)
400
300
200
100
0
SR1
2-1-2
2-1-3
2-1-4
3-1
3-3
3-5
3-9
Dehisced
Germinated

Figure 6.5. Digrams showing the structure of APG-*Ac* fusions.
The plasmids contain transcriptional fusions of the -530 and -1501 bp promoter fragments, isolated by PCR amplification, to the *Ac* transposase. The plasmids are based on the plasmid pSLJ1101 as described in the text.


SacI
XbaI
BamHI
SmaI
EcoRI
PstI
SmaI
HindIII
SacI
HindIII
EcoRI
HindIII
XhoI
BamHI
PstI
5' APG
Ac TRANSPOSASE
OMEGA
pAPGAc0.5 (7.6 kb)
AMP


EcoRI
SacI
KpnI
SmaI
PstI
EcoRI
XhoI
SacI
HindIII
EcoRI
HindIII
XhoI
BamHI
PstI
5' APG
Ac TRANSPOSASE
OMEGA
pAPGAc1.5 (8.6 kb)
AMP

6.3 Discussion

6.3.1 The APG Promoter Drives Developmentally Complex Anther-Specific Expression of ß-Glucuronidase

An examination of the temporal expression of the *B. napus* APG genes by northern blotting (Fig. 5.7) showed that APG message is detectable in flower buds of 1-2 to 4-5 mm. Using the established correlation of Scott *et al.* (1991a) it was predicted that APG expression increases throughout microspore development (which occurs in 1.8-3 mm buds) peaking around the time of the microspore mitosis (3-4 mm buds) before declining dramatically during pollen maturation (buds over 4 mm). APG message was not detected in mature pollen of *B. napus* (>5 mm buds) or *Arabidopsis* (Fig. 5.9). The relationship between *N. tabacum* microgametogenesis and bud length shown in the results is much less pronounced than that seen in *B. napus*. However, it was possible to recognize a number of developmental stages which could be predicted with a good degree of reliability from the length and appearance of the buds on the plant. This enabled the selection of buds containing a wide range of points of gametogenesis with minimum destruction of material. Although values of activity cannot be interpreted simply when one considers that the histochemical analysis showed that APG is active preferentially within different cell types of the anther, the fluorimetric assays carried out on staged anthers show that in entire anthers, temporal expression in transgenic *N. tabacum* mirrors that determined by northern analysis of *B. napus*. The data in Fig. 6.2 show increasing GUS enzyme activity during microspore development, to a point in late microspore development where maximum activity occurs, before a reduction after microspore mitosis. Activity was still detectable in fully mature *N. tabacum* pollen, though histochemical analysis showed that not all pollen grains expressed GUS (Fig. 6.3i). Histochemical analysis of *N. tabacum* anthers transformed with *gus* fusions to the -530 and -1501 base pair promoters showed indistinguishable patterns of activity. Gene activity was found to be differentially regulated in the tapetum, the stomium, the anther wall and the gametophyte. The timing of appearance and disappearance of GUS activity, whilst not co-ordinate in these tissues, correlates with various developmental and cytological markers. In sporophytic cell types, activation of APG follows microspore release, GUS activity first being detectable in the tapetum. The tapetum and the anther wall stain positive for GUS enzyme until the time of tapetal dissolution, which occurs in buds of 16-22 mm (Koltunow *et al.*, 1990). This stage also corresponds to the time of microspore mitosis, and is when APG activity increases in the male gametophyte from the low level detected during microspore development. This up-regulation in the gametophyte is concurrent with down-regulation in the stomium, the last sporophytic tissue in which APG is expressed. Gametophytic expression appeared to be reduced following the expansion of the pollen grains in 25-30 mm buds. The apparent difference

in expression of APG in mature pollen of *B. napus* and *Arabidopsis* compared to *N. tabacum* may be a reflection of the type of pollen these plants produce. The Brassicaceae produce trinucleate pollen, whereas *N. tabacum* forms binucleate pollen, and thus the developmental switch to inactivate APG may differ in this species. It is interesting to note with regard to this matter that the reduction in APG-*gus* expression in *N. tabacum* pollen does correlate with a morphological change in the appearance of the pollen grains as they mature (c.f. Fig. 6.3g and 6.3h).

The lack of detectable mRNA in male sterile *B. napus* anthers indicates that the Ogura cms factor affects gene expression throughout the anther, not only in the gametophyte where the major phenotypic difference is manifest. The implications of this finding will be discussed in Chapter 7.

Thus, the transcriptional activity of the APG gene, although at first sight simple, is in fact regulated in a complex manner. Such activity is presumably controlled by various *cis*-acting sequences. Several such regulatory elements could exist in the APG promoter. There may be independent sequences which control expression in each particular cell type where activity is required, under the control of cell-specific *trans*-acting factors. Perhaps more likely in view of current understanding of *trans* -acting factors (e.g. reviewed by Glover, 1989) is that a small number of regulatory elements exist which are under the control of a single or few regulatory proteins. These proteins would be differentially activated by specific second messenger pathways in the different cell types, perhaps under the influence of some cell-cell interaction which regulates cell and tissue identity. Since the 0.54 kb promoter in PSG2 is capable of driving expression in all relevant cell types, all necessary *cis* elements for this complex developmental regulation must reside within the -530 base pair region. Nucleotide sequence analysis detected no sequences previously implicated in pollen- or anther-specific transcription, such as those identified by Twell *et al.* (1991), highlighting the complexity of gene expression required for anther development.

The reasons for this complexity presumably relate to the *in vivo* function of the APG protein. The sequence analysis discussed in Chapter 5 led to the suggestion that APG is a structural protein, particularly because of it's proline-richness and N-terminal signal peptide. *gus* fusion experiments have demonstrated that maximum gene expression is not co-ordinate in the sporophyte and gametophyte, but that peak activity in the sporophyte occurs during mid to late microspore development, and in the gametophyte during early pollen maturation. As previously discussed, it is during microspore development that the majority of the intine layer of the pollen wall is laid down (Vithanage and Knox, 1976). APG may thus have a role in intine formation, protein being secreted from the tapetum during microspore development and then from within the binucleate pollen grain after tapetal dissolution. However, during microspore development, synthesis of the outer exine layer is advanced in most species, and APG protein secreted from the tapetum may

well not reach the intine. Structural cell wall proteins of various kinds will naturally also be required in the anther wall and stomium, although why synthesis should occur over such a short and defined period of development is puzzling. An examination of the spatial and temporal expression of APG thus casts doubt on the original proposition that the protein has a simple structural role, though this is by no means excluded by the data. In *N. tabacum* plants, GUS activity could be detected in the gametophyte at all stages of development from mid microspore development up to anthesis. At anthesis, staining showed that around 50% of pollen grains were GUS-positive, whereas in the same plants, all microspores expressed the gene. Thus, transcription and translation had ceased in the GUS-negative pollen grains. Furthermore, germination of dehisced pollen did not lead to a reacivation of APG-*gus* fusions. In *Arabidopsis* anthers exhibiting strongest staining, the reaction product is presumably located in sporophytic cells surrounding the microspores, predominantly the tapetum (see Fig. 6.3j). Other anthers from buds of a similar size were found to contain GUS activity only in the gametophyte, possibly binucleate pollen. No expression was detected in any cells of anthers from older buds, many of which would contain trinucleate pollen. The combined observations of APG-*gus* activity in *N. tabacum* and *Arabidopsis* lead to the conclusion that expression ceases before the mitotic division of the generative nucleus in the pollen grain. This is the same conclusion as was drawn from northern analysis of the homologous APG1 and APG2 genes in *B. napus.*

6.3.2 Relative Strength of APG Expression in Transgenic *N. tabacum*

Fluorimetric detection of GUS activity is a sensitive method of detection of GUS activity in plants (Jefferson *et al.*, 1987), but no significant levels above the background in untransformed *N. tabacum* plants were detected in any organs other than anthers. Due to the expression of APG-*gus* in more than one cell type, it was only possible to assess GUS activity in a single cell type in the case of maturing pollen. Assays of anthers at stages of early and late pollen maturation showed that significant levels of APG-*gus* expression were maintained in the binucleate gametophyte, presumably accounting for all anther activity during this period. The detected activity of between 100 and 400 pmole 4-MU/min/mg protein is comparable with the levels of GUS activity directed by more commonly used promoters, despite the fact that it is at least an order of magnitude lower than was present in whole anthers during microspore development. For example, in the description of *gus* fusions by Jefferson *et al.* (1987) the "constitutive" CaMV 35S promoter in pBI121 produced 321 pmole 4-MU/min/mg protein in leaves, and 577 pmole 4-MU/min/mg protein in roots, whilst the ribulose bisphosphate carboxylase oxygenase (RuBisCO) small subunit (rbc) promoter in pBI131 produced 1525 pmole 4-MU/min/mg protein in leaves. The levels of transcription driven by APG in anthers during microspore development are thus particularly high, and highlight the extreme degree of metabolic activity involved in gametogenesis. Since no other genes which are

expressed in a similar manner to APG have been characterized by *gus* fusion analysis, no direct comparison of expression in the gametophyte between APG and other promoters can be made, though high levels of expression are typical of tapetum-specific promoters. Strong levels of expression tend to be indicative of genes coding for structural proteins, which by definition need to produce large amounts of protein, rather than of genes coding for proteins with an enzymatic function, though as discussed above, the temporal and spatial expression of APG presents a confusing picture regarding this matter.

6.3.3 APG-*Ac* Fusions

Reporter gene experiments incorporating the APG promoter demonstrated significant transcriptional activity during microspore development. The APG promoter thus proved a potentially good candidate for use in transposition experiments, as was suggested previously. One factor which is unclear regarding the suitability of APG to drive transposition of *Ac* and *Ds* is the exact relationship between APG expression and DNA replication. It is not known when during the development of the haploid microspore DNA replication occurs, but it is clear that APG is only expressed in the later stages of development. Thus, if DNA replication were to take place soon after microspore release, then transposase under the control of APG would not be present at the appropriate time to promote high levels of transposition. The constructs described in the results section are designed to produce maximum levels of *Ac* transposase enzyme in the developing gametophyte of plants carrying an APG-*Ac* transgene, enabling *trans* -activation of *Ds* elements prior to or during microspore mitosis. The use of these constructs will be discussed in more detail in the following chapter.

6.4 Summary

The APG promoter drives transcription in specific sporophytic cell types of the anther and the gametophyte during microspore development and early pollen maturation. This gene expression is regulated in a complex temporal and spatial fashion. Overall expression is maximal during late microspore development, but falls dramatically when microspore mitosis takes place. Expression is maintained only in the gametophyte after this cell division, but ceases during the interphase preceding the pollen mitosis. Maximal expression levels are extremely high compared to other "model" promoters such as CaMV 35S. Although the majority of the specific GUS activity detected originates from sporophytic tissues, significant activity is found in the gametophyte around the time of microspore mitosis, and the APG promoter might thus prove useful for driving germinal transposition of engineered transposable elements.

CHAPTER 7

DISCUSSION

7.1 General Discussion

The primary objective of the studies described in this thesis is to improve the accessibility of transposon tagging as a means to identify plant genes of interest. In the first instance, the species *Linum usitatissimum* (flax) was chosen as a host plant, and the *Ac*/*Ds* transposable element family chosen as potential tags. The *Ac* and *Ds* elements had previously been shown to be active in other plant species besides *Z. mays*, their source, and activity in *L. usitatissimum* is demonstrated in this thesis. However, during the course of these and similar studies by other workers, it became apparent that there were major inadequacies in current transposon tagging strategies. One of these was the difficulty in obtaining plants with an inserted transposed element at the same locus in all cells of the plant. This problem can be overcome by selecting geminal excision events, which lead to the production of plants carrying the same transposition event in all cells. Natural germinal excision frequency is low, but this thesis describes the cloning of an anther-specific gametophyte-expressed gene which may facilitate increased germinal excision frequency and somatic stability when fused with transposase genes.

7.2 Transposition in *L. usitatissimum*

The studies presented in Chapter 3 of this thesis indicate that the *Z. mays Ac* element is active in callus tissue produced following transformation of *L. usitatissimum* hypocotyls with *Agrobacterium*. However, this activity appears to be limited to excision (at a moderately high frequency of 35%) of the transposon from the introduced T-DNA, since no reintegrated elements were detected. This second finding is contrary to the observed trend for *Ac* in other transgenic plant species, where all major published studies report reintegration events associated with a significant proportion of excisions (reviewed by Haring *et al.*, 1991). This review also reports however, that in some species, for example lettuce, *Ac* transposition is very difficult to detect. Thus, the frequency of excision of the *Ac* element varies from 1-80%, depending upon the assay employed, the species and individual, and the stage of development assayed. Since an excision frequency of 35% as found in *L. usitatissimum* is reasonably high, there remains a discrepancy between excision and reintegration in the callus system employed in this study.

This may result from a number of factors. Firstly, and perhaps most likely, since only nine independent lines were analysed, it is quite possible that if reintegration were only 10-20% for example, then by random chance, nine lines which did not contain reintegrated elements could have been selected, although reinsertion had occurred in a number of other lines. Other possible explanations are more complex, and relate to the nature of *L. usitatissimum* tissue culture. The experiments described in Chapter 3 rely on the DNA status of callus cells to approximate that in differentiated plant tissues to permit a good assessment of how transposable elements may behave in whole plants. Clearly this is not the case, as is understood from the universally observed phenomenon of somaclonal variation (Evans, 1989). Such variations may be especially severe in *L.*

usitatissimum, as the genome of this species is particularly susceptible to changes induced by various environmental factors (e.g. Durrant, 1971; reviewed by Cullis, 1986). Lines with unusual phenotypes arise commonly after environmental changes and are known as genotrophs. Cullis and Cleary (1986) showed that normal *L. usitatissimum* plants, and particularly genotrophic lines, exhibited significant levels of DNA rearrangement in callus tissues derived from them. Such instability of the DNA in *L. usitatissimum* callus cultures may well affect reintegration frequencies of excised elements such as the introduced *Ac* element in these studies.

The discussion of the affect of tissue culture on transposition can be extended further, since it has been shown that tissue culture can increase the activity of resident transposable elements in *Z. mays* (Peschke *et al.*, 1987). For this reason, one might argue that a callus system represents a situation where transposition may be at a high level, making possible the selection of a large number of excision events. An investigation of the literature infers that this may not necessarily be true however. Peschke *et al.* (1987) argue that although most somaclonal variation is caused by chromosome breakage and associated rearrangements, significant variability must be caused by other factors, such as transposon activity. These two mutagenic events need not be separate phenomena however. Many genetic studies of transposon activity have recognized that elements can undergo reversible inactivation (for a review see Dellaporta and Chomet, 1985). Molecular studies have shown that inactivation is associated with methylation of transposon DNA sequences, and that conversely, reactivation can be a result of demethylation (e.g. Schwartz and Dennis, 1986; Chomet *et al.*, 1987). Reactivation of such "cryptic" elements is often associated with chromosome breakage, be this as a result of ionizing or ultra-violet radiation (Doerschug, 1973), infection by BSMV (barley stripe mosaic virus) or WSMV (wheat streak mosaic virus) (Sprague *et al.*, 1963) or by the breakage-fusion-bridge cycle, as originally described by McClintock (1950). Thus transposon activity induced by tissue culture is probably a secondary effect of chromosome breakage events which result from culture. If this activation results from a DNA demethylation process associated with replication and repair mechanisms, then tissue culture activation would not apply to hypomethylated elements recently introduced by transformation. If, on the other hand, DNA modification and reactivation by chromosome breakage were independent phenomena, then one would be correct in the original assumption that selecting for excision in callus tissues might be the most efficient way of obtaining lines containing transposed elements.

It is clear that the selection of *L. usitatissimum* cell lines containing T-DNAs from which the *Ac* element excised early during the callus generation procedure is not likely to represent the best method for obtaining plants containing loci mutated by transposon insertion. Possible reasons for the failure to detect reintegrated elements in *L. usitatissimum* callus are discussed above, and analysis of more lines containing excised

elements may be revealing. It might also be informative to investigate transposition in transformed and regenerated *L. usitatissimum* plants. In the present study, no major effort was made to obtain whole regenerated plants, though regeneration has been reported (Basiran *et al.*, 1987). Before initiating any transposon tagging experiment in *L. usitatissimum* using the *Ac* element, it would be necessary to generate stably transformed *L. usitatissimum* plants containing intact *Ac*-reporter gene constructs, and to assess the behaviour of the element in differentiated tissues.

7.3 Microspore-Specific Messenger RNAs

The aims of the thesis were changed following the preliminary work with *L. usitatissimum* to investigate an approach to improve the efficacy of transposon tagging in model systems. To exploit existing knowledge in the laboratory, it was decided to attempt to express transposase in the male gametophyte to increase germinal transposition frequency. In order to obtain the promoter of a gene active specifically during microspore development, which would be suitable for use in the proposed transposition experiments, the expression of anther-specific cDNAs of *B. napus* was investigated. Four cDNA clones which hybridized to RNA from microspores and immature pollen grains were initially selected for analysis. Northern blotting experiments using staged bud RNA preparations showed that the genes corresponding to the four clones were all expressed in a similar fashion, in that messages were first detectable after meiosis, with levels reaching a peak during microspore development before declining during pollen maturation. Three of the cDNAs hybridized to mRNAs with almost identical expression patterns, whilst the fourth represented a gene expressed slightly later in development. Interestingly, all four of these microspore-expressed mRNAs were not detectable in RNA extracted from buds of the cytoplasmic male sterile variety, Ogura. This was also found to be true of two related cDNAs, CEX and CEX2, which were subsequently isolated and also showed expression during microspore development, in this case with a pattern of mRNA accumulation intermediate between the patterns observed for the first four clones. The lack of expression of all six genes in the cms variety Ogura was taken to suggest that the messages were microspore-specific. In view of the low proportion of anther-expressed genes which have been estimated to be microspore-specific (see Chapter 1) this was perhaps a surprising result, and, as was later demonstrated, the assumption that only microspore-specific genes would not be expressed in Ogura anthers was an over-simplification. The analysis of the APG gene promoter detailed in Chapter 6 showed that although the *Arabidopsis* and *B. napus* APG genes were initially classified as microspore-specific (as discussed in Chapter 5), APG expression is also found in other parts of the anther. Similarly, a *B. napus* cDNA termed E2 (first reported by Scott *et al.*, 1991a) which is also expressed in the microspores of fertile plants, but not in male-sterile plants, has been shown to be expressed in the tapetum as well as the microspores (Foster *et al.*, manuscript submitted). Indeed, many

forms of cms are first manifest in the tapetum rather than the gametophyte, such as *Petunia* cms mutants and *Z. mays* Texas cytoplasm (Bino, 1985; Warmke and Lee, 1977). Thus, the mRNAs corresponding to cDNAs #14, #17 and G2 may not necessarily be microspore-specific. This information leads one to conclude that the Ogura cms factor affects gene expression in all tissues of the anther, even though the only phenotypically altered cell type is the gametophyte. Molecular evidence has shown that the Ogura cms phenotype is caused by mitochondrial DNA rearrangements, which particularly affect transcription of the *atpA, atp6* and *coxI* ORFs (Makaroff and Palmer, 1988). How such changes might influence the expression of specific genes such as those identified in this study is unknown. Since *Arabidopsis* genomic clones which clearly corresponded to the anther-specific homologues in *B. napus* were not obtained for the #14, #17 and G2 cDNAs, no further work was carried out with them.

The I3 cDNA probe, however, did facilitate the isolation of an *Arabidopsis* anther-specific gene. For this reason I3 was analysed in detail. Of the four initial cDNA clones, I3 was the one which is most likely to be truly microspore-specific. In addition to it's lack of expression in the cms variety, I3 is strongly expressed later in development than the other genes, mRNA levels being maximal in buds of 3-5 mm in length. At this stage the tapetum is degenerating, and is virtually absent in buds of 4-5 mm. I3 is thus unlikely to be expressed in the tapetum, where the highest rates of transcription within the anther are found. Another line of evidence which suggests that I3 is microspore-specific comes from sequence analysis. The prediction of the amino acid sequence of I3 enabled the protein to be identified as being homologous to seed oleosin proteins. These oil body membrane proteins have to date been studied only in seeds, which in many plants, including *B. napus*, store large amounts of lipid for metabolism during germination. *B. napus* pollen grains are also known to store significant reserves of lipid in oil bodies (Evans *et al.*, 1990b) and one can envisage a parallelism between the structures of oil bodies in seeds and in pollen. However, the structure of I3 is significantly different from the seed oleosins of *B. napus* and other species, suggesting that it is a member of a different gene family to the genes expressed during embryogenesis. Evans *et al.* (1990a) report that different lipid biosynthesis pathways are employed in the gametophyte and sporophyte, primarily the tapetum, lipids being stored in elaioplasts rather than oil bodies in the tapetum. That a function has been assigned to the I3 gene product which is presumably required only in the gametophyte is indicative of the transcript being microspore-specific. Future studies involve further attempts to isolate the I3 gene from *B. napus* or *Arabidopsis*.

It is worth considering at this point the implications for transposition experiments of the expression of genes in other anther tissues in addition to the microspores. The primary aim of isolating a microspore-specific promoter was to enable the construction of a recombinant transposase gene, which would drive transposition of a *Ds* element in developing microspores but not in the vegetative organs of the plant. This would permit

the production of seeds carrying a range of unique stable insertions, which could be analysed without the complications of somatic activity. These requirements would still be met if the promoter was also active in the sporophytic tissues of the anther, since the use of such a promoter would not be detrimental to the experimental design. Indeed, any gene which is expressed in the gametophyte at the appropriate stage of development might then be used to express transposase, as long as any additional expression was limited to terminal organs not contributing to the germ cells, such as the petals, ovary wall *etc.*

7.4 *Arabidopsis* Genes Homologous to Anther-Specific cDNAs

The four original microspore-expressed cDNA clones, #14, #17, G2 and I3, were used as probes to screen a genomic library of *Arabidopsis thaliana*, a process which resulted in the isolation of clones hybridizing to each of the cDNAs except #17. This preliminary result supported the expectation that as members of the same phylogenetic group, *Arabidopsis* and *B.* would possess genes homologous at the nucleotide sequence level. However, complications in their relationships became apparent upon subsequent analysis. The *Arabidopsis* G2 coding region was located on a 1 kb *Bam*HI restriction fragment, but when used as a probe in northern blot experiments, this sequence did not hybridize to an anther-specific message. Such a small region of DNA is unlikely to include more than one gene, but presumably the *Arabidopsis* G2 gene hybridizes to *B. napus* mRNAs to which the cDNA probe does not. Whether or not this is because the *Arabidopsis* genomic clone is not in fact the equivalent of the *B. napus* G2 gene was not determined. Similarly, of a family of four #14-homologous genes isolated from *Arabidopsis*, none showed specific hybridization to any differentially regulated transcript in *B. napus* buds. Again, with probe fragments ranging in size from 0.7 to 2.0 kb, one would not anticipate more than one coding region to have been labelled. *Arabidopsis* did not, therefore, turn out to be as simple a system as was supposed with regard to cloning the G2 and #14 genes. The same library has nonetheless been used by Scott and co-workers to successfully isolate *Arabidopsis* genes corresponding to various other *B. napus* anther-specific cDNAs.

The cDNA which did hybridize to an anther-specific *Arabidopsis* gene (the APG gene) was the I3 oleosin cDNA. Northern blotting experiments and sequence analysis however, indicated that APG was not an oleosin gene, but encoded a novel proline-rich protein. APG and I3 are not related in any way other than in a short stretch of coincidental nucleotide similarity which enabled detection by hybridization with labelled cDNA probe. The fact that APG and its *B. napus* homologues, APG1 and APG2, are specifcally expressed in the anther during microspore development would appear entirely fortuitous. The predicted amino acid sequence of the APG protein included many features which suggested a possible role in cell wall architecture (see Chapter 5), the most significant of these being a signal peptide sequence, the high proline content and its arrangement in repeated

sequences.

gus gene fusion analysis established that in transgenic *N. tabacum*, the APG promoter is active in several cell types of the anther, namely the microspores, tapetum, stomium, and anther wall, and that expression is regulated in a complex developmental manner. The expression pattern of the APG gene is unique amongst those of other anther-specific genes. Of the genes which have been identified to date that are expressed during the microspore development phase, all are confined either to the sporophyte or to the gametophyte. Of those specific to the sporophyte, many are tapetum-specific (Koltunow *et al.*, 1990; Smith *et al.*, 1990; Scott *et al.*, 1991a) whilst the *N. tabacum* TA20 and TA56 genes are expressed in the anther wall, endothecium and epidermis, and the connective and stomium respectively (Koltunow *et al.*, 1990). Another group of genes are expressed after microspore development in mature and germinating pollen and in several instances additionaly in the anther wall. These include the polygalacturonase gene, P2, of *Oenothera organensis* (Brown and Crouch, 1990) and the Zmg13 gene of *Zea mays* (Hamilton *et al.*, 1989) which are expressed in maturing and germinating pollen grains, and LAT52, LAT56 and LAT59 from *L. esculentum* (Ursin *et al.*, 1989) which are expressed in confined regions of the anther wall and maturing and germinating pollen grains. The *B. napus* genes described by Albani *et al.* (1990; 1991) are maximally expressed during microspore development and were concluded to be microspore-specific, though no evidence to eliminate the possibility of expression in other anther cell types was presented. Another cDNA, Bcp1, which was isolated from *B. campestris* appears to hybridize to message(s) present during pollen maturatrion in both the tapetum and in bi- and tri-nucleate pollen (Theerakulpisut *et al.*, 1991). Thus APG represents a novel class of gene, which is expressed in both male gametophyte and sporophyte during earlier stages of microgametogenesis than previously described genes. There are several functions involved in microspore development which may require expression of the same or similar genes in both the tapetum and microspore, such as the production of sporopollenin for example. As mentioned in Chapter 6, a second cDNA, E2, has been isolated from *B. napus* which also hybridizes to transcripts present in both the microspore and tapetum during microspore development, though these messages may not be encoded by the same gene. This also applies to the expression of many of the clones described above, the expression patterns of which have been elucidated by means of *in-situ* hybridization. In the case of APG, the reporter gene fusion experiments prove that the same gene is active in both gametophyte and sporophyte.

Interestingly, a number of proline-rich proteins are known to be localized in the epidermal cell layer of *Helianthus annuus* (sunflower) anthers (Evrard *et al.*, 1991) but these proteins do not bear any significant similarity with the APG and CEX sequences. That the timing of expression in the microspores and tapetum is concurrent with intine deposition is in part consistent with the contention that the protein may be sequestered in the microspore wall as a structural molecule, though whether protein secreted from

the tapetum might reach the intine is unclear. Preferential expression in the stomium might also be interpreted to infer a structural role, as the cells of the stomium may have a wall composition which is different to that of the rest of the anther in order that it be specifically degraded to facilitate dehiscence. Other genes have been demonstrated to be expressed in stomium cells at later stages of development, such as the TA56 *N. tabacum* thiol peptidase (Koltunow *et al.*, 1990) and possibly the LAT56 and LAT59 *L. esculentum* pectate lyases (Wing *et al.*, 1989) but these all have putative hydrolytic activities and are therefore presumed to play a direct role in the dehiscence process. The only two genes which have thus far been reported to be preferentially expressed in the developing microspore are the Bp4 and Bp19 genes of *B. napus* (Albani *et al.*, 1990;1991). Although the function of the Bp4 genes is unknown, the occurrence of a putative signal peptide and a high content of lysine and cysteine in the products of the family led the authors to suggest a possible structural function for these proteins. The Bp19 gene appears to encode a protein with pectin esterase activity. This gene encodes a protein with a signal peptide sequence plus an additional large N-terminal extension, leading to the suggestion that the protein is stored in an inactive form in the pollen wall to fulfil some function during the germination of the pollen tube. Thus a number of anther-specific genes with either structural or hydrolytic functions have been found to be expressed in a broadly similar manner to the APG gene. Presumably these two types of protein are representative of highly expressed classes of genes, which are thus those identified in differential screening of cDNA libraries.

7.5 A Microspore-Expressed *Ac* Transposase

Investigations of *Ac/Ds* behaviour in transgenic plants have been carried out in detail in a number of different species, as discussed in Chapter 1. In each of these studies, either wild type *Ac* elements were used, or modified *Ac* elements were created by deletion of various regions of the transposase gene. Whilst several groups are well advanced in the study of the affect of expressing *Ac* transposase from heterologous promoters such as CaMV 35S, *nos*, *ocs* and a number of inducible or developmentally regulated promoters, to date, no results from these studies have been published. It has been demonstrated, however, that the use of such recombinant genes does alter the pattern of somatic and germinal transposition of the *Ac* element or *trans*-activated *Ds* elements. This finding is in accord with the published observation that in *N. tabacum*, *trans*-activation of a *Z. mays dSpm* (suppressor-mutator) element via a CaMV 35S-*Spm* transposase gene alters the pattern of transposition compared to that seen using a stabilized wild type *Spm* transposase helper (Masson and Fedoroff, 1989).

It has been speculated (see Chapter 1) that production of transposase enzyme to promote optimal germinal transposition is required in microspores in the interphase preceding microspore mitosis. This is because *Ac* and *Ds* tend to be active mainly during DNA replication. Transposition prior to meiosis would be less useful, since a smaller number

of independent transposition events would be transmitted into the next generation, though this statement assumes that frequencies would be equal for pre-meiotic and post-mitotic transposition. Genes expressed in the generative nucleus before the pollen mitosis might theoretically be useful, but it is believed that very little transcription occurs in the generative cell other than that required for the process of mitosis itself, and it would thus be very difficult to clone genes active at this point in development. This is one major reason why genes active during microspore development were targeted. The efforts described in this thesis resulted in the cloning of one microspore-expressed gene (APG) but this gene was found not to be microspore-specific. The APG gene is, however, anther-specific, and expression of transposase in the sporophytic tissues of the anther should not interfere with the benefits provided by the intended system, *vis-a-vis* the generation of independent germinal transpositions and subsequent transposon stability during vegetative growth. Quantitative analysis of ß-glucuronidase reporter gene activity under the control of APG promoter fragments established that in transgenic *N. tabacum*, expression occurred prior to and after microspore mitosis, having risen from undetectable levels at microspore release. Thus activity of the APG promoter exhibits the attributes required for the control of germinal transposition, and work to test APG-*Ac* fusions, as described below, is underway at the time of writing.

It is true that the promoters of the *B. napus* Bp4 and Bp19 genes identified by Albani *et al.* (1990; 1991) would also fulfil the requirements for dirving germinal transpositions, but these genes are protected under patent, and thus were not available for use in this study.

7.6 Future Work

7.6.1 Pollen Oleosins

The cloning of the microspore-specific I3 oleosin brings to light a potential new area of research in lipid biochemistry. Whilst the molecular biology of the oil bodies of seeds has been studied in detail because of the economic importance of the variability in lipid composition, to my knowledge, no work on their counterparts in pollen grains has been undertaken. This may be worthwhile in the future, since several groups have shown a biochemical correlation between the lipid composition of pollen grains and the seeds of sibling plants (e.g. Evans *et al.*, 1987, 1990a). The identification of the I3 gene now provides a starting point to begin such a study. Initially it will be necessary to establish more definitely whether or not the I3 protein is synthesized only in pollen, and not in seeds, as is suggested by the evidence collected thus far. Five oleosin isoforms have been purified from *B. napus* seeds and peptide sequenced, but all are significantly different to the predicted I3 protein (Dr. D. Murphy, personal communication) which at 15 kDa is smaller than previously identified oleosins. Straightforward Southern blot analysis employing a range of washing stringencies should show the relatedness of I3 to other genes in *B.* species. It should be possible to conclusively distinguish between the

expression of seed oleosins and I3 In northern analysis by the use of a gene-specific probe. For example, the 3-prime end of the I3 cDNA, which contains only the sequence coding for the non-conserved C-terminus and the untranslated region of the mRNA should only hybridize to I3 message. The I3 cDNA sequence conveniently contains a *Pst* I restriction site which would enable the simple production of such a probe. These questions might be approached concurrently by Western blotting experiments. A number of antisera specific to oleosins are available, which were raised to proteins purified from the oil bodies of various species, including *B. napus.* Oleosin antibodies have been demonstrated to cross-react with related isoforms from a wide range of species, and the same study also showed that there are two immunologically distinct classes of high and low molecular weight oleosins in all species investigated (Tzen *et al.*, 1990). An examination of the numbers and classes of oleosin isoforms in protein extracts from microspores and pollen should therefore be possible using antisera raised to seed oleosins. Banding patterns obtained in such experiments could be compared to those seen in protein extracts from seeds (preferably included on the same Western blot) to distinguish directly between isoforms expressed in the different tissues. It is intended that such experiments will be carried out in the near future in collaboration with Dr. Dennis Murphy of the 'B. and Oilseeds Research Group' of the Institute for Plant Science Research, Norwich, UK.

A further line of experiments, which may provide proof of the *in vivo* function of the I3 gene product, are over-expression studies. The introduction into plants of an APG-I3 gene fusion, for example, might affect the production of oil bodies in pollen. Since APG is expressed before I3, one might observe premature or abberant formation of oil bodies, which can easily be visualized using the histochemical stain Nile red. Similarly, expression of I3 oleosin from a tapetum-specific promoter may provide information on lipid metabolism in this tissue by restricting the amount of lipid available for normal functions, for example sporopollenin production, which probably requires lipid precursors. Similar experiments using the CaMV 35S promoter might also create interesting phenotypic affects in other plant organs.

7.6.2 Expression of the APG Gene

The results presented in Chapter 6 showed that the APG promoter regulates a complex temporal and spatial expression regime in the anther, and that the elements responsible for this activity are located within the -530 region of the APG upstream sequence. All the conclusions which were drawn regarding the pattern of APG expression come from GUS assays on transgenic *N. tabacum*, but before further work is undertaken it is important that the constructs are tested more comprehensively in transgenic *Arabidopsis.* Following this, further experiments may be undertaken to resolve the question of the complexity of regulation of the promoter in terms of the number of *cis* -acting sequences present in the -530 bp region. One approach would be to investigate

the activity of further *gus* fusions with different fragments of promoter DNA in transgenic tissues, thus narrowing down the possible locations of individual sequences which bind specific *trans*-acting factors. Typically in such studies, successive deletions to the 5-prime and 3-prime ends of full length promoters are produced, which are then fused to reporter genes and assayed either transiently or in stably transformed plants for loss or reduction of activity. In constructs exhibiting loss/reduction of function, one can deduce that the deleted region is important for gene expression under the particular assay conditions. This type of approach is commonly used to answer similar questions, and was for example successfully employed by Twell *et al.* (1991) to identify *cis* -acting regulatory elements required for correct activity of the promoters from the *L. esculentum* pollen-expressed LAT52, LAT56 and LAT59 genes. Twell *et al.* (1991) were able to identify several short nucleotide sequences which are necessary for pollen-specific expression of these genes, and demonstrated that one of these sequences, a 19 base pair fragment of LAT52 promoter DNA, was sufficient to enhance expression of the heterologous CaMV 35S promoter in pollen.

There is an additional possibility that the APG promoter may not be active in precisely the same fashion in *N. tabacum* as in *Arabidopsis* and *B. napus*, or that *N. tabacum* GUS assays did not produce an entirely true representation of the temporal variation in APG message levels in the Brassicacaea, due to the long half life (around 2 days) of GUS enzyme. The close relationship between the *Arabidopsis* and *B. napus* APG genes provides an opportunity to examine the expression of the gene by *in situ* localization of mRNA in thin sections of *B. napus* buds from different developmental stages should this prove necessary.

7.6.3 Germinal Transposition

Determination of the activity of *Ac* transposase during microspore development, as monitored by *trans*-activation of *Ds* elements, is the most pressing matter arising from this thesis. Despite the similarity of the behaviour of the -530 and -1501 bp promoters as monitored by *gus* fusion analysis, it was felt prudent to construct an *Ac* fusion with both promoters, in case subtle differences in expression in the microspores may have a bearing on transposition frequency. The APG/*Ac* construct described in Chapter 6, containing the 1.5 kb promoter, has been introduced into *N. tabacum* and *Arabidopsis* plants via *Agrobacterium*-mediated transformation at the time of writing. Work will also be in progress to transform plants with a second fusion, between *Ac* and the 0.53 kb promoter. Plants harbouring these two constructs will be used as preliminary starting material to test transposition capability and specificities, but will also need to be made homozygous to prevent meiotic segregation of the transposase allele before definite estimations of frequency can be made. Transgenic *N. tabacum* plants containing a *Ds* element inserted into a reporter gene have been kindly made available by Dr. J.D.G. Jones of the IPSR Sainsbury Laboratory, Norwich, UK. These plants carry a

construct (SLJ1615) which includes an engineered *Ds* element inserted into the untranslated leader sequence of a chimaeric CaMV 35S streptomycin phosphotransferase (SPT) gene. Excision of the *Ds* element restores streptomycin resistance. SPT was developed as a marker for plant transformation (Jones *et al.*, 1987) because it confers cell autonomous resistance to the antibiotic streptomycin, which causes bleaching of chloroplasts in sensitive cells. The gene was subsequently used to monitor germinal and somatic excision of the *Ac* element in *N. tabacum* (Jones *et al.*, 1989). Germination of seedlings on a medium containing streptomycin allows visual detection of green sectors of resistant cells, each sector deriving from a cell in which the transposable element excised from the T-DNA. Seedlings produced after a germinal excision in the parent can easily be identified since they are fully green.

If preliminary investigations are successful, and *Ds* elements are shown to be *trans* -activated by the APG/*Ac* transposase, then an estimation of transposition frequency must be made. To do this, it will be necessary to make plants carrying the two T-DNAs homozygous for these loci. This is achieved by firstly selfing heterozygotes and collecting their seed, which for single locus insertions will comprise kanamycin resistant and sensitive seeds in the ratio of 3:1, depending upon the presence or absence of a T-DNA copy. Of these resistant seeds, one third will be homozygous, the other two thirds being heterozygous, following classical Mendelian genetics. The homozygotes can be identified by growing and selfing resistant seedlings, and collecting seed from individual plants. When germinated on kanamycin, seeds of a homozygote will show 100% resistance, whilst those of heterozygotes will segregate 3:1. Homozygous lines can subsequently be maintained by repeated selfing. It would be wise to select several independent transformants for the *Ac* gene to work with, so that any ambiguities in transposase synthesis, caused by position effects for example, might be recognized. Once homozygous lines are obtained carrying both *Ac* transposase and *Ds* reporter constructs, these plants can be crossed with each other to produce double heterozygotes containing both T-DNAs (see Fig. 7.1). By selfing these plants, one induces transposition. However, because these plants are now heterozygous for each T-DNA, segregation will occur during meiosis, and only 25% of microspores will possess both constructs (Fig. 7.1). Thus the maximum frequency of production of streptomycin resistant seedlings is only 25% of the total seed population following selfing. Ideally, one would like eventually to produce a single T-DNA including both the *Ac* transposase and *Ds* reporter genes, so that segregation of the two constructs could not occur during microsporogenesis. Alternatively, for many applications, an autonomous microspore-active *Ac* element may be useful. This could be produced by replacing the wild type *Ac* promoter with the APG promoter, or possibly by introducing an APG/*Ac* fusion into a *Ds* element, which would then activate itself to enable germinal transposition from a single T-DNA.

The production of a functional anther-specific transposase will have major benefits to

Figure 7.1

PROCEDURE TO ENABLE
GERMINAL EXCISION DRIVEN BY APGAc

1. Identify homozygous transgenic plants, preferably containing single copy T-DNAs.

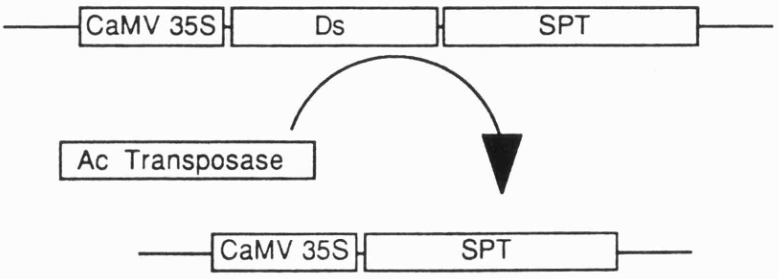
2. Cross these plants:

APGAc / APGAc ____ / ____ X DsSPT / DsSPT ____ / ____

APGAc / ____
DsSPT / ____

3. Selfing these plants leads to segregation of the T-DNAs such that 25% of microspores posses both Ac and Ds constructs.

4. In spores carrying both constructs there is potential for Ds transposition during gametogenesis.



Allows expression of streptomycin phosphotransferase and thus selection of resistant seedlings.

workers currently using transposable elements. The main use is likely to be in transposon tagging experiments, as discussed in Chapter 1. Here, the production of many plants containing different transposed *Ds* elements would greatly decrease the need to produce large numbers of transformed plants, or to select F2 plants containing germinal transpositions from sectored plants. The stability of the resulting insertions during vegetative growth is also important, as this feature may simplify the cloning of disrupted genes. Basic research into the behaviour of transposable elements should also benefit. For example, by mapping the site of insertion of transposed elements following germinal transposition from a single locus, it will be possible to learn more about the distance and direction of *Ac* and *Ds* transposition. Of course, studies need not be limited to *Ac* and *Ds*, since the same promoter may work equally well when fused to different transposase genes to drive transposition of the corresponding non-autonomous elements.

7.7 Summary

Investigations were carried out to determine the feasibility of a transposon tagging experiment in *Linum usitatissimum*. The excision of the *Z. mays* transposable element *Activator* (*Ac*) from the genome of transgenic *L. usitatissimum* callus was demonstrated, whilst a *Dissociation* element (*Ds*) was found to be stable. However, reintegration of excised *Ac* elements was not detected, and this barrier to gene tagging led to an examination of procedures which might improve the general applicability of transposon tagging.

A recombinant *Ac* transposase gene was constructed in order to achieve a high germinal transposition frequency in transgenic plants; this feature is an essential component of an efficient transposon tagging strategy. The *Ac* construct was produced by fusing the promoter of an anther-specific gene to the transposase coding region. The anther-specific gene, APG, was cloned from *Arabidopsis thaliana*, following the identification of four putative microspore-specific mRNAs from *B. napus*. Of these mRNAs, one, termed I3, was analysed in detail and found to encode a novel oleosin protein, and was apparently confined to the developing microspore. The I3 cDNA was used as a molecular probe to clone the APG gene, which encodes a proline-rich protein of unknown function. A small gene family encoding proteins highly homologous to the APG protein was identified in *B. napus* via the isolation of three cDNAs termed CEX1, CEX2 and CEX6.

Promoter fragments of the APG gene were demonstrated to drive expression of a ß-glucuronidase reporter gene in the male gametophyte, tapetum, stomium and anther wall of *Nicotiana tabacum* and *Arabidopsis* during the microspore development stage of gametogenesis. It is hoped that restriction of transposition to these cells will permit the production of a seed population containing a wide range of unique, transposon inserts which will be stable during vegetative growth.

APPENDICES

Appendix I : Bacteriological Media

Liquid media

NB (nutrient broth) was purchased as a pre-made powder from Difco Laboratories and dissolved in water and sterilized as per their instructions.

NZY:	per litre -	NZ Amine*	10g
		Bacto-Yeast extract	5g
		NaCl	5g
		Caesin amino acids	1g
		$MgSO_4$	2g

(*NZ Amine is type 2 caesin hydrolysate)
Made to pH 7.5 with sodium hydroxide and sterilised.

2xYT:	per litre -	Bacto-trypone	16g
		Bacto-Yeast extract	10g
		NaCl	5g

Made to pH 7.0 with sodium hydroxide and sterilized.

Solid Media

All media was solidified by the addition of Bacto-Agar to 1.5%.

Appendix II : Plant Tissue Culture Media

MSO: MSO salts as described by Murashige and Skoog (1962) were purchased pre-prepared from Flow Laboratories, dissolved in distilled water with 30 g/l sucrose and made to pH 5.8 with 0.1 M HCl. Oxoid Technical Agar was added to 0.8% and dissolved by heating in a steamer before sterilization.

MSD4x2: As MSO with the addition of:

0.1 mg/l naphthalene acetic acid
1 mg/l 6-benzyl amino purine

Hoaglands' Solution:

Component	Concentration (g/l)
KNO_3	101.1
$Ca(NO_3)_2.4H_2O$	236.16
$NH_4H_2PO_4$	115.08
$MgSO_4.7H_2O$	246.49
KCl	3.728
H_3BO_3	1.546
$MnSO_4.H_2O$	0.338
$ZnSO_4.7H_2O$	0.575
$CuSO_4.5H_2O$	0.125
H_2MoO_4	0.081
FeEDTA	6.922

Appendix III : Molecular Biology Buffers and Solutions

TE : 10 mM Tris-Cl pH 7.5, 1 mM ethylenediamine tetracetic acid (EDTA)

20x SSC : 3 M NaCl, 0.3 M sodium citrate

20x SSPE : 3.6 M NaCl, 0.2 M sodium phosphate, 0.02 M EDTA, pH 7.7

50x Denhardts : 1% BSA, 1% Ficoll 400, 1% PVP

Phenol/Chloroform : Phenol:Chloroform:Isoamyl Alcohol, 25:24:1 plus 0.1% 8-hydroxyquinoline, equilibrated and overlayed with 0.1 M Tris-HCl pH 8.0.

10x MOPS Buffer : 0.2 M 3-[N-morpholino]-propane sulphonic acid, 0.05 M sodium acetate pH 7.0, 0.01 M EDTA

SM : 0.5 M Tris-HCl pH 7.5, 0.1% gelatin, 0.1 M NaCl, 15 mM $MgSO_4$

Depurinating solution : 0.25 M hydrochloric acid

Denaturing solution : 0.5 M NaOH, 1.5 M NaCl

Neutralizing Solution : 3.0 M NaCl, 0.5 M Tris-Cl pH 7.4

DNA hybridization solution : 50% formamide, 50 mM NaH_2PO_4, 5x SSC, 5x Denhardts, 0.2 mg/ml sheared herring sperm DNA

RNA hybridization solution : 50% formamide, 5x SSPE, 5x Denhardts, 0.5% SDS, 0.2 mg/ml sheared herring sperm DNA

Wash solution A : (3x SSC, 0.5% SDS)

Wash solution B : (0.5x SSC, 0.5% SDS)

Oligolabelling stop solution : 20 mM NaCl, 20 mM Tris-Cl pH 7.4, 2 mM EDTA, 0.25% SDS

RNA denaturing solution : 65% formamide, 8% formaldehyde, 1.3x MOPS

2x CTAB : 0.1 M Tris-Cl pH 8.0, 1.4 M NaCl, 20 mM EDTA, 2.8 µl/ml β-mercaptoethanol, 2% CTAB

10% CTAB : 10% CTAB, 0.7 M NaCl

5x Ligation Buffer : 0.25 M Tris-HCl pH 7.6, 50 mM $MgCl_2$, 50 mM dithiothreitol, 250 µg/ml BSA, 5m M ATP

CIP buffer : 50 mM Tris-HCl pH 8.0, 0.1 mM EDTA, 20 mM $ZnCl_2$, 0.5 mM $MgCl_2$

Oligolabelling Buffers :

Oligolabelling Buffer A : Combine 625 µl 2 M Tris-HCl pH 8.0, 25 µl 5 M $MgCl_2$, 350 µl dH2O, 18 µl ß-mercaptoethanol, 5 µl each of 3 mM dATP, dGTP, dTTP

Oligolabelling Buffer B : 2 M HEPES pH 6.6

Oligolabelling Buffer C : random hexadeoxyribonucleotides (Pharmacia) suspended in 3 mM Tris-HCl, 0.2 mM EDTA pH 7.0 at 90 OD_{260} units/ml.

OLB is made by mixing buffers A,B and C in the ratio 2:5:3.

Gel loading buffer : 0.25% bromophenol blue, 0.25% xylene cyanol, 25% Ficoll (type 400)

10x TAE : 0.4 M Tris-acetate pH 8.0, 10 mM EDTA

Lysis buffer (25 mM Tris-Cl pH 8.0, 10 mM EDTA, 0.5 M glucose)

Alkaline SDS : 0.2 M NaOH, 1% SDS

10x TCM : 0.1 M Tris-Cl pH 7.5, 0.1 M $CaCl_2$, 0.1 M $MgCl_2$

10x PCR buffer : 67 mM Tris-Cl pH 8.8, 16.7 mM ammonium sulphate, 6.7 mM $MgCl_2$, 10 mM β-mercaptoethanol, 6.7 µl EDTA pH 8.0, 1.5 mM each dNTP, 170 µg/ml BSA

10x TBE : 121 g Tris base, 7.4 g EDTA, 53.4 g boric acid per litre distilled water

DEB : 10% glycerol, 60 mM Tris-HCl pH 6.8, 5% ß-mercaptoethanol, 0.1% SDS

DRB : 67 mM Tris-malate pH 7.1, 42 mM $MgCl_2$, 0.4 M NH_4Cl, 10 µM ATP, 25 µM kanamycin, 10 mM NaF, 15 µCi γ-^{32}P-ATP

GEB : 10% glycerol, 40 mM EDTA, 150 mM NaCl, 100 mM NH_4Cl, 10 mM Tris-HCl pH 7.5, 3 mg/ml dithiothreitol, 0.2 mg/ml leupeptin

GRB : 67 mM Tris-malate pH 7.1, 42 mM $MgCl_2$, 400 mM NH_4Cl

GUS extraction buffer : 50 mM NaPO4 pH 7.0, 10 mM EDTA, 0.1% Triton-X 100, 0.1% Sarkosyl, 10 mM β-mercaptoethanol

Bradford's solution : 600 mg/ml Serva blue G-250 stain in 2% perchloric acid and filtered

REFERENCES

ALBANI D, ROBERT LS, DONALDSON PA, ALTOSAAR I, ARNISON PG, FABIJANSKI SF (1990) Characterization of a pollen-specific gene family from *Brassica napus* which is activated during early microspore development. Plant Mol. Biol. 15: 605-622

ALBANI D, ALTOSAAR I, ARNISON PG, FABIJANSKI SF (1991) A gene showing similarity to pectin esterase is specifically expressed in developing pollen of *Brassica napus*. Sequences in its 5' flanking region are conserved in other pollen-specific promoters. Plant Mol. Biol. 16: 501-513

BAKER B, SCHELL J, LORZ H, FEDOROFF NV (1986) Transposition of the maize controlling element '*Activator*' in tobacco. Proc. Nat. Acad. Sci. USA 83: 4844-4848

BAKER B, COUPLAND G, FEDOROFF NV, STARLINGER P, SCHELL J (1987) Phenotypic assay for the excision of the maize controlling element *Ac* in tobacco. EMBO J. 6:1547-1554

BASIRAN N, ARMITAGE P, SCOTT RJ, DRAPER J (1987) Genetic transformation of flax (*Linum usitatissimum*) by *Agrobacterium tumefaciens* : Regeneration of transformed shoots via a callus phase. Plant Cell Reports 6: 396-399

BEDINGER PA, EDGERTON MD (1990) Developmental staging of maize microspores reveals a transition in developing microspore proteins. Plant Physiol. 92: 474-479

BEHRENS U, FEDOROFF N, LAIRD A, MULLER-NEUMANN M, STARLINGER P, YODER J (1984) Cloning of the *Zea mays* controlling element *Ac* from the *wx-m7* allele. Mol. Gen. Genet. 194: 346-347

BHANDARI NN (1984) The Microsporangium. In: Embryology of Angiosperms, Ed. Johri BM (Springer Verlag) pp. 53-121

BINO RJ (1985) Histological aspects of microsporogenesis in fertile, cytoplasmic male sterile and restored fertile *Petunia hybrida.* Theor. Appl. Genet. 69: 425-428

BIRNBOIM HC, DOLY J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7: 1513-1523

BONAS U, SOMMER H, HARRISON BJ, SAEDLER H (1984) The transposable element *Tam1* of *Antirrhinum majus* is 17 kb long. Mol. Gen. Genet. 194: 138-143

BRADFORD MM (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilising the principle of protein dye binding. Anal. Biochem. 72: 248-254

BRINK RA, NILAN RA (1952) The relationship between light variegated and medium variegated pericarp in maize. Genetics 37: 519-544

BROWN SM, CROUCH ML (1990) Characterization of a gene family abundantly expressed in *Oenothera organensis* pollen that shows sequence similarity to polygalacturonase. The Plant Cell 2: 263-274

CARNIEL, K (1963) Das antherentapetum. Ein kritischer uberblick. Oesterr. Bot. Z. 110: 145-176

CASSAB GI, VARNER JE (1988) Cell wall proteins. Ann. Rev. Plant Physiol. Plant Mol. Biol. 39: 321-353

CHEN J, GREENBLATT IM, DELLAPORTA SL (1987) Transposition of *Ac* from the *P* locus of maize into unreplicated chromosomal sites. Genetics 117: 109-116

CHOMET PS, WESSLER S, DELLAPORTA SL (1987) Inactivation of the maize transposable element *Activator* (*Ac*) is associated with its' DNA modification. EMBO J. 6: 295-302

CHOUREY PS, DeROBERTIS GA, STILL PE (1988) Altered tissue specificity of the revertant *shrunken* allele upon *Dissociation* (*Ds*) excision is associated with loss of expression and molecular rearrangement at the corresponding non-allelic isozyme locus in maize. Mol. Gen. Genet. 214: 300-306

COUPLAND G, BAKER B, SCHELL J, STARLINGER P (1988) Characterization of the maize transposable element *Ac* by internal deletions. EMBO J. 7: 3653-3659

COUPLAND G, PLUM C, CHATTERJEE S, POST A, STARLINGER P (1989) Sequences near the termini are required for transposition of the maize transposon *Ac* in transgenic tobacco plants. Proc. Natl. Acad. Sci. USA 86: 9385-9388

CULLIS CA (1986) Phenotypic consequences of environmentally induced changes in plant DNA. Trends Genet. 2: 307-309

CULLIS CA, CLEARY W (1986) DNA variation in flax tissue culture. Can. J. Genet. Cytol. 28: 247-251

DELLAPORTA SL, CHOMET PS (1985) The activation of maize controlling elements. In: "Genetic Flux in Plants" Ed. Hohn B, Dennis ES, pp. 169-216 Vienna/New York: Springer Verlag.

DENNIS ES, SACHS MM, GERLACH WL, BEACH L, PEACOCK WJ (1988) The *Ds1* transposable element acts as an intron in the mutant allele *Adh1-Fm335* and is spliced from the message. Nucleic Acids Res. 16: 3815-3828

DEVEREUX J, HAEBERLI P, SMITHIES O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 12: 387-395

DITTA G, STANSFIELD S, CORBIN D, HELINSKI DR (1980) Broad host range DNA cloning system for Gram negative bacteria - Construction of a gene bank of *Rhizobium melilotis*. Proc. Nat. Acad. Sci. USA 77: 7347-7351

DOERSCHUG EB (1973) Studies of *Dotted*, a regulatory element in maize. I. Induction of *Dotted* by chromatid breaks. II. Phase variation of *Dotted*. Theor. Appl. Genet. 43: 182-189

DOONER HK, WECK E, ADAMS S, RALSTON E, FAVREAU M (1985) A molecular genetic analysis of insertions in the *Bronze* locus in maize. Mol. Gen. Genet. 200: 240-246

DOONER HK, BELACHEW A (1989) Transposition pattern of the maize element *Ac* from the *bz-m2(Ac)* allele. Genetics 122: 447-457

DOONER HK, KELLER J, HARPER, RALSTON E (1991) Variable patterns of transposition of the maize element *Activator* in tobacco. The Plant Cell 3: 473-482

DORING H-P, FREELING M, HAKE S, JOHNS MA, KUNZE R (1984a) A *Ds* mutation of the *Adh1* gene in *Zea mays* L. Mol. Gen. Genet. 193: 199-204

DORING H-P, TILLMAN E, STARLINGER P (1984b) DNA sequence of the maize transposable element *Dissociation*. Nature 307: 127-130

DORING H-P, STARLINGER P (1986) Molecular genetics of transposable elements in plants. Ann. Rev. Genet. 20: 175-200

DORING H-P, NELSEN-SALZ B, GARBER R, TILLMAN E (1989) Double *Ds* elements are involved in specific chromosome breakage. Mol. Gen. Genet. 219: 229-305

DORING H-P, PAHL I, DURANY M (1990) Chromosomal rearrangements caused by the abberant transposition of double *Ds* elements are formed by *Ds* and adjacent non-*Ds* sequences. Mol. Gen. Genet. 224: 40-48

DOWE MF, ROMAN GW, KLEIN AS (1990) Excision and transposition of two *Ds* transposons from the *bronze mutable* 4 derivative 6856 allele of *Zea mays* L. Mol. Gen. Genet. 221: 475-485

DRAPER J, SCOTT RJ, ARMITAGE P, WALDEN R, *Ed.* (1988) Plant Genetic Transformation and Gene Expression. A laboratory manual. Blackwell Scientific Publications, Oxford.

DURRANT A (1971) Induction and growth of flax genotrophs. Heredity 27: 277-298

ECHLIN P, GODWIN H (1968) The ultrastructure and ontogeny of pollen in *Helleborus foetidus* L.: Pollen grain development through the callose special wall stage. J. Cell Sci. 3: 175-186

ELLIS JG, LAWRENCE GJ, PEACOCK WJ, PRYOR AJ (1988) Approaches to cloning plant genes conferring resistance to fungal pathogens. Ann. Rev. Phytopathol. 26: 245-263

EVANS DE, ROTHNIE NE, PALMER MV, BURKE DG, SANG JP, KNOX RB, WILLIAMS EG, HILLIARD EP, SALISBURY PA (1987) Comparative analysis of fatty acids in pollen and seed of rapeseed. Phytochemistry 26: 1895-1897

EVANS DA (1989) Somaclonal variation - genetic basis and breeding applications. Trends Genet. 5: 46-50

EVANS DE, SANG JP, COMINOS X, ROTHNIE NE, KNOX RB (1990a) A study of phospholipids and galactolipids in pollen of two lines of *Brassica napus* L. (rapeseed) with different ratios of linoleic to linolenic acid. Plant Physiol. 92: 418-424

EVANS DE, SINGH MB, KNOX RB (1990b) Pollen development: applications in biotechnology. In "Microspores: evolution and ontogeny" (Academic Press Limited) pp.309-338

EVRARD J-L, JAKO C, SAINT-GUILY A, WEIL J-H, KUNTZ M (1991) Anther-specific, developmentally regulated expression of genes encoding a new class of proline-rich proteins in sunflower. Plant Mol Biol 16: 271-281

FEDOROFF N, WESSLER S, SHURE M (1983) Isolation of the transposable maize controlling elements *Ac* and *Ds*. Cell 35: 235-242

FEDOROFF N, FURTEK D, NELSON O (1984) Cloning of the *bronze* locus in maize by a

simple and generalizable procedure using the transposable controlling element *Activator* (*Ac*). Proc. Natl. Acad. Sci. USA 81: 3825-3829

FEINBERG AP, VOGELSTEIN B (1984) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 137: 266-267

FLAVELL AJ, SMITH DB, KUMAR A (1992) Extreme heterogeneity of *Ty1-copia* group retrotransposons in plants. Mol. Gen. Genet. 231: 233-242

FLOR HH (1956) The complementary genic systems in flax and flax rust. Adv. Genet. 8: 29-54

FREELING M (1984) Plant transposable elements and insertion sequences. Ann. Rev. Plant Physiol. 35: 227-298

FREY M, TAVANTZIS SM, SAEDLER H (1989) The maize En1/Spm element transposes in potato. Mol. Gen. Genet. 217: 172-177

GALLIE DR, SLEAT DE, WATTS JW, TURNER PC, WILSON TMA (1987) The 5'-leader sequence of tobacco mosaic virus RNA enhances the expression of foreign gene transcripts *in vitro* and *in vivo*. Nucleic Acids Res. 15: 3257-3273

GLOVER CVC (1989) Sequence-specific protein-DNA recognition by transcriptional regulatory proteins. Plant Mol. Biol. Reporter 7: 183-208

GOLDBERG RB, HOSCHECK G, VODKIN LO (1983) An insertion sequence blocks the expression of a soybean lectin gene. Cell 33: 465-475

GREENBLATT IM, BRINK RA (1962) Twin mutations in medium variegated pericarp maize. genetics 47: 489-501

GREENBLATT IM, BRINK RA (1963) Transpositions of Modulator in maize into divided and undivided chromosome segments. Nature 197: 412-413

GREENBLATT IM (1984) A chromosomal replication pattern deduced from pericarp phenotypes resulting from movements of the transposable element, *Modulator,* in maize. Genetics 108: 471-485

GUERRERO FD, CROSSLAND L, SMUTZER GS, HAMILTON DA, MASCARENHAS JP (1990) Promoter sequences from a maize pollen-specific gene direct tissue-specific transcription in tobacco. Mol. Gen. Genet. 224: 161-168

HAMILTON DA, BASHE DM, STINSON JR, MASCARENHAS JP (1989) Characterization of a pollen-specific genomic clone from maize. Sex. Plant Reprod. 2: 208-212

HANSON DD, HAMILTON DA, TRAVIS JI, BASHE DM, MASCARENHAS JP (1989) Characterization of a pollen-specific cDNA clone from *Zea mays* and its' expression. The Plant Cell 1: 173-179

HARING MA, ROMMENS CMT, NIJKAMP HJ, HILLE J (1991) The use of transgenic plants to understand transposition mechanisms and to develop transposon tagging strategies. Plant Mol. Biol. 16: 449-461

HATZOPOULOS P, FRANZ G, CHOY L, SUNG RZ (1990) Interaction of nuclear factors with upstream sequences of a lipid body membrane protein gene from carrot. Plant Cell 2: 457-467

HAUSER C, FUSSWINKEL H, LI J, OELLIG C, KUNZE R, MULLER-NEUMANN M, HEINLEIN M, STARLINGER P, DOEFLER W (1988) Overproduction of the protein encoded by the maize transposable element *Ac* in insect cells by a baculovirus vector. Mol. Gen. Genet. 214: 373-378

HEHL R, BAKER B (1989) Induced transposition of *Ds* by a stable *Ac* in crosses of transgenic tobacco plants. Mol. Gen. Genet. 217: 53-59

HEHL R, NACKEN WKF, KRAUSE A, SAEDLER H, SOMMER H (1991) Structural analysis of *Tam3*, a transposable element from *Antirrhinum majus*, reveals homologies to the *Ac* element from maize. Plant Mol. Biol. 16: 369-371

HESLOP-HARRISON J (1963) Tapetal origin of pollen-coat substances in *Lilium*. New Phytol. 67: 779-786

HESLOP-HARRISON J (1968) Wall development within the microspore tetrad of *Lilium longiflorum*. Can. J. Bot. 46: 1185-1191

HUSSEY PJ, LLOYD CW, GULL K (1988) Differential and developmental expression of ß-tubulins in a higher plant. J. Biol. Chem. 263: 5474-5479

JEFFERSON RA, KAVANAGH TA, BEVAN MW (1987) GUS fusions: ß-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J. 6: 3901-3907

JONES JDG, SVAB Z, HARPER EC, HURWITZ CD, MALIGA P (1987) A dominant nuclear streptomycin resistance marker for plant cell transformation. Mol. Gen. Genet. 210: 86-91

JONES JDG, CARLAND FM, MALIGA P, DOONER HK (1989) Visual detection of transposition of the maize element *Activator* (*Ac*) in tobacco seedlings. Science 244: 204-207

JONES JDG, CARLAND FM, LIM E, RALSTON E, DOONER HK (1990) Preferential transposition of the maize element *Activator* to linked chromosomal locations in tobacco. The Plant Cell 2: 701-707

JORDAN E, SAEDLER H, STARLINGER P (1969) Strong-polar mutations in the transferase gene of the galactose operon in *E. coli*. Mol. Gen. Genet. 100: 296-306

JOSHI CP (1987) An inspection of the domain between putative TATA box and translation start site in 79 plant genes. Nucleic Acids Res. 15: 6643-6653

KAMALAY JC, GOLDBERG RB (1980) Regulation of structural gene expression in tobacco. Cell 19: 934-946

KOLTUNOW AM, TURETTNER J, COX KH, WALLROTH M, GOLDBERG RB (1990) Differential temporal and spatial gene expression patterns occur during anther development. The Plant Cell 2: 1201-1224

KNAPP S, COUPLAND G, UHRIG H, STARLINGER P, SALAMINI F (1988) Transposition of the maize transposable element *Ac* in *Solanum tuberosum*. Mol. Gen. Genet. 213: 285-290

KNOX RB, HESLOP-HARRISON J (1970) Pollen wall proteins: localization and enzymic activity. J. Cell Sci. 6: 1-27

KUNZE R, STOCHAJ U, LAUFS J, STARLINGER P (1987) Transcription of the transposable element *Activator* (*Ac*) of *Zea mays* L. EMBO J. 6: 1555-1563

KUNZE R, STARLINGER P, SCHWARTZ D (1988) DNA methylation of the maize transposable element *Ac* interferes with its transcription. Mol. Gen. Genet. 214: 325-327

KUNZE R, STARLINGER P (1989) The putative transposase of transposable element *Ac* from *Zea mays* L. interacts with subterminal sequences of *Ac*. EMBO J. 8: 3177-3185

KYTE J, DOOLITTLE RF (1982) A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157: 105-132

LASSNER MW, PALYS JM, YODER JI (1989) Genetic transactivation of dissociation elements in transgenic tomato plants. Mol. Gen. Genet. 218: 25-32

LI M-G, STARLINGER P (1990) Mutational analysis of the N-terminus of the protein of maize transposable element *Ac*. Proc. Natl. Acad. Sci. USA 87: 6044-6048

LUDWIG SR, OPPENHEIMER DG, SILFLOW CD, SNUSTAD DP (1988) The α_1-tubulin gene of *Arabidopsis thaliana* : primary sequence and preferential expression in flowers. Plant Mol. Biol. 10: 311-321

MARAKOFF CA, PALMER JD (1988) Mitochondrial DNA rearrangements and transcriptional alterations in the male sterile cytoplasm of Ogura radish. Molec. Cellular Biol. 8: 1471-1480

MARTIN CR, PRESCOTT A, LISTER C, MacKAY S (1989) Activity of the transposon *Tam3* in *Antirrhinum* and tobacco: possible role of DNA methylation. EMBO J. 8: 997-1004

MASCARENHAS JP (1989) The male gametophyte of flowering plants. The Plant Cell 1: 657-664

MASCARENHAS JP (1990) Gene activity during pollen development. Ann. Rev. Plant Physiol. Plant Mol. Biol. 41: 317-338

MASSON P, FEDOROFF NV (1989) Mobility of the maize *suppresor-mutator* element in transgenic tobacco cells. Proc. Natl. Acad. Sci. USA 86: 2219-2223

McCLINTOCK B (1946) Maize genetics. Carnegie Inst. Washington Year Book 45: 176-186

McCLINTOCK B (1948) Mutable loci in maize. Carnegie Inst. Washington Year Book 47: 155-169

McCLINTOCK B (1950) The origin and behaviour of mutable loci in maize. Proc. Natl. Acad. Sci. USA 36: 344-357

McCLINTOCK B (1956) Mutation in maize. Carnegie Inst. Washington Year Book 55: 323-332

McCLINTOCK B (1965) The control of gene action in maize. Brookhaven Symp. Biol. 18: 162-184

McCORMICK SM, SMITH C, GASSER CS, SACHS K, HINCHEE M, HORSCH R, FRALEY RT (1987) Identification of genes specifically expressed in reproductive organs of tomato.

In Tomato Biotechnology, Ed. Nevins DJ and Jones RA (New York: Liss) pp. 255-265

McDONNELL RE, CLARK RD, SMiTH WA, HINCHEE MA (1987) A simplified method for the detection of neomycin phosphotransferase II activity in transformed plant tissue. Plant Mol. Biol. Reporter 5: 380-386

MEPHAM RH, LANE GR (1969) Formation and development of the tapetal plasmodium in *Tradescantia bracteata.* Protoplasma 68: 175-192

MERCKELBACH A, DORING H-P, STARLINGER P (1986) The aberrant *Ds* element in the *Adh1-2F11::Ds* allele. Maydica 31: 109-122

MEYEROWITZ EM, (1989) *Arabidopsis*, a useful weed. Cell 56: 263-269

MULLER-NEUMANN M, YODER JI, STARLINGER P (1984) The DNA sequence of the transposable element *Ac* of *Zea mays* L. Mol. Gen. Genet. 198: 19-24

MURAI N, LI ZJ, KAWAGOE Y, HAYASHIMOTO A (1991) Transposition of the maize *Activator* element in transgenic rice plants. Nucleic Acids Res. 19: 617-622

MURASHIGE T, SKOOG F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. Plant Pysiol. 15: 473-497

MURPHY DJ, KEEN JN, O'SULLIVAN JN, AU DMY, EDWARDS E-W, JACKSON PJ, CUMMINS I, GIBBONS T, SHAW CH, RYAN AJ (1991) A class of amphipathic proteins associated with lipid storage bodies in plants. Possible similarities with animal serum apolipoproteins. Biochem. Biophys. Acta 1088: 86-94

NEVINS JR (1983) The pathway of eukaryotic mRNA formation. Ann. Rev. Biochem. 52: 441- 446

OGURA H (1968) Studies on the new male sterility in Japanese radish with special reference to utilization of sterility towards the practical raising of hydrid seed. Mem. Fac. Agric. Kagoshima Univ. 6: 39-78

OSBORNE BI, CORR CA, PRINCE JP, HEHL R, TANKSLY SD, McCORMICK S, BAKER B (1991) *Ac* transposition from a T-DNA can generate linked and unlinked clusters of insertions in the tomato genome. Genetics 129: 833-844

PAZ-ARES J, WIENAND U, PETERSON PA, SAEDLER H (1986) Molecular cloning of the *c* locus of *Zea mays*, a locus regulating the anthocyanin pathway. EMBO J. 5: 829-833

PEREIRA A, SCHWARZ-SOMMER Zs, GIERL A, BERTRAM I, PETERSON PA, SAEDLER H (1985) Genetic and molecular analysis of the *Enhancer* (*En*) transposable element system of *Zea mays.* EMBO J. 4: 17-23

PESCHKE VM, PHILLIPS RL, GENGENBACH BG (1987) Discovery of transposable element activity among progeny of tissue culture-derived maize plants. Science 238: 804-807

PLEGT L, BINO RJ (1989) ß-glucuronidase activity during development of the male gametophyte from transgenic and non-transgenic plants. Mol. Gen. Genet. 216: 321-327

POHLMAN RF, FEDOROFF NV, MESSING J (1984) The nucleotide sequence of the maize controlling element *Activator.* Cell 37: 635-643

QU R, HUANG AHC (1990) Oleosin KD 18 on the surface of oil bodies in maize: genomic

and cDNA sequences, and the deduced protein structure. J. Biol. Chem. 265: 2238-2243
RAINES CA, LLOYD JC, CHAO S, JOHN UP, MURPHY GJP (1991) A novel proline-rich protein from wheat. Plant Mol Biol 16: 663-670
RAJHATHY T (1976) Haploid flax revisited. Z. Pflanzenzuchtg. 76: 1-10
SAEDLER H, NEVERS P (1985) Transposition in plants: a molecular model. EMBO J. 4: 585-590
SAMBROOK J, FRITSCH EF, MANIATIS T (1989) Molecular Cloning: A laboratory manual. Cold Spring Harbour Press, Cold Spring Harbour, NY.
SANGER F, NICKLEN S, COULSON AR (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74: 5463-5467
SCHIEFELBEIN JW, FURTEK DB, RABOY V, BANKS JA, FEDOROFF N, NELSON OE (1985) Exploiting transposable elements to study the expression of a maize gene. In: Plant Genetics, Ed. Freeling M. (New York: Liss) pp. 445-459
SAUER N, CORBIN DR, KELLER B, LAMB CJ (1990) Cloning and characterization of a wound-specific hydroxyproline-rich glycoprotein in *Phaseolus vulgaris*. Plant, Cell and Environment 13: 257-266
SCHMIDT R, WILLMITZER L (1989) The maize autonomous element *Activator* (*Ac*) shows a minimal germinal excision frequency of 0.2% - 0.5% in transgenic *Arabidopsis thaliana* plants. Mol. Gen. Genet. 220: 17-24
SCHMIDT RJ, BURR FA, BURR B (1987) Transposon tagging and molecular analysis of the maize regulatory locus, *opaque-2*. Science 238: 960-963
SCHRAUWEN JAM, de GROOT PFM, van HERPEN MMA, van der LEE T, REYNEN WH, WETERINGS KAP, WULLEMS GJ (1990) Stage-related expression of mRNAs during pollen development in lily and tobacco. Planta 182: 298-304
SCHWARTZ D (1989) Pattern of *Ac* transposition in maize. Genetics 121: 125-128
SCOTT R, DAGLESS E, HODGE R, PAUL W, SOUFLERI I, DRAPER J (1991a) The identification of anther-specific genes in *Brassica napus*. Plant Mol. Biol. 17: 195-207
SCOTT R, HODGE R, PAUL W, DRAPER J (1991b) The molecular biology of anther differentiation. Plant Sci. 80: 167-191
SHIRSAT AH (1988) A transposon-like structure in the 5' flanking sequence of a legumin gene from *Pisum sativum*. Mol. Gen. Genet. 212: 129-133
SHWARTZ D, DENNIS ES (1986) Transposase activity of the *Ac* controlling element in maize is regulated by its degree of methylation. Mol. Gen. Genet. 205: 476-482
SMITH AG, GASSER CS, BUDELIER KA, FRALEY RT (1990) Identification and characterization of stamen- and tapetum-specific genes from tomato. Mol. Gen. Genet. 222: 9-16
SMYTH DR (1991) Dispersed repeats in plant genomes. Chromosoma 100: 355-359
SOMMER H, CARPENTER R, HARRISON BJ, SAEDLER H (1985) The transposable element *Tam3* of *Antirrhinum majus* generates a novel type of sequence altertion upon excision. Mol. Gen. Genet. 199: 225-231

SPRAGUE GF, McKINNEY HH, GREELEY L (1963) Virus as a mutagenic agent in maize. Science 14: 1052-1053

STACE C (1991) New flora of the British Isles (Cambridge University Press) pp. 335-336

STINSON JR, EISENBERG AJ, WILLING RP, PE ME, HANSON DD, MASCARENHAS J (1987) Genes expressed in the male gametophyte of flowering plants and their isolation. Plant Physiol. 83: 442-447

STINSON J, MASCARENHAS JP (1985) Onset of alcohol dehydrogenase synthesis during microsporogenesis in maize. Plant Physiol. 77: 222-224

SUTTON WD, GERLACH WL, SCHWARTZ D, PEACOCK WJ (1984) Molecular analysis of *Ds* controlling element mutations at the *Adh1* locus of maize. Science 223: 1265-1268

THEERAKULPISUT P, XU H, SINGH MB, PETTITT JM, KNOX RB (1991) Isolation and developmental expression of Bcp1, an anther-specific cDNA clone in *Brassica campestris*. The Plant Cell 3: 1073-1084

TWELL D, WING R, YAMAGUCHI J, McCORMICK S (1989) Isolation and expression of an anther-specific gene from tomato. Mol. Gen. Genet. 217: 240-245

TWELL D, YAMAGUCHI J, McCORMICK S (1990) Pollen-specific gene expression in transgenic plants: coordinate regulation of two different tomato gene promoters during microsporogenesis. Development 109: 705-713

TWELL D, YAMAGUCHI J, WING RA, USHIBA J, McCORMICK S (1991) Promoter analysis of genes that are coordinately expressed during pollen development reveals pollen-specific enhancer sequences and shared regulatory elements. Genes Devel. 5: 496-507

TZEN JTC, LAI Y-K, CHAN K-L, HUANG AHC (1990) Oleosin isoforms of high and low molecular weights are present in oil bodies of diverse seed species. Plant Physiol. 94: 1282-1289

UPADHYAYA KC, SOMMER H, KREBBERS U, SAEDLER H (1985) The paramutagenic line *niv-44* has a 5kb insert, *Tam2*, in the chalcone synthase gene of *Antirrhinum majus*. Mol. Gen. Genet. 199: 210-207

URSIN VM, YAMAGUCHI J, McCORMICK S (1989) Gametophytic and sporophytic expression of anther-specific genes in developing tomato anthers. The Plant Cell 1: 727-736

VAN SLUYS MA, TEMPE J, FEDOROFF NV (1987) Studies on the introduction and mobility of the maize *Activator* element in *Arabidopsis thaliana* and *Daucus carota*. EMBO J. 6: 3881-3889

VANCE VB, HUANG AHC (1987) The major protein from lipid bodies of maize. Characterization and structure based on cDNA cloning. J. Biol. Chem. 262: 11275-11279

VELTEN J, VELTEN L, HAIN R, SCHELL J (1984) Isolation of a dual plant promoter fragment from the Ti-plasmid of *Agrobacrerium tumefaciens*. EMBO J. 3: 2723-2730

VERGNE P, DUMAS C (1988) Isolation of viable wheat male gametophytes of different stages of development and variations in their protein patterns. Plant Physiol. 88: 969-972

VITHANAGE HIMV, KNOX RB (1976) Pollen wall proteins: Quantitative cytochemistry of the origins of intine and exine enzymes in *Brassica oleracea*. J. Cell Sci. 21: 423-435

VON HEINJE G (1983) Patterns of amino acids near signal-sequence cleavage sites. Eur. J. Biochem. 133: 17-21

VOYTAS DF, KONIECZNY A, CUMMINGS MP, AUSUBEL FM (1990) The structure, distribution and evolution of the *Ta1* retrotransposable element family of *Arabidopsis thaliana*. Genetics 126: 713-721

WARMKE HE, LEE S-LJ (1977) Mitochondrial degeneration in Texas cytoplasmic male-sterile corn anthers. J. Hered. 68: 213-222

WESSLER SR, BARAN G, VARAGONA M, DELLAPORTA SL (1986) Excision of *Ds* produces *waxy* proteins with a range of enzymic activities. EMBO J. 5: 2427-2432

WIENAND U, WEYDEMANN U, NIESBACH-KLOSGEN U, PETERSON PA, SAEDLER H (1986) Molecular cloning of the *c2* locus of *Zea mays*, the gene coding for chalcone synthase. Mol. Gen. Genet. 203: 202-207

WILLING RP, MASCARENHAS JP (1984) Analysis of the complexity and diversity of mRNAs from pollen and shoots of *Tradescantia*. Plant Physiol. 75:865-868

WILLING RP, BASHE D, MASCARENHAS JP (1988) An analysis of the quantity and diversity of messenger RNAs from pollen and shoots of *Zea mays*. Theor. Appl. Genet. 75: 751-753

WING RA, YAMAGUCHI J, LARABELL SK, URSIN VM, McCORMICK S (1989) Molecular and genetic characterization of two pollen-expressed genes that have sequence similarity to pectate lyases of the plant pathogen *Erwinia*. Plant Mol. Biol. 14 : 17-28

YODER J, PALYS J, ALPERT K, LASSNER M (1988) *Ac* transposition in transgenic tomato plants. Mol. Gen. Genet. 213: 291-296

ZAMBRYSKI P, JOOS H, GENETELLO C, LEEMANS J, VAN MONTAGU M, SCHELL J (1983) Ti plasmid vector for the introduction of DNA into plant cells without alteration of their normal regeneration capacity. EMBO J. 2: 2143-2150

ZHOU JH, ATHERLY AG (1990) In situ detection of the maize controlling element (*Ac*) in transgenic soybean tissues. Plant Cell Reports 8: 542-545

ZHOU JH, MYERS A, ATHERLY AG (1991) Functional analysis of the 3'-terminal sequence of the maize controlling element (*Ac*) by internal replacement and deletion mutagenesis. Genetica 84: 13-21