Isolation of expressed gene sequences from the stigma of brassica oleracea

Booker, Jonathan Peter

Isolation Of Expressed Gene Sequences From The Stigma Of *Brassica Oleracea.*

Jonathan Peter Booker

Thesis for the degree of PhD, 1996

Abstract

Molecular characterisation of the stigma and style in the Cruciferae has concentrated exclusively on the components of the SI system, and other members of their sequence families, while neglecting other sequences involved in tissue development and compatible pollination. To address this imbalance work has been carried out to isolate expressed gene sequences from the stigma of *Brassica oleracea.* cDNA clones were isolated from a cDNA library and characterised together with two clones isolated previously. Four out of eleven clones represented previously isolated S-gene and S-related gene sequences which have been extensively studied by other workers. One other sequence that could be positively identified encodes a putative glycine-rich cell wall protein which is expressed strongly in the stigma and more weakly in the style and petal, suggesting that it plays a specific role in the former.

Of the other isolated clones, two cross-hybridise with closely related transcripts that are stress-induced in vegetative tissue. This allows parallels to be drawn with the Solanaceae, in which proteins, normally induced by stress, are also expressed in healthy carpels. However, the *Brassica* species differ from these in that they are expressed in all floral whorls. Neither stress-induced clone can be positively identified although their encoded proteins share homology with the major latex protein of poppy and two bacterial protein sequences.

Of the remaining four sequences, two were enriched in stigmas, but contained insufficient sequence data to allow identification. The third sequence was expressed throughout the mature plant, though expression was higher in floral tissue, while expression of the fourth could not be determined.

From these sequences no information concerning the molecular basis stigma development and function can be deduced. Therefore other approaches to studying this tissue are discussed.

Isolation Of Expressed Gene Sequences From The Stigma Of *Brassica Oleracea*.

Jonathan Peter Booker

Thesis for the degree of PhD
University of Durham.
Department of Biological Sciences.
1996

Declaration

I confirm that no part of the material offered has previously been submitted by me for a degree in this or in any other university. If material has been generated through joint work, my independent contribution has been clearly indicated. In all other cases material from the work of others has been acknowledged and quotations and paraphrases suitably indicated.

J.P.Booker.

Acknowledgements

I have the following people to thank with respect to my work for the PhD and the writing of this thesis: Firstly I would like to thank Ron Croy for offering me the chance to work on this project, for supervising me throughout and for hints and criticism on this thesis. Secondly Charlie Scutt for isolating 15H11 and 13G6(i) so that the project existed in the first place.

For assistance with nucleic acid techniques I'd like to thank Liz Croy, Tony Fordham-Skelton, Russell Swinhoe and Nigel Robinson. Also Lesley Sinclair, Colin Illet and Duncan Robertson for their invaluable help with the protein work.

I also thank Julia Bartley for her continuously fine DNA sequencing and John Gilroy for primer synthesis for the PCR. For their provision of various plants, nucleic acids and libraries, I'd like to thank Rudolpho Bernadi, Kevin Powell, Tim Daniell and Martin Trick.

Dedication

I dedicate this thesis to the memory of Bailey, Sapphyre and Tinker who helped keep me going through the last four and a half years and who I'll miss for many more years to come

CONTENTS.

FIGURES AND TABLES

ABBREVIATIONS

BCIP	5-bromo-4-chloro-3-indoyl phosphate
bp.	base pair(s)
BSA	bovine serum albumin
cDNA	complementary deoxyribonucleic acid
CaMV	cauliflower mosaic virus
ConA	concanavalin A
CTAB	cetyltrmethylammoniumbromide
DEPC	diethylpyrocarbonate
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide(s)
DTE	dithreothreitol
ECM	extra cellular matrix
EDTA	ethylenediaminetetraacetic acid
EM	electron microscope
EST	expressed sequence tag
FUE	far upsteam element
g	gramme(s)
GRP	glycine rich protein
IAA	isoamyl-alcohol
IEF	iso-electric focussing
M	molar
min	minute
MLP	major latex protein
MOPS	3-[Nmorpholino]propanesulphonic acid
mRNA	messenger ribonucleic acid
NBT	nitroblue tetrazolium
nt	nucleotide(s)
ORF	open reading frame
PAS	periodic-acid-schiff
PCR	polymerase chain reaction
pfu	plaque forming unit(s
RF	reading frame
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
RNase	ribonuclease
RT-PCR	reverse transcriptase polymerase chain reaction
s	second
SAR	systemic acquired resistence
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SI	self incompatibility
SLG	S-linked glycoprotein
SLR	S-locus related sequence
SLSG	S-linked style glycoprotein
SRK	S-related kinase
SSC	(1x) 0.15M sodium chloride, 0.015M sodium citrate, pH7.0
TAE	0.04M Tris-acetate, 0.001m EDTA, pH8.0
TE	10mM TrisHCl (pH8.0), 1mM EDTA
TFA	trifluoroacetic acid
TMV	tobacco mosaic virus
Tris	tis[hydroxymethyl]aminomethane
U	enzyme units
UHQ	ultra-high quality
UTR	untrantslated region
UV	ultraviolet

SEQUENCE CONVENTIONS

Standard IUPAC single letter codes for nucleic acids and amino acids are used. They are

Nucleic acids

A	adenosine
C	cytidine
G	guanosine
T	thymidine
U	uridine
N	unknown

Amino acids

A	alanine
C	cysteine
D	aspartate
E	glutamate
F	phenylalanine
G	glycine
H	histidine
I	isoleucine
K	lysine
L	leucine
M	methionine
N	asparagine
P	proline
Q	glutamine
R	arginine
S	serine
T	threonine
V	valine
W	tryptophan
Y	tyrosine
x	unknown

1.Introduction

1.1 The stigma and style as organs for study.

Sexual reproduction is a key element in promoting genetic variability and is therefore a highly desirable trait. Such a factor poses problems for plants whose stationary lifestyle inhibits interaction between the gametes of individuals. In higher plants the principal approach to this problem is to have a small mobile male gamete (pollen) which is released into the environment in order to come into contact with and fertilise a stationary female gamete.

In the angiosperms sexual reproduction is mediated by flowers which are the specialised reproductive structures of these organisms. All flowers have the basic structure shown in fig. 1a, consisting of four concentrically arranged whorls of the four floral organs. The outer two whorls consist of the sepals and petals which are vegetative organs in that they do not contain, or interact with, the male or female gametes. Inside these are the male organs, the stamens, consisting of an anther, which is the organ in which pollen are generated, supported by the filament.

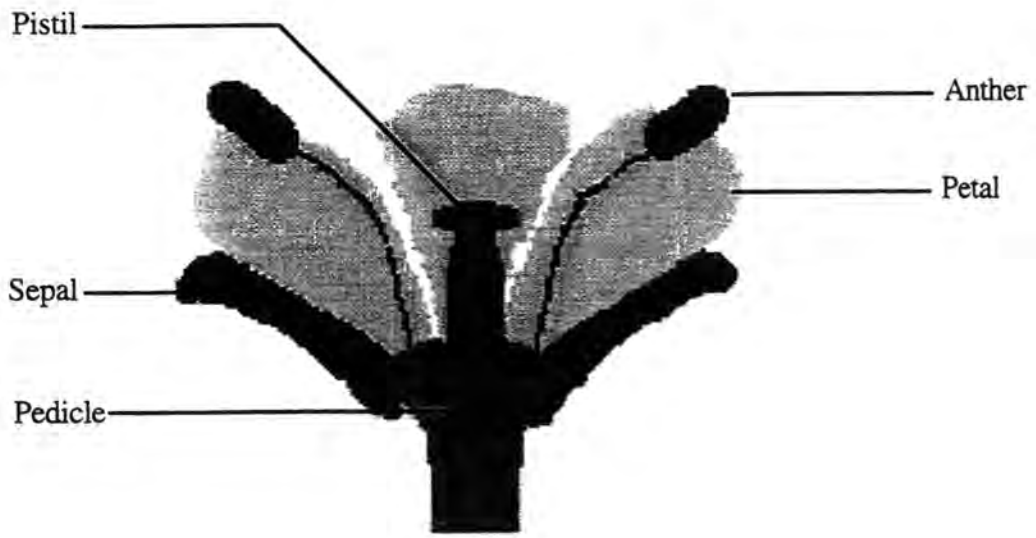
Inside the stamen whorl lies the gynoecium which is the female reproductive part of the flower. Each gynoecium consists of one or more carpels which are distinct female reproductive organs consisting of an ovary, containing the female gametes, together with the stigma and style. The latter two organs are occasionally referred to as the pistil (e.g. Wemmer et al., 1994), but as this term is also a synonym for carpel it will not be used here.

The flower provides an ideal system for the study of developmental processes in plants. The accessibility of the flower, coupled with the highly differentiated nature and precise patterning of its organs has lead to it becoming a model system for the study of plant development, via the isolation and characterisation of mutants (Coen and Carpenter, 1993; Okamuro et al., 1993).

The process of reproduction, pollination, is also a suitable model system for studying plant development. This process, whereby the male gamete (pollen grain) fertilises the female gamete (ovule), can be summarised as interaction with the stigma, germination, growth down the style through a specialised tissue known as the transmitting tract and then fusion with an ovule to achieve fertilisation (fig. 1b). These events involve cell recognition, controlled uptake into the cell, differentiation, directional growth and cell fusion. The accessibility of the floral organs means that these processes can be observed and studied more easily than in the vegetative organs where they are often inaccessible. Pollination is also amenable to study as the precise timing of this process can controlled, via emasculation, by the experimenter This enables precise temporal mapping of events during pollination and so isolation of the organs at various stages of the process.

The stigma and style contain tissues in which much of the cellular interactions involved in pollination occur. They are also the sites of pollen recognition. Therefore much of the work on the pollination process has concentrated on these two organs.

(a)



(b)

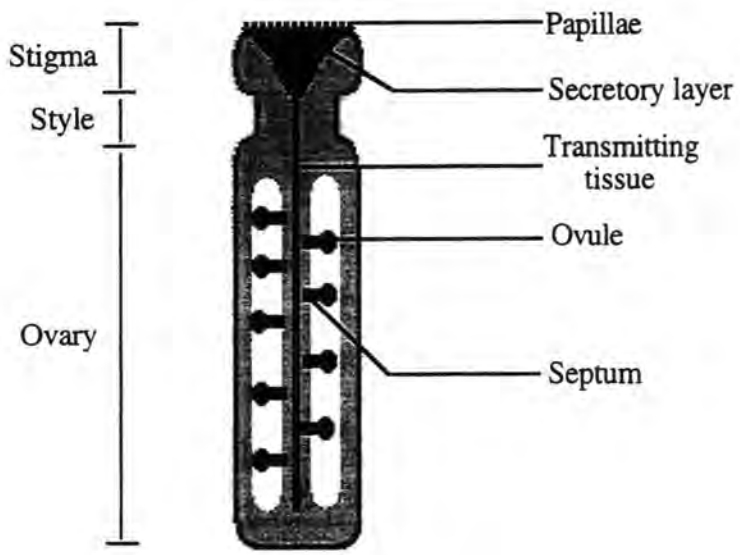


Fig.1. Morphology of an idealised flower. (a) morphology of a simple single-carpelled flower such as found in the Solanaceae, Cruciferae and Lillaceae. (b) simplified diagram illustrating the major components of a simple solid styled carpel. Open stylar carpels have a hollow core in place of the transmitting tract.

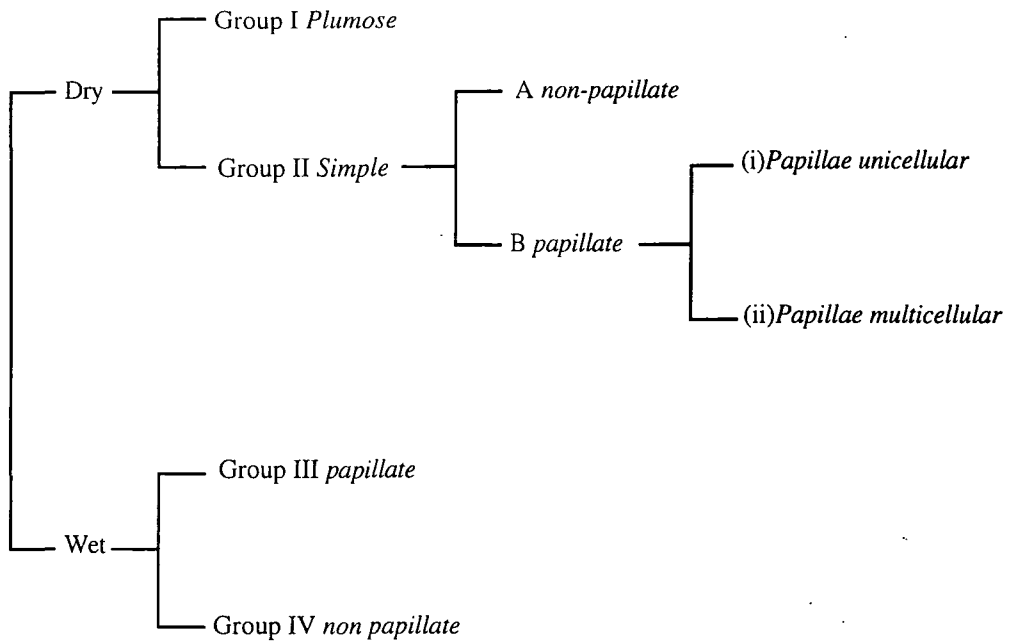


Fig.2. Characterisation of angiosperm stigmas according to Heslop-Harrison and Shivanna (1977).

1.2 GENERAL BIOLOGY OF THE STIGMA AND STYLE.

1.2.1 Heterogeneity between the stigmas and styles of different species.

The angiosperm female tissues differ widely in both their physical and biological attributes. In all plant species, however, the stigma and style contain four major tissues types: the parenchyma, the vascular bundles, the epidermis and the transmitting tissue (Bell and Hicks, 1976) with the first two often being classified together as cortex. However, despite this similarity in structure differences in morphology, anatomy, physiology and biochemistry occur between different plant species. The most extensive characterisation of these differences has been for the first two aspects where differences can be directly observed.

Most of the differences seen between the organs of the carpel are seen in the stigma. The principal differences between stigmas are shape, wet or dry surface and presence/absence of papillae. The first of these is dictated by whether the stigma shape is simple or forms a multiple branching structures (plumose stigma), the latter being predominantly found in the grasses (Heslop-Harrison, 1981).

Wet/dry classification is determined by the nature of the stigma surface (Heslop-Harrison and Shivanna, 1977). Wet stigmas are characterised by an exudate covering the stigmatic surface, although this exudate need not contain a high percentage of water (Konar and Linskens, 1966a). The dry stigma is covered by an adhesive surface known as the pellicle which contains protein carbohydrate and lipid (Mattson 1974).

The final means of classification is based on the nature of the papillae. These take the form of raised hair-like structures on the stigma surface consisting of one or several cells.

These differences have been used by Heslop-Harrison and Shivanna (1977) to produce a classification system for stigmas which is illustrated in fig.2. This classification can be applied to both monocots and dicots even though neither group of plants contains examples of all the categories.

Sub-classification of the reproductive organs can also be performed using the nature of the style. The style can be open, i.e. hollow with the pollen tube growing over the transmitting tissue or closed with the pollen growing intercellularly. A third intermediate class having features of both open and closed styles is limited to the *Cactaceae* which have not been studied at the molecular level (Knox, 1984). The open style is more common in monocots (Knox, 1984) while closed style plants predominate in the dicots (Knox, 1984). The open style consists of a canal lined with extracellular matrix which may or may not be covered by a cuticle (Knox, 1984). The ECM is present between cells in the closed style and can be subdivided into a stigmatic zone and a transmitting tract on the basis of morphological studies (Herrero and Dickinson, 1979).

Biochemical and physiological differences have also been observed between stigmas and styles of different species. These have been catalogued by Knox (1984) and include the presence of proteins, lipids and carbohydrates on the stigma surface and in the extracellular exudate as indicated by the use of stains. The presence and absence of classes of enzymes such as esterases and phosphatases also differ widely between plants.

1.2.2 The stigma and style are active in promoting pollen tube growth.

The principal role of the stigma and style are to capture pollen, support its germination and provide a conduit for pollen tube growth to the ovary. The behaviour of pollen during these events has been most extensively studied using *in vitro* pollen growth assays (reviewed in Steer and Steer, 1989 and Mascarenhas, 1993). Such a system has the principal advantage over working *in vivo* in that developing pollen tubes can be easily observed and/or harvested, an option not available if they are growing through the carpel. Experiments with various pollens with this system have revealed that pollen tube elongation occurs at the tip (Rosen, 1961) with calcium playing a key role in growth (Steer and Steer, 1989).

However, the *in vitro* approach has disadvantages, principal of which is that tube growth is both slower and less extensive than that seen *in vivo*. Pollen in these assays also lacks directionality in its growth. Therefore the stigma and style must have properties which redress these quantitative and qualitative difference between pollen tube growth between the *in vitro* and *in vivo* systems.

Initial experiments confirming this came from the addition of carpel extracts to *in vitro* pollen growth assays which could attract pollen tubes (Rosen, 1961). However, these results could simply be due to the presence of specific nutrients in the extracts as the active factors in the extracts were not fully characterised. More conclusive experimental evidence for the role of the tissues of the stigma and style came from the work of Sanders and Lord (1989) who applied latex beads to the transmitting tissues of *Raphanus raphinastrum, Vicia faba* and *Hemerocallis flava*, then observed their behaviour. These inert beads moved down the transmitting tissue at the same rate as pollen tubes in the respective plant species.

These observations lead to the hypothesis that the transmitting tissue was active in moving the beads, and by inference pollen cells, towards the ovary. This hypothesis draws analogies to animal

systems, where cells can move by interaction with another biological surface (Hynes and Landers, 1992). The force for the movement of pollen tube was proposed to come from the interaction of the tube wall with the ECM of the transmitting tissue. As the pollen nuclei are moved down, the pollen tube is left as a trail with callose plugs separating the pollen cytoplasm from the rest of the tube.

This experiment and its subsequent interpretation is particularly interesting as it involved angiosperms with different carpel morphologies: *Raphanus* and *Vicia* have solid styles with dry and wet stigmas respectively while *Hemerocallis* has an open stylar canal. Therefore its implications can be applied to a variety of plants, rather than being restricted to one morphology/physiology.

Further evidence that the carpel may be active in pollen growth comes from microscopical observations on *Vicia*, *Raphanus* and *Phaseolus* that indicated that pollen tube growth only occurred on restricted regions of the stylar ECM (Lord and Kohorn, 1986; Hill and Lord, 1987 and Lord and Sanders, 1992). A physical basis for such controlled growth could also provide the basis of a pollen guidance system.

1.2.3 Recognition between the pollen and stigma/style.

Recognition of the pollen is an important aspect of the pollination process. The stigma/style not only differentiates between its own pollen and that of other species, but in many species can block self pollen from achieving fertilisation. the latter system is desirable as self-fertilisation would lead to inbreeding. These two types of pollen rejection have been reported in many species of plants and are referred to as heteromorphic incompatibility (foreign pollen rejection) and homomorphic incompatibility (self-pollen rejection).

1.2.3.1 Heteromorphic incompatibility.

All plants grown in the field will be exposed to the pollen of other species. The interaction between the carpel and foreign pollen varies between different species with four main classes of interaction being reported (Knox, 1984):

i) complete lack of support for foreign pollen which fails to germinate on the stigma e.g.. *Gladiolus* stigmas and *Gloriosa* pollen.

ii)support for pollen germination but not penetration (e.g.. *Gladiolus* stigmas and *Crocosmia* pollen).

iii)pollen germinates and produces a tube but arrest occurs in the style (e.g.. in cross species interactions in *Crocus*).

iv)pollen germinates and grows to the ovary where gamete fusion occurs but the resulting zygote is sterile.

Work on the physiological basis of these processes is rare. Some workers have looked for possible recognition factors in the first class of plants: Proteins were proposed to be involved in recognition in *Brassica*, where protease treatment of the stigma inhibited pollen binding (Stead et al., 1980) whereas in *Gladiolus* a carbohydrate epitope is involved in recognition as concanavalin A competes with pollen for stigma binding sites (Knox, et al., 1979). In *Brassica* the recognition factors could be developmentally

as germination of *Arabidopsis* pollen is supported by immature but not mature carpels (Kandasamy et al., 1994). The isolation of *Arabidopsis* mutants, deficient in surface lipid biosynthesis, which cannot support pollen binding suggests that a lipidic factor might be involved in these plants (Preuss et al., 1993) although it is unclear whether this mutant eliminates the recognition factor(s) or prevents proper presentation of these factors.

1.2.3.2 Homomorphic incompatibility

More substantial work has been carried out on homomorphic incompatibility. While some plants inhibit self-fertilisation by physical methods such as dioecy (reviewed in Dellaporta and Calderon-Urrea, 1993) others use physiological methods, in which cases the term self-incompatibility (SI)is used. This process is widespread throughout the angiosperms with presence/absence not being dictated by phylogeny. Due to this distribution SI has been proposed to be an ancestral condition to prevent inbreeding, indeed many SI plants such as *Brassica oleracea* show the symptoms of inbreeding depression if their SI mechanism is over-ridden (personnel observation). However how non-SI relatives of these plants, e.g. the Crucifer *Arabidopsis thaliana*, have overcome the inbreeding problem has not been investigated.

Genetic studies on a variety of plants have revealed the existence of two different systems depending on the origin of the incompatibility phenotype in the male gamete (DeNettancourt, 1977). In the gametophytic SI system the phenotype is controlled by the haploid pollen phenotype while in sporophytic SI compatibility is controlled by the diploid genotype of the anther (fig.3). Unlike the presence of SI the occurrence of these systems follows taxonomic lines and tend to be related to the physiology of the stigma (table.1). These rules are not hard and fast, the most notable exception being *Papaver rhoeas* which has a dry stigma but has a gametophytic SI system. A detailed description of two of the most studied examples of the two systems is given later.

1.2.4 The molecular investigation of the stigma and style.

The significant heterogeneity in the biology of stigmas and styles has been highlighted by the numerous studies on the structure and physiology of these organs in a variety of species. In contrast to this molecular biological approaches to studying the development and involvement in pollination of the various tissues, via the isolation and characterisation of expressed sequences, has been limited in its scope as illustrated in table 2.

Concentration on closely related species allows integration of different results which would be difficult to justify cross-species given the biological heterogeneity between different species. However it also means that any results/models are only applicable to the species they were produced in, so that a general model for the molecular basis of the stigma/style biology cannot be drawn from these results.

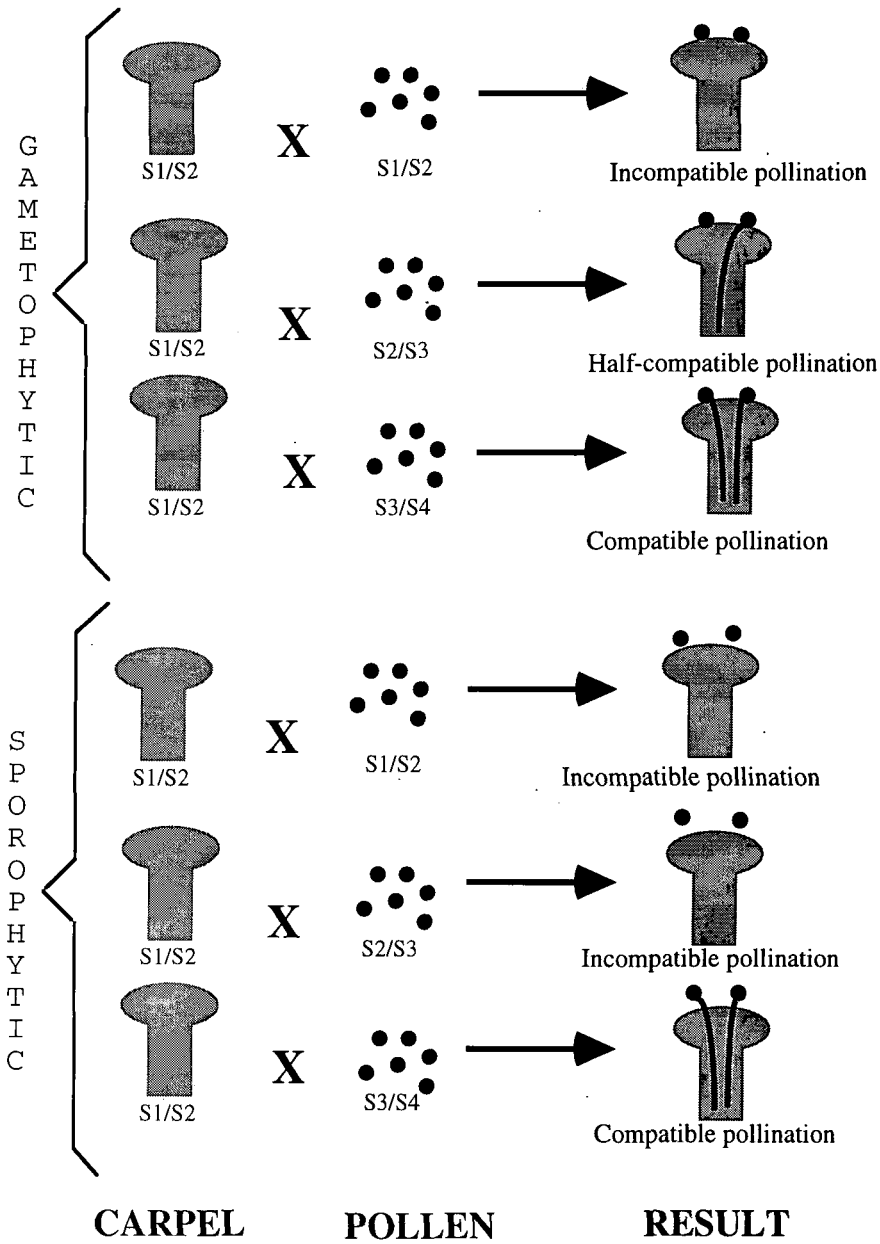


Fig 3. Compatible and non-compatible pollinations in the two principal forms of self-incompatibility. In the gametophytic system the pollen phenotype is determined by the pollen genotype so that in crosses involving two plants that share only one common allele half of the pollen grains will germinate. In the sporophytic system the phenotype of the pollen is determined by the sporophyte genotype. Therefore every pollen grain contains S-determinants from both parent-plant alleles, only one of which is required to induce an incompatible reaction, hence half-compatible pollinations are not possible.

	Sporophytic	Gametophytic
No. of loci	one	one or more
Control of reaction of pollen and stigma	parental	pollen
Stigmatic surface	dry	generally wet
Pollen cell no	tricellular	bicellular
Source of male component	tapetum	pollen
Site of pollen arrest	stigma	style, rarely stigma
Reaction to incompatible pollen	callose deposition	callose deposition

Table 1.Principal differences between gametophytic and sporophytic self-incompatibility systems in plants. After Knox (1984) and Heslop-Harrison and Shivanna (1977).

Family	Fields in which molecular studies have been carried out			
	Self-incompatibility	Pollen-carpel adhesion proteins	'Active style' components	Others
Solanaceae	yes	yes	no	yes
Cruciferae	yes	no	no	yes
Papaveraceae	yes	no	no	no
Lilaceae	no	no	yes	no

Table 2. Summary of the main areas of stigma/style biology in which isolation of cDNA's or proteins has been reported. Compiled from the references quoted in sections 1.3, 1.4 and 1.5.

The most extensive molecular work has been carried out on the Solanaceae and Cruciferae. Therefore the biology of these two families is discussed together with the molecular work carried out to investigate them.

1.3 STUDIES ON THE STIGMA/STYLE AND POLLINATION IN THE SOLANACEAE.

1.3.1 The Solanaceae as model systems.

Most of the molecular work on pollination has focused on the plant family Solanaceae. This family consists of several important crop plants such as *Nicotiana tabacum*, *Solanum tuberosum*, *Lycopersicon peruvianum* and *Lycopersicon esculentum* as well a several ornamental species such as *Nicotiana alata*, *Petunia hybrida* and *Petunia inflata*. The choice of these plants for molecular work has largely been dictated by their economic importance together with their well developed molecular biology. Transformation of *Nicotiana*, for promoter analysis with reporter genes (Jefferson et al., 1987) or the introduction of antisense constructs for inactivation of expression (Smith et al., 1988), is now routine using the Agrobacterium method (Marton et al., 1979) which has also been applied to the other members of the family (Klee et al., 1987).

As well as their economic importance, the biology of these plants has also promoted their use. Many species have large carpels (e.g.. stigma/style lengths are *L. peruvianum* ~12mm (Cresti et al., 1976), *P.hybrida* 25-30mm (Konar and Linskens, 1966a), *N.alata* ~60mm (Bell and Hicks, 1976)) and develop multiple flowers per plant so facilitating the harvest of tissue for protein/nucleic acid isolation. Also many species are self-incompatible although the pattern of distribution for SI is not phylogenetic.

1.3.2 Structure and physiology of the solanaceous stigma and style.

1.3.2.1 The stigma.

The Solanaceae have simple gynoecia containing only one carpel which consists of a simple stigma/style/ovary arrangement. All members of the Solanaceae have wet stigmas which have a degraded

cuticle and a rich exudate (reviewed in Knox, 1984). Under the classification of Heslop-Harrison and Shivanna (1977) the members of the Solanaceae predominantly fall into the group III classification although *Datura* and *Schizanthus* are classified as group IV.

Characterisation of the stigma has been reported for *P.hybrida* (Konar and Linskens, 1966a), *N.tabacum* (Cresti et al., 1986) and *Nicotiana sylvestris* (Kandasamy and Kristen, 1987). All contain four major cell layers: papillae, epidermis, secretary zone and storage zone. Two major structural configurations occur in the tissues, pre- and post-anthesis. Pre-anthesis the papillae consist of one (*P.hybrida* and *N.tabacum*) or two (*N.sylvestris*) cells and occur at low density as defined by the group III classification of these plants. In *N.sylvestris* the papillae in immature plants stain strongly for lipid protein and carbohydrate in their cytoplasm. The epidermis consists of a single cell layer and is coated by a thin cuticle which also coats the papillae. Exudate, which defines the Solanaceae as wet stigma-ed plants lies in a layer between the cellulose wall of the papillae, the epidermis and the cuticle. In *P.hybrida* this exudate is believed to come from the epidermal and papillar cells as vesicles in these cells contain the same electron opaque material as the exudate and can be seen fusing with the plasma membrane, presumably so as to deposit the material. At various points the cuticle ruptures and allows exudate to spill onto the stigma surface, usually as 4-6 major drops on each stigma.

The secretary zone and storage zone are both initially compact tissues in the immature organ. The latter subset of cells were defined in *P.hybrida*, by their differing morphology, compared to the secretary zone and underlying parenchyma, although no further differentiation studies have been reported. As the stigma matures large, schizogenous, cavities open up, in both zones, which fill with substantial quantities of exudate. In *N.sylvestris* the initial disintegration of these tissues occurs with the breakdown of the middle lamellae, although it is not known which event causes the other

Changes in these tissues become most pronounced post-anthesis. The cuticle is thrown off as flakes such that the entire stigma surface is covered in exudate. The exudate in the secretary layer is also given out between the epidermal cells which appear to loosen, although it is not known whether the exudate forces open the epidermis or whether changes in the epidermis allow release. The exudate washes away the last remains of the cuticle so forming a continuous layer on the stigma surface.

The stigma surface exudates of *P.hybrida and N.tabacum* have been partially characterised by Konar and Linskens (1966b) and Cresti et al. (1986) The electron-opaque nature of the exudate in *Petunia* suggested that it was lipid and this was confirmed by isolation of the exudate and heating to 105°C which failed to significantly reduce its mass. This component was estimated to constitute 80-85% of the total mass of the exudate. A similar situation occurs in *Nicotiana* where the exudate stains strongly for lipid.

Some water does occur on the stigma surface, beneath the lipid exudate as demonstrated by repeatedly blotting the stigma surface with cobalt chloride paper. Initial blottings removed exudate which had no effect on the paper but after the lipid had been removed the water could be blotted and induce the paper to change colour (Konar and Linskens, 1966b).

The components of the *P.hybrida* exudate were analysed by thin later chromatography which indicated that the oily component was pure lipid, being free of sterols, free acids and phospholipids. Three principal sugars, are sucrose, glucose and fructose in *Petunia* (Konar and Linskens, 1966b) while glucose

and mannose predominate in *N.tabacum* (Cresti et al., 1986). The *Petunia* exudate also contained low levels of free amino acids (Konar and Linskens, 1966b). No proteins were detected in the exudate by Konar and Linskens (1966b) although Herrero and Dickinson (1979) have reported acid phosphatase activity staining which occurred in small, presumably aqueous, droplets in the stigma-surface exudate. In *N.tabacum* the exudate stained only weakly with Coomassie (a blue protein stain), but some proteins were extracted and shown to have molecular masses of 10-60kDa by SDS-PAGE (Cresti et al., 1986).

1.3.2.2 The style.

The style, in the Solanaceae, is solid and consists of three tissue layers, epidermis, cortex (parenchyma and vascular) and transmitting tissue (Bell and Hicks, 1976), the latter being the one involved in the pollination process, has been the most extensively studied.

In *Petunia* two distinct types of cell have been reported in the transmitting tissue. At the neck of the style the cells are large and spherical while further down the cells are smaller, elongate and loosely packed (Herrero and Dickinson, 1979). In *L.esculentum* (Cresti et al., 1976), *N.alata* (Gane et al., 1994) and *N.tabacum* (Bell and Hicks, 1976) the loosely packed cells have also been reported but not the neck cells, although it is unclear whether this is due to true absence or a failure to observe them.

All the cells are surrounded by a fluid extracellular matrix which can be induced to flow from the plant after wounding (Cresti et al., 1976). This matrix is rich in carbohydrates, principally pectin, rather than cellulose, as indicated by its staining after pectinase and cellulase treatment (Cresti et al., 1976). The matrix also contains protein which in *L.esculentum* appears to undergo a period of rapid deposition just prior maturity (Cresti et al., 1976). In *Petunia* this protein includes acid phosphatases and peroxidases as indicated by histochemical staining (Herrero and Dickinson, 1979). The role of this protein is unclear as the neck cells in *Petunia* are particularly sensitive to protease treatment, while in the smaller cells are not (Herrero and Dickinson, 1979). In *Lycopersicon* a similar result has also been reported in the non-neck cells (Cresti et al., 1976).

Lying adjacent to the extracellular fluid are the cell walls of the transmitting tissue cells, which are predominantly cellulose, as indicated by their removal by cellulases (Cresti et al., 1976). In *N.alata* this was also confirmed by isolation of cell walls and H_2SO_4 hydrolysis which released significant amounts of glucose (>50% total sugar) while TFA, which hydrolyses cellulose poorly, did not (Gane et al., 1994). Other sugars detected in the wall by staining include arabinogalactans, β-D-Glucans and polyanions (pectins) while lignins and callose were absent (Gane et al., 1994). The protein content in the cell walls is higher (4.1% against 2.1%) than the rest of the style in *N.alata* (Gane et al., 1994).

1.3.2.3 Cell ultrastructure in the stigma and style

The ultrastructure of the cells of the stigma and transmitting tissue have also been investigated. Extensive smooth ER and golgi are present in the cells of the stigma and style respectively, with many of the observed vesicles fusing with the plasma membrane (Cresti et al., 1976; Konar and Linskens, 1966a; Bell and Hicks, 1976, Kandasmany and Kristen, 1987). This suggests that the cells are metabolically

active with much of this activity contributing to the generation of extracellular material, be it stigmatic exudate or the intercellular material of the transmitting tract.

As well as the extensive vesicle traffic, all of these specialised cells contain chloroplasts, with extensive grana, and starch bodies which indicates that they are likely to be photosynthetically active (Konar and Linskens, 1966a; Bell and Hicks, 1976). In *N.tabacum* this is contrast to the cortex where the thykaloids are poorly developed and starch grains small or absent (Bell and Hicks, 1976), which correlates with the hypothesis that they have a high level of metabolic activity. The starch bodies are believed to be a metabolic store to drive the production of exudate as the immature papillar cells, in *N.sylvestris*, are rich in stainable bodies which decrease as the carpel approaches maturity.

Two final aspects of note are the presence of crystalline bodies surrounded by a double membrane and myelin-like structures in *N.alata*, the latter being often associated with the ER. (Bell and Hicks, 1976). Neither of these structures have been reported in the transmitting tissues of the other Solanaceae suggesting they may be a specialised component in this plant, although what role they play is unknown.

1.3.2.4 Compatible pollination in the Solanaceae

When pollination occurs the first component of the carpel that a pollen grain interacts with is the stigmatic exudate. After landing in this the pollen grain begins to germinate after only 30min (Herrero and Dickinson, 1979) which would suggest that the exudate supports germination of the pollen. However experiments, *in vitro*, suggested that this is not the case: Pollen suspended in a drop of exudate failed to germinate and could only be induced to develop by the addition of water, while removal of this water and replacement with silica gel blocked germination (Konar and Linskens, 1966b). These results were interpreted as indicating that the exudate played no role in supporting germination but is a means of trapping pollen which sinks to the aqueous layer that lies between the exudate and stigma surface where rehydration and germination occur. Konar and Linskens also predicted that the exudate may act as a pollen-friendly protective cover for the stigma that stops transpiration in the absence of a cuticle. This is consistent with the observation that immature stigmas can support pollen germination, even though they lack exudate (Shivanna and Sastri, 1981; Kandasmany and Kristen, 1987)

After germination the pollen tube emerges, traverses the stigma surface and enters the transmitting tract, through which it grows to the ovary. Germination of the tube stimulates distinct changes in the transmitting tissue, both around and in front of the growing pollen tube in *P.hybrida* (Herrero and Dickinson, 1979). Cellular activity increases with a rise in the number of polyribosomes and vesicles. The latter contain a grey fibrillar matrix which is released into the extracellular space. A slight increase in starch reserves is also seen in front of the growing tube which is followed by a significant fall in reserves as the pollen tube passes the cells, contiguous with an increase in the relative volume of the vacuole compared to the rest of the cell. In the neck region of the transmitting tract cellular degeneration, indicated by an increase in intercellular disorganisation, release of cytoplasmic contents and cytoplasmic granularisation, has also been observed (Herrero and Dickinson, 1979)

The changes that pre-empt the arrival of the pollen tube must be triggered by some signal in the carpel. The nature of this signal is, however, currently unknown.

1.3.2.5 Self-Incompatibility in the Solanaceae.

Several members of the Solanaceae have a single-locus gametophytic SI system in which pollen tube arrest occurs in the transmitting tissue. More than 40 alleles are believed to occur in natural populations (Lawrence, 1975). The ultrastructural aspects have been most extensively reported in *L.peruvianum* (DeNettancourt et al., 1973) where the gross differences between compatible and incompatible pollinations only occur after the pollen tube has grown through the first third of the style. The first detectable changes are that the inner pollen tube wall, which consists of callose, (Miekle et al., 1991), 1992) disappears while the outer arabinan wall (Rae et al., 1985) thickens. Callose is deposited around the incompatible tube, while numerous particles consisting of an electron dense core and a less dense coat accumulate in the pollen cytoplasm. When the cytoplasm became completely loaded with these particles the tube tip bursts and releases its contents into the extracellular space of the transmitting tissue.

While the most obvious differences between compatible and incompatible pollinations occurs in the pollen tube detectable differences have also been reported in the carpel (Herrero and Dickinson, 1979). The cellular starch reserves are not depleted to the extent of that seen in compatible pollinations. However polyribosome and vesicular transport increases in the stylar neck still occur during incompatible pollinations although, as the pollen tube growth through this region was physiologically similar to that seen in compatible pollinations, this was as expected.

1.3.3 Molecular Characterisation of the female tissues of the Solanaceae.

Given their amenability to genetic manipulation, as discussed in section 1.3.1, the majority of the work on the molecular characterisation of the female tissues has concentrated on the Solanaceae. From the structure of the relevant tissues in these plants some predictions about the genes and proteins, that would expected to be identified, can be made: In the stigma enzymes involved in lipid production together with proteins involved with extracellular transport would be expected as much of the tissue in this sub-organ is involved in production of exudate. In the style components that constitute the specialised nature of the transmitting tract occur, which would be predicted to include specialised cell wall and apoplast proteins, both structural and metabolic, so as to produce an environment through which pollen tubes can grow. Biosynthesis of such proteins would also require enzymes involved in the generation of glycan moieties and protein export. Given the possible active nature of the style, adhesion molecules such as arabinogalactan proteins or vitronectin homologues would be expected together with factors involved in pollen guidance. Finally the proteins involved in the recognition and rejection of self pollen are predicted to reside in the transmitting tissue

In studying the female tissue in this family, the majority of the isolated sequences have been obtained by focusing on cDNA's or proteins, initially identified by their appearance in the stigma/style and absence in other tissues. However targeted studies have also been carried out, predominantly on isolation of the female component of the Solanaceous SI system. As the relevant genes and proteins involved in this system have been the most extensively studied these will be discussed first.

1.3.3.1 The Solanaceous gametophytic S-locus.

Much of the work on the genes and proteins of the carpel has concentrated on isolation of the components of the S-locus. The Solanaceae are amenable to such a study as they have a single S-allele which simplifies the generation of mono-allelic strains.

Candidates for the female S-factor in *Nicotiana alata* were identified by comparison of the stigma/style protein patterns of different S-genotypes by 2D gel electrophoresis (Anderson et al., 1986). Putative S-linked proteins stained for carbohydrate and so were termed SLSG's (S-linked style glycoproteins) Sequence information from one of these proteins was then used to isolate a cDNA encoding the protein (Anderson, 1986). This nucleotide sequence, in turn, allowed other S-alleles from *N.alata* (Anderson et al., 1989) *S.tuberosum* (Kaufmann et al., 1991), *P.inflata* (Clark et al., 1990) and *L.peruvianum* (Tsai et al., 1992) to be isolated. Many alleles showed greater inter-species than intra-species homology, suggesting that this gameophytic SI system predates speciation in the Solanaceae (Tsai et al., 1992).

Initial evidence that these cloned genes constituted the female component of SI came from their genomic organisation and expression. In *N.alata* the putative S-genes were encoded by a single locus and co-segregated with their respective S-haplotype. Immunolocalisation indicates that the putative SLSG is found on the stigma surface, in the transmitting tract and on the inner epidermis of placenta (Anderson et al., 1989). which correlates with the path of pollen tube growth and encompasses the site of incompatible pollen arrest. However this expression is not specific as SLSG has also been detected in immature anthers by RT-PCR (Clark and Sims, 1994; Dodds et al., 1993). Reporter gene experiments also suggest that the SLG gene may also be active in petals (Clark and Sims, 1994)

Absolute confirmation of the role of the SLSG in SI came from studies in which transformation of both *N.alata* (Murfett et al., 1994) and *P.hybrida* (Lee et al., 1994) with SLSG genes, not present in the transformed plants, conferred the S-phenotype associated with the introduced gene. Similarly introduction of an antisense S3 S-RNase construct into S2S3 *P.hybrida* eliminated the ability of carpels to reject pollen carrying this allele (Lee et al., 1994).

1.3.3.2 Function of the Solanaceous S-proteins

While the above experiments confirmed the fact that the SLSG was the sole female determinant in the gametophytic SI system in the Solanaceae, they provided no data on how the protein mediates recognition or pollen arrest. Initial clues to these problems came from analysis of the primary structure of the SLSG protein in which two short stretches of amino acid sequence show high homology to sequences around the catalytic domains of fungal ribonucleases from *Aspergillus oryzae* and *Rhizopus niveus* (McClure et al., 1989; Kawata et al., 1990). Evidence that the SLSG might also be an RNase come from fractionation of the stylar proteins on *N.alata* indicated that 40-80% of stylar RNase activity copurified with the SLSG (McClure et al., 1989).

The role of such an enzyme in SI was investigated *in vitro* Addition of an SLSG to incompatible pollen led to pollen arrest and death, similar to that observed *in vivo*, allied to degradation of the pollen tube rRNA (Jahnen et al., 1989; McClure et al., 1990) although the effects on pollen in this assay were

phenotype is also the only female factor required to induce pollen arrest. Therefore the programme of particle formation and callose deposition is mediated by the pollen as opposed to the transmitting tissue.

These results were also interpreted by the authors to indicate that the rejection of self-pollen is mediated via degradation of pollen RNA. Evidence supporting this hypothesis came Kowyama (1994) who showed that a self-compatible line of *L.peruvianum* lacked stylar RNase activity. Similarly, transformation of S1S2 *P.inflata* with S3 genes, containing mutations in the two histidines conserved between all the characterised SLG's and the fungal RNases failed to confer an S3 phenotype on the plant, even though SLG3 protein levels in these plants were as high as when wild-type S3 genes were used (Huang et al., 1994). This failure to confer phenotype was also mirrored by a lack of RNase activity in the mutant proteins.

While these results appeared to confirm a role for the RNase in pollen rejection they did not explain other aspects of the rejection response, i.e. how particle accumulation and callose deposition are induced, although it is possible that the cessation of the production of certain proteins may trigger these responses. A second objection to this model comes from the fact that heat treatment of the SLSG destroys ~98% of RNase activity but does not affect the effect on pollen growth *in vitro* (McClure et al., 1989). This would appear to suggest that rRNA degradation is only a side effect of the SLSG whose principal function lies in other, uncharacterised properties of the protein. If this is correct then the inactivation of SLG by mutation of its conserved histidines presumably acts via conformational change influencing other properties of the protein as well as its RNase.

To determine the means by which SLSG triggers the SI response, in pollen, the male recognition component(s) needs to be isolated and its interaction with SLSG characterised. However no experimental work which identifies or characterises the pollen factors that recognise SLSG allowing it to enter to the pollen tube and trigger the pollen response has yet been reported. Attempts to identify the male component via identification of pollen proteins that co-segregate with the S-phenotype have proven unsuccessful (Newbigin et al., 1993). One hypothesis that would explain these problems is that the male component may only be expressed once the pollen tube has started growing (Dodds et al., 1995). This would explain the fact that pollen tubes germinate and begin growing before they are arrested in gametophytic plants. Therefore a different approach such as mapping the genomic region around the SLSG gene for other components in the locus, as has been used in sporophytic plants (Boyes and Nasrallah, 1995; discussed in section 1.4.3.1) could be the next logical approach.

1.3.3.3 'Stress responsive' genes and proteins in the carpel.

One of the largest class of expressed sequences identified in the solanaceous carpel comprise cDNA's and proteins that are members of sequence families, previously identified by their role in the response to stress. A number of such sequences have been identified in the stigma and style of various members of the Solanaceae.

(1,3)ß-glucanases.

The first stress-responsive protein of this class to be identified in the stigma and style was (1,3)ß-glucanase. A polyclonal antibody, raised against an acidic form of this enzyme induced by TMV infection,

The first stress-responsive protein of this class to be identified in the stigma and style was (1,3)ß-glucanase. A polyclonal antibody, raised against an acidic form of this enzyme induced by TMV infection, cross-reacted with a protein present in *N.tabacum* carpels (Lotan et al., 1989). Immunolocalisation revealed that this protein was expressed in the stylar transmitting tract and in a dumbbell-shaped subset of cells below the stigma surface. Contradictory reports on the sub-cellular localisation of the glucanase occur as Lotan et al. (1989) report vacuolar expression while Ori et al. (1990) claim that the protein is present in the apoplasts. The latter would appear more likely as other acidic (1,3)ß-glucanases have been localised to the apoplast, while vacuolar expression is normally restricted to the basic forms of these enzymes (Bowles, 1990).

Chitinases

The presence of chitinase in floral tissue was first detected by probing northern blots with probes for previously described genes (Memelink et al., 1990). Proteins that cross-reacted with anti-acidic and anti-abscission zone endochitinases have been reported in *N.tabacum* ovary, anther, sepal and pedicle tissue (Lotan et al., 1989). and the stigma, but not the style of *Phaseolus vulgaris* (Campillo and Lewis, 1992).

Molecular characterisation of stigma/style chitinases has been reported for *S.tuberosum* and *L.esculentum*. The former was initially identified as the most abundant basic protein in stylar extracts, which was purified and partially purified to obtain sequence information which was used to clone the relevant cDNA (Wemmer et al., 1994). The latter was isolated as a cDNA clone in a differential screen of a carpel library which was subsequently used to isolate a genomic clone (Harikrishna et al., 1996). Endochitinase activity has been directly demonstrated for the former while high levels of enzyme activity have been detected in the style which was interpreted as indicating that the latter is also a chitinase, although this has not been conclusively proven

Both cDNA's share >66% identity to class I basic chitinases while the level of identity between them is 91%. Localisation using a reporter gene construct using the promoter of the *L.esculentum* gene and antibodies raised to the protein for the *S.tuberosum* indicate both are expressed in the stylar transmitting tissue.

The *L.esculentum* cDNA was used to probe northern blots containing carpel tissue from other species including *N.tabacum*, *P.hybrida*, *B.napus* and *Antirrhinum majus*. No hybridisation at low stringency was reported suggesting that the presence of these chitinases is restricted to the Solanoidae sub-family. However this is contrast to the work of Wemmer et al. (1994) who detected chitinase activity in the styles of *N.tabacum* and *P.inflata*, although only the *Petunia* chitinase cross-reacts with antibodies raised against the *Solanum* protein. This suggests that chitinases are present in the members of the Solanaceae but sequence divergence means that they don't cross-hybridise with the *L.esculentum* cDNA. However, the mode of chitinase expression in the Solanaceae is not ubiquitous as chitinase activity in *Petunia* is restricted to the stigma and is not detectable in the style (Leung, 1992).

Proteinase inhibitors

Homologues of proteinase inhibitors were first detected in floral tissue by the use of antibodies raised against wound induced *Solanum* class II inhibitors. These cross react with proteins in potato buds

differential screening of a *N.alata* cDNA library. A class of transcript with ~50% identity to the potato class II inhibitor was found to constitute 3% of the screened clones which encoded a protein consisting of six homologous domains (Atkinson et al., 1993). The clone hybridised with two differently sized transcripts in the stigma/style which are also present in the ovary, petal and sepal. Hybridisation, *in situ*, indicated the expression in the stigma/style was restricted to the upper layers of the stigma (secretary zone and stigma surface).

An inhibitor protein was isolated from the stigma, by its activity against trypsin and chymotrypsin and subsequently shown by N-terminal sequencing to match a polypeptide in all six repeats of the protein, predicted from the cDNA (Atkinson et al., 1993). As the mass of the isolated protein was only 6kDa it was proposed to have arisen by post-translational processing of the cDNA's putative protein, a situation mirrored in other inhibitors (Pearce et al., 1982). Evidence, backing this hypothesis, came from antibodies raised against the 6kDa protein which cross reacted with higher molecular weight proteins. No further work on the expression processing and role of this protein has been published.

γ-thionins

A fourth class of possible stress-response gene has been isolated from *Lycopersicon* by differential screening (Milligan and Gasser, 1995). The predicted protein sequence of this clone is 66% identical to a γ-thionin from tobacco flowers that had not been shown to be carpel specific (Gu et al., 1992). The γ-thionins are a family of small acidic proteins which show sequence and structural differences from other classes of thionins, which have been implicated in defence against pathogens (Garcia-Olmedo et al., 1989). The precise role of the γ-thionins is unkown but they have been shown to inhibit protein synthesis in cell free systems (Mendez et al., 1990).

PR-gene promoters and the tissues of the carpel

As well as the direct isolation of carpel genes and proteins that are normally only stress induced in vegetative tissue, evidence for the role of such factors in the reproductive tissue has come from studies using the promoters of previously characterised stress-induced genes. An *Oryza sativa* chitinase promoter drives strong expression in healthy floral tissue when introduced into tobacco (Zhu et al., 1993). Expression is highest in the stigma and lower in the ovary and pollen while in the rest of the flower expression is much lower than in the other two organs. The promoter also confers inducement by wounding and pathogen attack to the vegetative tissues, which mirrors the expression of the wild type gene in *Oryza*. Similarly a hydroxyproline rich glycoprotein promoter from *Phaseolus* drives both stigma-specific and wound-induced expression of a linked reporter gene in *N.tabacum* (Wycoff et al., 1990).

These results suggest that promoters that drive stress-responsive expression can also be active in healthy floral tissue in the Solanaceae. However, the relevance of these results to the wild-type plant must be considered, as the introduced promoters are essentially foreign sequences to *N.tabacum* and so may not reflect the behaviour of native genes. Such data is also difficult to relate to the cDNA's and proteins described above, all of which differ from previously characterised stress-induced sequences and so are not transcribed/translated from the same genes. Similarly the reporter-GUS fusions used in these experiments can give misleading results. One example of this is the *N.tabacum* PR1a promoter which is highly active

in pollen when linked to a GUS reporter gene but appears inactive in this tissue as part of the wild type gene (Uknes et al., 1993).

Such problems are not relevant to a third 'stress' promoter active in the carpel. This is the *Solanum* PR10a gene whose promoter can confer strong expression in the stigma, particularly the upper layer, as well as in response to pathogen and wound inducement (Constabel and Brisson, 1995). This is backed by immunolocalisation of the PR10a protein to the stigma which suggest that this gene does drive both stress-responsive and carpel-specific expression. Unfortunately the function of the PR10a protein is unknown, although outside the stigma it is often associated with the vascular bundles. Therefore its role in the tissues of the carpel is less easy to speculate about than the other identified stress-induced genes and proteins.

1.3.3.4 Possible roles of PR-proteins in the carpel.

The role of the identified PR-proteins in the stigma and style is still a cause for speculation. One possible role for them is defensive, analogous to their PR function (Wemmer et al., 1994; Atkinson et al., 1993): The wet stigmas and transmitting tissue of the Solanaceae are nutrient-rich and should provide an attractive target for pathogenic fungi and bacteria, which would be a particular problem if they can be transmitted by pollinating insects. However infection of the carpel is rare (Atkinson et al., 1993) which may be due to constitutive expression of proteins that can protect against invasion. Evidence for a defensive role comes from the behaviour of pathogenic fungi which can infect the upper parts of a carpel but are inhibited in the transmitting tissue of the style (Jung et al. 1956). Such behaviour has been used as evidence for a defensive role for the (1,3)ß-glucanase and *Solanum* and *Lycopersicon* chitinases which are both restricted in expression to this part of the reproductive tissue (Lotan et al., 1989; Wemmer et al., 1994; Harikrishna et al., 1996). The expression within the carpel has been proposed to enable the plant to counteract the development of resistance to the defence proteins by only attacking the pathogen at a late stage in development.

However a defensive role is inconsistent with the pattern of expression of the *Petunia* chitinase, the class II proteinase inhibitor and PR10a which are all found at or under the stigma surface. It could be argued that these genes are involved in defence against different types of pathogen or stress, though confirmation would require characterisation of these stresses which has not been reported.

A second possible role for these sequences is one directly involved in the physiology of the healthy carpel, where the PR proteins play a role in the pollination process. Such a model is consistent with a growing body of evidence that PR-proteins may play an active role in other aspects of plant development. The most conclusive evidence for such a role comes from germination in *Daucus carota*, where a temperature sensitive mutant, that arrests embryo development, can be rescued by media conditioned by wild type embryos (de Jong et al., 1992). The active component of this media was isolated and found to be a 32kDa glycoprotein with endochitinase activity. Further evidence that stress-induced sequences may play a role in healthy embryo development comes from the activity of a (1,3)ß-glucanase promoter during *N.tabacum* germination (Vögeli-Lange et al., 1994). Expression is localised in the

promoter during *N.tabacum* germination (Vögeli-Lange et al., 1994). Expression is localised in the endosperm, prior to the penetration of the radicle, although no direct role has been established in germination for this gene.

Direct evidence that a non-defensive role for these sequences occurs in the carpel comes from the *Petunia* chitinase which increases in activity by 400% after anther dehiscence (Leung, 1992), which correlates with the peak of carpel receptivity to pollen. however the relevance, if there is any, of this rise to the capacity of the carpel to support pollen growth is unknown. Moreover, this rise in activity can bear little relevance to other plants due to the different patterns of expression of chitinases between *Petunia*, *Lycopersicon* and *Solanum*.

The major theoretical argument against such a role is that the substrates for many of the enzymes are not present in the carpel. Callose (ß(1,3)glucan) is deposited around the pollen tube rather than being degraded and there is no reports of chitin occurring in any plant tissue (de Jong et al., 1992) and the class II proteinase inhibitor is only believed to act on fungal and insect proteinases (Ryan et al.,1984). It is however possible that these proteins may have different biochemical functions in the stigma and style to those detected in defence. Evidence that this may be so comes from the work on *D.carota* germination where the chitinase was essential for development even though no chitinous substrates were present in the relevant plant tissues (de Jong et al., 1992). This was interpreted as indicating that the chitinase was not acting in germination via its chitinase activity. Support for this hypothesis comes from the discovery that N-acetylglucosamine-linked lipo-oligosaccharides (normally associated with the induction of nodulation by *Rhizobium* (Spaink, 1992)) can substitute for the chitinase in rescuing *D.carota* embryos (de Jong et al., 1993). Therefore the chitinase may be acting to produce similar molecules from a novel substrate *in vivo* (Spaink et al., 1993).

It is also possible that the stress-responsive sequences in the carpel, may also have physiological roles involving different biochemical activities than those recorded for the related stress-responsive sequences. However as more of these sequences are isolated from the reproductive organs, it becomes more unlikely that all of the carpel sequences will have different activities to those of their relatives in defence

Therefore the stress-induced sequences provide more questions than answers about the molecular processes that underlie the physiological nature of the carpel and its role in supporting pollination. Given the routine genetic manipulation techniques available for the Solanaceae manipulation of the relevant genes should be relatively simple allowing analysis of their role as discussed in section 5.2.

1.3.3.5 Proline/Hydroxyproline-Rich Proteins

Proteins with a high proline or hydroxyproline (PRP's) content have been hypothesised as being major components of plant cell walls for many years due to the detection of high levels of these amino acids in crude protein fractions from cell walls (Lamport, 1969). Molecular characterisation of these proteins has led them to be classified into three families by their sequence or mode of glycosylation (Showalter, 1993). The extensins (hydroxyproline rich glycoproteins or HRGP's) are characterised by SP_4 repeats where the prolines are hydroxylated and glycosylated. The Arabinogalactan proteins (AGP's) are characterised by their precipitation by Yariv reagent (1,3,5-tris[ß-glucopyranosyl-oxyphenolazo]-

1975). The Proline-rich protein (PRP) group constitutes other cell wall proteins with a high proline content.

Examples of all three proline-rich protein families have been isolated from the stigma and style of Solanaceous plants as catalogued in table 3. The majority of the sequences were isolated by random approaches (differential screening or isolation of highly expressed proteins) with only NaPRP4 and NaPRP3 being obtained by workers specifically looking for that class of sequence. This suggests that there is a bias for proline-rich sequences in the *Nicotiana* stigma and style as opposed to glycine-rich proteins, for which no examples have yet been identified.

Although all the sequences were derived from *Nicotiana* spp. there is evidence that they occur in other Solanaceous stigmas and styles. Carpel AGP's were detected by diffusion into a gel containing ß-glucosyl Yariv reagent in *N.tabacum*, *N.alata* and *L.esculentum* (Hoggart and Clarke, 1984), while proteins, that cross react with antibodies raised against GaRSGP have been detected in the styles of various *Nicotiana* and *Lycopersicon* spp. The AGPNa3 cDNA, however, only hybridises to *Nicotiana* carpel RNA and not that of *L.peruvianum.*

1.3.3.6 Role of proline-rich proteins in pollination.

Unlike the available data on the stress-induced sequences, there is a growing body of evidence suggesting that some of the proline-rich proteins are involved in the pollination process. Some of this comes from changes in the expression of the sequences in response to pollination. The level of class I extensin transcript increases on pollination while that of class II and III falls (Goldman et al., 1992). Similarly pollination causes a fall in the level of AGPNa3 (Du et al., 1996) although crude analysis of the level of AGP's in the stigma and style suggest that they increase in the former but remain constant in the latter (Gell et al., 1986). This response to pollination suggests that the encoded proteins may effect the pollination process either positively (stigma AGP's, class I extensin) or negatively (class II and III extensin, AGPNa3). However the time scale for these changes is not always consistent with a role in pollination. For AGPNa3 and the extensins the changes in expression occur over a period of 5 days which is much longer than the time taken for *Nicotiana* pollen to achieve fertilisation and in fact stops just before stigma/style abscission on day 6. Therefore it is probable that the changes are involved in, or the result of the death of the stigma and style tissues.

More substantial evidence for a role in pollination comes from the TTS proteins. Addition of purified TTS1 to *in vitro* pollen growth media accelerates pollen tube growth by a maximal three-fold at a concentration of 2µg/ml and also acts as an attractant to the growing tubes (Cheung et al., 1995). Coupled to this behaviour was the fact that TTS1 bound to and was apparently de-glycosylated by pollen tubes, the latter being detected by migration rates on SDS-PAGE gels (Wu et al., 1995). Attempts to remove this de-glycosylating activity from the pollen tube by washing were unsuccessful indicating that it must act via binding of the TTS protein. This de-glycosylation has been proposed to de-activate TTS as the de-glycosylated protein cannot influence pollen tube growth *in vitro* (Cheung et al., 1995).

		Type of Sequence	Dominant motif	Expression	Notes	Reference
EXTENSINS	CLASS I	cDNA	SP4	transmitting tissue		Goldman et al., 1992
	CLASS II	cDNA	SP4	transmitting tissue		
	CLASS III	cDNA	SP4	transmitting tissue		
	CELP 1c	cDNA/ genomic	XP(3-5) and XP	Between transmitting tissue and cortex in pistil. In the vascular tissue in rest of flower	Contains extensin and cysteine rich domains	Wu et al., 1993
	CELP 2c	cDNA				
	CELP 3c	cDNA				
	CELP 4c	cDNA				
	NaPRP3	cDNA		transmitting tissue	Contains extensin and AGP motifs	Chen et al., 1992
AGPS	AGPNa3	cDNA protein	Hyp/ProAsx/ Glx rich	Stigma, weak in transmitting tract	Protein sequence is not of the classical AGP type	Du et al., 1996
	AGPNa1	cDNA protein	Hyp/ProAla/ Ser rich	Multiple plant tissues	Has a classical AGP protein structure	Du et al., 1994
	TTS1	cDNA protein	KP2	extracellular space of transmitting tissue	Non-AGP protein structure but reacts with Yariv	Cheung et al., 1993
	TTS2	cDNA protein	KP2	extracellular space of transmitting tissue	Non-AGP protein structure but reacts with Yariv	Cheung et al., 1993
OTHERS	NaPRP4	cDNA	KP2	style	~97% identity at nucleic acid and predicted protein level to TTS cDNA's/proteins	Chen et al., 1993
	GaSRP	protein	KP2	inner cell wall of transmitting tract	Putative protein encoded by NaPRP4. Non reactive to Yariv reagent so is not an AGP	Sommer-Knudsen et al., 1996

Table 3. Proline rich protein sequences isolated from the stigma and style of Solanaceous plants. All sequences were derived from *N.alata* except for the class I-III extensins, the CELP cDNA and genomic sequences. and the TTS sequences which were derived from *N.tabacum*. References are for the sequence and expression of the cDNA's and proteins and are not an exhaustive list for each sequence.

To determine the relevance of these *in vitro* results to the TTS proteins in the carpel, glycosylation levels were studied *in vivo*. This provided three lines of evidence suggesting TTS proteins interacted with pollen:

(i) The level of TTS glycosylation, as measured by size on SDS-PAGE gels, decreases after pollination (Wu et al., 1995).

(ii) Isolation of TTS from different parts of the style shows that there is a TTS glycosylation gradient being highest at the base of the style and lowest at the stigma/style junction (Wu et al., 1995).

(iii) Elimination of TTS expression in the style by antisense technology reduces the pollen growth rate but by variable amounts (Cheung et al., 1995).

These results indicate that TTS may play a role in pollen tube growth and guidance, although the exact mechanism of this is not known. Wu et al. (1995) have proposed that TTS proteins bind to the pollen via an electrostatic interaction between the lysines of the TTS protein and pectins which are abundant in the pollen coat (Li et al., 1994). By interacting with pollen the TTS proteins act as a chemo-attractant, or may influence the pollen cell wall increasing flexibility and so promoting growth around the tip.

The role of de-glycosylation and its effect on binding is also unknown. Wu et al. (1995) predict that if TTS is a chemo-attractant then the glycosylation gradient may act as an indicator of the level of sugars in the style inducing the pollen to grow towards the ovary. Likewise pollen-mediated glycosylation may be used to deactivate the chemo-attractant ability of the TTS proteins. The removed sugars could then be utilised for metabolic energy for the growing pollen tube. This however is all speculation and requires further elucidation.

1.3.3.7 Other genes expressed in the solanaceous stigma and style.

As well as the genes and proteins, described above, several other sequences have been isolated from the female tissue, of the Solanaceae, which may play a role in the biology of the reproductive organs. Two of these, isolated from *L.esculentum* encode cell-wall metabolising enzymes: pectate lyase (Budelier et al., 1990) and endo β(1,4)glucanase (Milligan and Gasser, 1995) which could act to generate the extra-cellular environment of the specialised tissues of the stigma/style. The former is represented by two cDNA's, sharing 94% identity which were shown to be restricted in their expression to a subset of cells in the transmitting tissue. *In situ* hybridisation indicated that transcripts cross-hybridising with these two cDNA's occurred in the outer layers of the transmitting tract but only in the upper two thirds of the style. The upper third of the style showed strong expression while in the middle third a gradient of expression occurred, decreasing towards the ovary. This pattern of expression lead to the hypothesis that the transmitting tissue is not homogeneous, but exists as distinct biochemical zones.

The transmitting-tissue expression suggests that the pectate lyase could act on the pectins that have been reported in the apoplast of this tissue (Cresti et al., 1976). Digestion of these polysaccharides

could be involved in promoting pollen growth through the tissue, but if this is so than why the cDNA is only expressed in the upper part, rather than the entirety of the style is not known.

The glucanase, however has not been localised, even to the sub-organ level, although it is known to be expressed at the RNA level only in immature carpels and so is unlikely to play a role in modifying the extracellular environment in the mature pollinated flower.

Several other carpel-expressed cDNA's were also isolated from the same screen that produced the β(1,4)glucanase. These include two putative leucine aminopeptidases, a putative Fe2+/ascorbate dependent oxidase, a protein similar to the taste-modifying protein miraculin and an unknown sequence (Milligan and Gasser, 1995). None of these sequences are carpel-specific and none have been localised to any of the carpel organs. Therefore no predictions on the roles that these sequences could play in the stigma and style can be made.

1.3.3.8 The molecular description of the stigma/style in the Solanaceae.

The molecular studies carried out on the stigma and style of the Solanaceae have met with mixed levels of success. Characterisation of the SI system is well advanced with the development of models to explain how the response works on the molecular level. Similarly work on the PRP's has produced information that provides clues to how the female tissues interact with compatible pollen. The stress-response family sequences, however, pose more questions than answers, while little information, so far, has been produced about enzymes that shape the transmitting tissue apoplast or produce the stigmatic exudate.

1.4 The stigma/style and pollination in the Cruciferae.

1.4.1 The Cruciferae as model organisms

Molecular analysis of pollination has also been carried out on the Cruciferae. As with the Solanaceae much of the work on this family of plants has been because it contains a number of economically important plant species: *Brassica napus* and *Brassica campestris* are both cultivated for the oil produced in their seeds while *Brassica oleracea*, *Brassica rapa*, *Raphanus raphinastrum* and *Raphanus sativus* are all important food crops.

This family also contains the species *Arabidopsis thaliana*. Unlike the plant species described above this plant has no economic value but is used as a model organism for molecular and genetic studies on plant biology (reviewed in Somerville, 1989). Four factors have promoted this plant as an experimental system:

(a) physical size. *Arabidopsis* is small enough such that individual plants can be grown in a single test tube. Therefore large numbers of plants can be cultivated for screening of mutations with minimal space requirements.

(b) genome size. *Arabidopsis* has a relatively small genome size, being approximately one tenth of that of *B.oleracea* and one fiftieth of *N.tabacum* (Arumaganathan and Earle, 1991). A smaller amount of target DNA means that less mutants need to be

screened for desirable phenotypes in random mutation studies. The small genome size has also lead to work to map all the genes of this organism in the hope that a complete genetic map, if not the complete sequence of the genome will be obtained (Pruitt et al., 1986).

(c) self-compatibility. *A.thaliana* is naturally self-compatible so enhancing the generation of homozygous plants after mutation.

(d) Fast life cycle. *A.thaliana* can go from seed to seed in 5 weeks (Alberts et al.,1989). This coupled with (a) and (c) makes the plant ideal for crossing to generate homozygous mutants and to check for complementation between different mutants.

The utilisation of these properties is well illustrated by the isolation and characterisation of genes involved in flower patterning from *Arabidopsis*, where mutants were identified then the relevant genes mapped and cloned (Coen and Carpenter, 1993; Okamuro et al., 1993).

However *Arabidopsis* has been used sparsely for the study of the pollination process with only some biological studies, where it has been studied in conjunction with other species (Elleman et al., 1992; Kanadasmany et al., 1994). A principal reason for this is that as *A.thaliana* is so small, harvesting of floral tissue for nucleic acid/protein harvesting would take substantially more work than from other species. Therefore physiological and molecular characterisation of the pollination process in these plants has concentrated on other species with larger flowers, predominantly *B.oleracea* and *R.raphinastrum.* Secondly *A.thaliana* has no SI system, which has been the most intensively studied aspect of pollination in the Cruciferae which has lead to molecular work concentrating on the other species described above.

As very little molecular work has been carried out on aspects of the carpel and pollination, other than SI, then any discussion must focus on the recognition system. However before this molecular work is described the carpel morphology and the physiology of the pollination process in these plants will be outlined.

1.4.2 Biology of the cruciferous carpel.

1.4.2.1 Anatomy and physiology of the stigma and style

As with the Solanaceae, the Cruciferae gynoecium consists of a single carpel consisting of a simple stigma/style/ovary combination. The carpel does however differ significantly from the previously discussed species. Predominant in the differences when considering pollination is the nature of the stigma which is dry and papillate which has led to a group IIa classification (Heslop-Harrison and Shivanna, 1977). As the stigma surface is densely covered in papillae it is the surface of these cells with which the pollen grain must first interact and so which has been most intensively investigated. These investigations have led to the identification of three distinct components lying outside the plasma membrane of the papillar cells (fig.4):

The pectocellulosic wall appears under the electron microscope as an amorphous fibrillar matrix (Roberts et al., 1984a). This is not a single entity but instead consists of two distinct layers observable under the EM around the tip of the papillar cells (Elleman et al., 1988). Near the base of the cells these layers become less distinct so that the wall appears to consist of a single layered structure. Regional differences also occur between the contact point between the plasma membrane and the pectocellulosic

layer: Near the tip of the papillae the plasma membrane is more irregular than near the base, and its junction with the pectocellulosic layer is often interrupted by vesicular inclusions

The cuticle is a waxy covering, similar to that seen over non-reproductive tissues in the plant. This layer has been reported to be discontinuous, forming what have been described as globular droplets (Roberts et al., 1984a), or bars and rods (Mattson et al., 1974), on the papillar surface. These droplets contain intrusions of the pectocellulosic layer (Elleman et al., 1988).

The pellicle is a proteinaceous layer that covers the cuticle and was first identified by Mattson et al., (1984) who stained the surface with both 1-ANS (protein stain) and fluoroscein diacetate (non-specific esterase stain). Both these stains identified a previously unreported, pronase sensitive, lipase insensitive layer ensheathing the cuticle. The pellicle could also be observed under the EM after freeze-drying (Mattson et al., 1984) or by the use of lanthanum nitrate as a surface tracer (Heslop-Harrison et al., 1975). Staining with the latter indicates that the pellicle runs from papillae to papillae forming a continuous surface over the whole stigma surface (Heslop-Harrison et al., 1975). The pellicle is not a specific component of the cruciferous stigma but has been detected in a variety of species and so has been proposed to be a common element in plants in stigma-groups I and II (Heslop-Harrison et al., 1975).

The extracellular proteins of the papillae have a high rate of turnover, compared to the rest of the cell as initially indicated by ^{3}H-leucine incorporation (Roberts et al., 1984b). Quantitation of this turnover came from incubation with the protein synthesis inhibitor cyclohexamide which reduced the level of protein in the papillar cell wall by three times the amount seen in other plant cells

Beneath the papillar cells lies the transmitting tissue which occurs in both the stigma and style. This tissue, in the Cruciferae, differs from that of the Solanaceae in that it constitutes less of the volume of the style, though, structurally the tissues in the two families appear similar under the microscope (Sassen, 1974) consisting of loosely packed cells which can contact to form a middle lamellae but more often are separated by a fibrillar extracellular matrix (Elleman et al., 1988). Analysis of the constituents of the cell walls and ECM of the style reveal similar carbohydrate composition to the Solanaceae with both alcian blue and PAS giving positive results (Hill and Lord, 1987). While there is no detectable

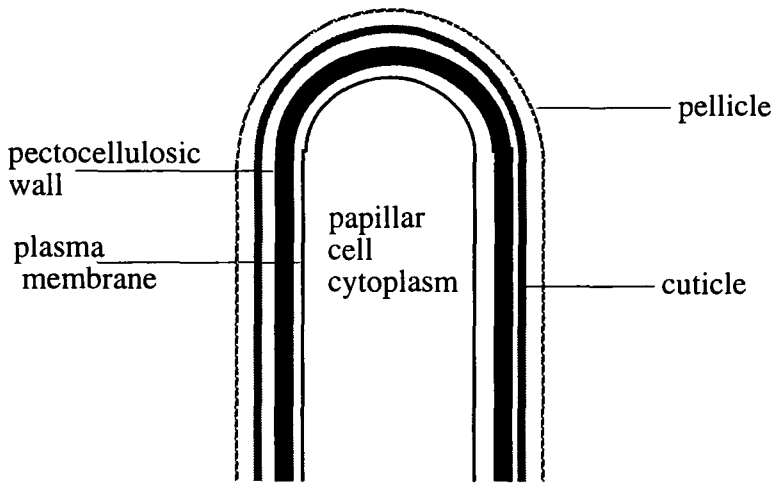


Fig.4. Diagrammatic representation of the layers surrounding a papillar cell in the Cruciferae. Derived from Elleman et al. (1988) and Mattson et al. (1974).

staining for lipid or callose. One difference between the two families is that in the Cruciferae the ECM fails to stain for protein (Hill and Lord, 1987).

1.4.2.2 Pollination in the Crucifers.

Pollination in the Cruciferae differs from that in the Solanaceae predominantly during germination and the early stages of growth. This is as would be expected as the cruciferous stigma surface is covered in well developed papillae which provide a more extensive barrier to penetration than is present in the Solanaceae. The papillae, however have developed such that rather than being a barrier to pollen they are essential for pollen capture and germination. Evidence indicating this key role came from genetic ablation studies where destruction of the papillae in *Arabidopsis* and *Brassica* rendered the stigma sterile (Thorsness, 1993; Kandasamy et al., 1993).

The initial stages of pollination in the Crucifers is pollen grain adhesion to the surface of a papillae cell. The site of initial adhesion must be the pellicle (Mattson et al., 1974) but as this layer is not visible in most of the microscopic reports on the pollination process no experimental description of its role has been reported.

In *A.thaliana* and *B.oleracea* the first visible occurrence after pollen has been brought into contact with the stigma is a spreading out of the exine of the pollen grain to form a foot or 'appressoria' (Ockendon, 1972) that appears to fuse with the cuticle to form a continuous electron opaque layer (Elleman et al., 1990). The exine in this foot has a modified structure as it is not soluble in cyclohexane as the coating in the free pollen is (Elleman et al., 1990). Similar fusion occurs between the exine and cuticle in *R.sativus*, although here the pollen coating is electron opaque prior to pollination (Dickinson and Lewis, 1973a). As the foot forms, the pollen protoplast develops a stratified fibrillar layer and there vesicles accumulate under the plasma membrane and between the plasma membrane and intine (Elleman et al., 1990). By placing the pollen in a humid atmosphere or on wet filter paper the these latter changes in the pollen grain can be artificially induced indicating that the foot is not a structure for taking up water so is probably a means of pollen anchorage.

Uptake of water on the stigma could be mediated by digestion of the cuticle and uptake through that tissue or uptake through one of the gaps in that layer. The fact that contact with the cuticle occurs in many of the microscopical observations (Dickinson and Lewis, 1973a; Elleman et al., 1988; Elleman et al., 1990 and Elleman et al., 1992 all indicate this occurrence) could be used to support the former model but contact with cuticle-free stigma at points of interaction, other than those in the plane of section cannot be ruled out. Confirmation that water uptake is through the cuticle does, however, come from differences in the rate of pollen germination on the carpels of different inbred strains of *B.oleracea* (Elleman et al., 1990). One strain, S63, allows more rapid germination of pollen than a second strain, S25. This was linked to cuticle thickness: the former has a cuticle significantly thinner than that of the latter which was interpreted to provide a less difficult barrier to water uptake. However the limited number of strains used in this study means is insufficient to confirm an a inverse correlation between cuticle thickness and pollen hydration.

The pollen tube starts to grow around the region of the foot in *B.oleracea* and *A.thaliana*. This developmental change in the pollen grain is contiguous with a change in the pectocellulosic cell wall. Layers I and II separate at the point of contact to give a 'blister' (Elleman et al., 1990) which fills with a fibrous matrix (Elleman et al., 1992). In the *B.oleracea* pollen tube the fibrillar layer moves into the tip of the developing tube, which then moves through the foot/cuticle and penetrates into the inter-pectocellulosic space in which it starts to grow (Elleman et al., 1988). In *A.thaliana* growth between these layers occurs with some pollen, while others grow between layer II and the papillar plasma membrane (Elleman et al., 1992). Pollen of *R.sativus* also penetrates the papillar wall but the studies on this species could not determine where in the papillar cell wall growth occurred (Dickinson and Lewis, 1973a).

The precise basis for the response of the papillar cell wall to the binding of compatible pollen and the change in the papillar surface allowing hydration and penetration is unclear. Elleman et al. (1992) suggested that it could be mediated by pollen enzymes. The presence of enzymes in the pollen coat has long been known (Vithanage and Knox, 1976) and these could be acting on the cuticle and pectocellulosic walls. In support of this idea, cutinase activity has been demonstrated and partially purified from *B.napus* while antibodies to a *Fusarium* cutinase, that cross-react with the pollen cutinases the enzyme has also been localised it to the pollen/papillae contact point just prior to pollen tube penetration (Hiscock et al., 1994).

However it is also apparent that in *B.oleracea*, stigma factors must be involved in directing the pollen tube during its initial penetration and growth as abnormal pollinations can be produced if immature carpels are used as pollen receptors. Mature pollen germinates on immature stigmas and often passes right through the pectocellulosic layer so that it grows between the pectocellulosic layer and the plasma membrane (Elleman et al., 1988). Many of these abnormal pollen tubes fail to breach the pectocellulosic layer at the base of the papillae so that they are unable to mediate pollination.

Tube growth occurs between the pectocellulosic layers until the point where the papillar cells contact each other. It is at this point that the cell wall becomes single-layered and this is matched by a change to intercellular growth for the pollen tube. No data on how this change in mode of growth occurs has yet been presented, although it has been reported that tube growth around this region and at the point of leaving the papillae is hard to follow microscopically (Elleman et al., 1988). Once the pollen tube has left the papillar region it enters the intercellular spaces in the transmitting tissue through which it grows towards the ovary (Elleman et al., 1988).

Much of this mode of growth is regulated by the age of the carpel. Pollination of immature carpels often leads to pollen tubes penetrating the entire pectocellulosic layer so that tube growth is often between the pectocellulosic layer

1.4.2.3 The sporophytic SI system in the Crucifers

The Cruciferae, where an SI system is present, have all been reported to have a single locus sporophytic system (Bateman, 1955). The most extensive studies have been on *Brassica* but some work has also been carried out on *Raphanus* which indicates that it has a closely related if not identical, system

(Dickinson and Lewis, 1973a). Genetically the SI system of the Cruciferae is more complex than that of the Solanaceae with over 60 alleles having been identified in *Brassica* (Ockendon, 1974). The behaviour of the pollen is often governed by a hierarchical system of allelic dominance (Thompson and Taylor, 1966) with different alleles having different phenotypic strengths. These alleles have been divided into two classes: class I alleles have a strong incompatibility phenotype (0-10 self pollen tubes germinate per selfed stigma) while class II have a weak incompatibility phenotype (10-30 pollen tubes develop) (Nasrallah et al., 1991). There is also evidence that genetic loci outside the S-locus can affect the strength of the S-phenotype (Nasrallah, 1974; Nasrallah et al., 1992).

Principal differences occur between the biology of SI in the Cruciferae and the Solanacaeae. Firstly, in the former the incompatibility reaction occurs in contact with the stigma surface (Ockendon, 1972). Incompatible pollen may initially fail to adhere to the papillae, adhere but fail to germinate, or germinate but generate abnormal tubes and fail to penetrate the papillae (Dickinson and Lewis, 1973a). The SI response can be overcome by high humidity which suggests that restriction of water to hydrate the pollen is a key factor in rejection (Carter and McNeilly, 1975). This is consistent with the observation that adhering incompatible pollen grains do form an appressoria, consistent with this structures uninvolvement with water uptake.

This site of action leads to the second difference which is the localisation of the SI recognition factors. In *Brassica* and *Raphanus* the carpel and pollen components of the sporophytic SI system have been proposed to lie in the protein pellicle (Mattson et al., 1974) and pollen exine (Dickinson and Lewis, 1973b). The former was identified on the site of pollen arrest, while the latter was identified by the property of its extracts to elicit a an SI response on the stigma surface.

The third principal difference is that unlike the solanaceous SI system, where the pollen is effectively killed, the Cruciferous SI response is reversible. This was illustrated by Roberts and Dickinson (1983) who reported that pollen incubated on a incompatible stigma could be removed and then germinated *in vitro*, with similar kinetics to untreated pollen. This work was supported by experiments on the effect of cycloheximide on pollen/stigma interactions. Treatment of carpels with cycloheximide, 2 hours prior to pollination, effectively blocks the SI response (Sarker et al., 1988). Treatment of stigmas with cycloheximide, post-pollination, led to a similar release of the SI system but only 2 hours after addition of the inhibitor, the lag presumably being the time needed for uptake of the inhibitor. The reversal has been demonstrated for pollen left on an incompatible stigma for upto 24 hours. These results indicated that pollen could undergo an incompatibility reaction yet still germinate *in vivo*.

One interpretation of these reversibility results is that the SI reaction could be restricted to the papillae as opposed to occurring in the pollen as appears to be the case in the Solanaceae. Evidence in support of this came from radiolabelling experiments which indicated that no female proteins appeared to enter the pollen (Roberts et al., 1984b). This contrasts with the Solanaceous SI system where the SI response appears to occur exclusively in the pollen where it is mediated by uptake of the female recognition factor. However the response, if it is restricted to the papillae must be localised at a sub-cellular level as each individual grain, in mixtures of compatible and incompatible pollen, behaves independently from rest. This was illustrated by placing pollen of different compatibility on the same

papillar cell where each grain continued to behave to phenotype and so was uninfluenced by the presence of the other (Sarker et al., 1988).

The results with cycloheximide also indicate that the SI response requires continuous protein synthesis. This is consistent with the high protein turnover in the pellicle and cell walls of the papillae (Roberts et al., 1984b). In the same experiment tunicamycin was shown to have no effect on the SI response but this was due to a failure by the carpel to take up this inhibitor as similar treatment of excised stigmas does inhibit the SI response (Sarker et al., 1988).

In all genotypes in *Brassica* and *A.thaliana* and *R.sativus*, pollen arrest is accompanied by callose deposition around the point of interaction with the papillae (Heslop-Harrison, 1975) which has been proposed to act as a barrier to water and nutrients. Such a deposition is similar to that seen in incompatible crosses in the gametophytic Solanaceae, suggesting that it may be a common link. The callose is also deposited around pollen in compatible pollinations in the Brassicaceae, however, which would appear to rule it out of a direct role in pollen arrest (Hodgkin et al., 1988). An elevation in the level of cellular calcium in incompatible pollinations has also been reported which has been proposed to be part of the intra-cellular signalling component of SI (Singh et al., 1989).

1.4.3 Molecular Characterisation of the stigma/style of the Cruciferae.

Unlike the Solanaceae, in which a number of different classes of expressed sequences have been derived from the stigma and style, work on the Cruciferae has concentrated almost exclusively on the components that mediate the SI response and related sequences.

1.4.3.1 The sporophytic S-genes of the Cruciferae.

Two sequences expressed in the female tissue have been identified from the Brassica S-locus. The first, the SLG (S-linked glycoprotein), was initially identified by one and two-dimensional protein electrophoresis where differently migrating proteins cosegregated with different S-genotypes (Nishio and Hinata, 1977; Roberts et al., 1979; Ferarri et al., 1981; Nasrallah and Nasrallah, 1984).

Cloning of potential S-protein cDNA's was initially performed by differentially screening a stigma cDNA library with stigma and seedling cDNA probes (Nasrallah and Nasrallah, 1985). This strategy was based on the relatively high abundance of the putative S-proteins in the stigma (~5% of total protein) compared to seedling tissue (undetectable) which was interpreted as being due to a relatively high abundance of the cDNA in the former. Therefore cDNA clones which hybridised strongly to the stigma probe but not to the seedling probe were selected as potential candidates. Other workers took a more targeted approach and sequenced the putative S-proteins, then designed oligonucleotide probes with which to isolate the relevant cDNA (Nasrallah et al, 1987; Scutt, 1990). All approaches produced similar sequences. Unlike SLSG, no biochemical activity can be assigned to the SLG on the basis of sequence homology with previously characterised protein sequences although as inhibition of glycosylation eliminates the SI response (Sarker et al. 1988), the carbohydrate moieties may be important for the function of the protein

As well as the freely soluble form of SLG, a cDNA encoding a putative membrane-bound form of the S2 allele has been isolated. Transformation of the S2 gene into tobacco confirmed that both transcripts were derived from the same gene, presumably by differential splicing as both free and membrane-bound proteins could be isolated from this heterologous system (Tantikanja et al., 1993).

The second S-linked sequence was isolated by its homology to the SLG gene. Stein et al. (1991) isolated a cDNA sequence from *B.oleracea*, for which the encoded protein shows three domains. The N-terminal S-domain, similar to SLG, is followed by a putative transmembrane helix, then the C-terminal domain, which shows homology to serine/threonine protein kinase domains. This lead the protein to be named SRK1 (S-related kinase 1). Kinase activity has been confirmed by expression of the kinase domain in *E.coli* (Goring and Rothstein 1992), where autophosphorylation on serine residues in the C-terminal domain occurred.

Both these species were isolated from *B.oleracea* but it is apparent that they also occur in other Crucifers which have SI. Sequences equivalent to SLG and SRK, originating from *B.campestris*, have been isolated and sequenced (Dwyer et al., 1991; Goring et al., 1991; Goring et al., 1992) while cross-hybridising sequences have been identified in the genome of *R.sativus* (Scutt, 1990; Kandasamy et al., 1989).

The biology of these two sequences is both consistent with them being female factors of the sporophytic SI system. RFLP patterns with different SLG and SRK sequences comigrate with their respective S-phenotypes suggesting linkage(Nasrallah et al., 1985; Stein et al. 1991). That both genes are closely linked has been confirmed by pulsed-field gel electrophoresis of *B.oleracea* genomic DNA which indicates that the two genes are approximately 200kbp apart (Boyes and Nasrallah, 1993).

The relationship between the different SLG alleles also suggests that they are the female S-factors. The SLG sequences from class I SI phenotype plants show greater levels of homology to each other than they do to the class II sequences (Nasrallah et al., 1991). This is mirrored by the binding of an antibody MAbH8, originally raised to the class I SLG6 protein which cross-reacts with all the class I proteins but none of the class II (Kandasamy et al., 1989).

The patterns of expression for the two sequences is also consistent with their predicted role as SI recognition factors. The SLG transcript is strongly expressed in the papillar cells as indicated by *in situ* hybridisation (Nasrallah et al. 1988) and reporter gene analysis (Sato et al. 1991), while immunolocalisation indicates that the SLG protein is secreted into the papillar cell walls (Kandasamy et al. 1989). This pattern of expression is consistent with the site of action of self-pollen rejection. The expression of SLG sequences during development also mirrors the onset of self-incompatibility (Nasrallah et al. 1985). Similarly, reporter gene experiments suggest that SRK is also expressed in both the carpel and anther (Nasrallah and Nasrallah, 1993), although exact localisation of the protein product has not been reported.

A third strand of evidence for the involvement of these sequences in SI comes from the study of self-compatible mutants. Two self compatible strains of *B.oleracea* and *B.napus* which were isolated from normally self-incompatible lines were both found to contain mutations in their SRK genes. In the

Unlike the solanaceous S-genes, however, absolute conformation of the role of these genes has not been produced as transformation of a *B.oleracea* and *B.campestris* S-gene into self-compatible *B.napus* does not confer the relevant S-phenotype (Nishio et al., 1992). This is not unexpected, given the presence of a second female S-gene (SRK), so the introduction of both SLG and SRK genes may be necessary to confer an S-phenotype, an experiment that has not yet been reported. The choice of *B.napus* cv Westar as a host for this transformation experiment is also open to criticism as this plant is naturally self-compatible which could be due to mutations in genes that are essential for the SI response but lie outside the S-locus, so that even if SLG/SRK could confer an S-phenotype it may be masked by deficiencies in other genetic loci. Therefore antisense suppression of the expression of existing S-genes may be the best approach to confirming their role in SI as performed by Lee et al. (1994) for the Solanaceae

1.4.3.2 Relationship between the sporophytic S-genes and the SI response

Some work has been carried out on possible signal transduction systems in SI, inspired by the presence of a kinase function in SRK. Treatment of mature (1-2 days post-anthesis) carpels of *B.oleracea*, *B.napus* and *B.campestris* with okadaic acid or microcystin-LR (inhibitors of class I and class IIA protein phosphatases) destroys their ability to support pollen germination (Rundle et al., 1993). Coupled with a lack of any effect on the SI system in carpels of this age this was interpreted to indicate that protein dephosphorylation was essential for pollen growth. This work was complicated, however by the effects of the inhibitors on younger (0-2 days pre-anthesis) carpels in which cross-pollination was unaffected but self-pollination was blocked. The fact that this was only observed in one genotype of plant (S22) lead the authors to conclude that this effect , while suggesting a different mechanism for promoting pollen tube growth in pre- and post anthesis carpels, was predominantly a result of the genetic background of S22 plants, possibly an inbreeding effect.

Meanwhile, Rundle and Nasrallah (1992) have isolated a class I protein phosphatase (PPI) cDNA from a *B.oleracea* stigma library via PCR using primers to conserved regions of the genes of other members of this protein family. They also indicated the presence of sequences in the *Brassica* genome that cross-hybridised with a rat PPI-inhibitor cDNA. Neither of these results, however, illuminate the processes occurring in the stigma, as even though the phosphatase was isolated from a stigma library its precise expression was not reported. Such data is desirable as, if the clone encodes a protein with a possible antagonistic role to SRK, then it must be expressed in the papillae. The inhibitor work is even less relevant as a definite sequence for this putative gene was not obtained leaving the doubt that the observed hybridisation may be to a different sequence altogether.

More evidence for the role played by SLG and SRK in SI, as with SLSG, might come from isolation of the male component and characterisation of its interaction with the two female components. However, as with the solanaceous SI system a convincing male recognition factor has yet to be reported. Three possible candidates have been proposed but none has been conclusively shown to be the male component. The candidates are:

(i) SLG. SLG-like transcripts in anthers have been detected by PCR (Heizmann et al., 1991) and shown to hybridise to a 4000nt. transcript from microspores. Together with reporter gene experiments

(Sato et al., 1991) in which both weak gametophytic and weak sporophytic expression in the anther was reported, this has lead to the hypothesis that the Lewis tripartite model of similar recognition components in both male and female tissues is essentially correct (Heizmann et al., 1991). Evidence for such a model comes from experiments where application of purified S2 glycoprotein to non-S2 pollen conferred an S2 phenotype to that pollen (Ferrari et al., 1981). However, this hypothesis is still not universally accepted due to the lack of any model detailing how like SLG components recognise each other.

(ii)SLG-binding protein. A polypeptide (pcp7) has been identified and isolated via its interaction with purified SLG (Doughty et al., 1993). This polypeptide induces a pI shift for SLG when the latter is pure or in crude stigmatic extracts, although the interaction differs in that in the former a series of new bands are produced on IEF gels whereas one discrete new band occurs with the latter. This suggests that other female components are required for a specific reaction. Purification and partial sequencing has been carried out on pcp7 (Hiscock et al., 1995), which has lead to the cloning of relevant genes (Stanchev et al., 1996). These sequences, however, do not appear to be S-linked so that it is unlikely that pcp7 is the male component of SI (Stanchev et al., 1996).

(iii) S-locus linked gene. A gene has been isolated from *B.oleracea* based on the co-segregation of a restriction fragment of the respective genomic DNA with the S-haplotype. Two complementary mRNA's are transcribed, at a low level, from the region, one of which is spliced and has the capacity to code for protein, whereas the second does not. Expression of the two transcripts during microspore development is antagonistic, a factor which has lead to the theory that the complementary expression has a regulatory function via antisense action.

While they are clearly S-linked, the genes encoded by this transcript may not encode be the pollen factor. The spliced, coding transcript has not been detected in tapetal tissue, which is the predicted site of synthesis of the male component, although the low level of expression may be contributing to this result. Nor does the predicted protein product contain any recognised signal sequence to direct export from the site of expression in the meiocytes to the pollen wall. Therefore, current evidence points to this S-gene not encoding a pollen coat protein, which is contradictory to the physiological localisation of the male S-factor (Dickinson and Lewis, 1973b).

1.4.3.3 Other expressed sequences from the Crucifer stigma and style.

Although the majority of work on the molecular basis of pollination in the Cruciferae has been targeted on the SI system, other expressed sequences have been isolated.

1.4.3.4 S-related genes.

A number of sequences similar to the SLG and SRK, but which do not appear to play a role in SI, have been isolated from the carpels of various Cruciferae. Three sequences, SLR1(SLR=S-locus-related, Lalonde et al. 1989), SLR2 (Boyes et al. 1991)and Ats1 (Dwyer et al., 1992) have been isolated by their similarity to the previously characterised S-genes. The predicted protein sequences of these three genes show 65%-90% homology to the SLG sequences. The protein encoded by SLR1, which was isolated from *B.oleracea* is localised in the papillar cell walls (Umbach et al., 1990) while the promoter of this gene

shows similar activity to that of the SLG (Hackett et al., 1992). Ats1, which from its sequence, is believed to be the *A.thaliana* equivalent of SLR1 has also be localised, via a promoter-GUS reporter construct, to the stigma (Dwyer et al., 1994)

None of these sequences are believed to involved directly in pollen recognition. Both SLR1 and SLR2 show extensive sequence homology but no electrophoretic variation between S-haplotypes and different *Brassica* species (Lalonde et al., 1989; Boyes et al., 1991; Umbach et al., 1990), while there is no co-segregation of their RFLP patterns with those of the SLG alleles (Dzelzkalns et al., 1992). The fact that Ats1 comes from a self-compatible species has been taken to infer a lack of involvement in SI (Dwyer et al., 1992).

The presence of sequences homologous to the S-genes in other organs and plant species is also worth noting. Dwyer et al. (1994) have isolated an Arabidopsis SRK homologue which also displays floral specific expression. Three vegetativly expressed SRK homologues have also been isolated from Arabidopsis (Tobias et al., 1992 and Walker, 1993), while an SLR sequence has been cloned from a carrot embryo cDNA library (van Engelen et al., 1993). Homologous sequences have also been found in monocot species. Three SLR (Zhang and Walker, 1993) and one SRK-like (Zhang and Walker, 1990) sequences have been isolated from maize. The SRK homologue is found in roots and silks, while at least one of the SLR sequences is also present in silks.

The presence of active S-related genes in vegetative and non-SI reproductive tissues suggests that they play roles in a variety of plant process. Such a role, if conserved may provide suitable basis for the elucidation of the role of SLG in pollen recognition. The pivotal role played by receptors in animal systems (Fantl et al., 1993) suggests that the SRK homologues, at least, may play a key part in plant function. However, no clear role for any of the isolated sequences has yet to be determined.

One proposal is that the reproductive tissue SLR sequences may play a role in compatible pollination (Nasrallah and Nasrallah, 1993). This hypothesis would explain the presence of SLR genes in non-SI *Arabidopsis* reproductive tissue.

1.5 Expressed sequences from other species.

Although the molecular analysis of the stigma and style has been dominated by the Solanaceae and Cruciferae work has been carried out on other species, although the range of expressed sequences from these is not extensive.

1.5.1 The Papaver SI genes.

Outside of the two main plant families, the most extensive molecular work has focused on the SI system of *Papaver rhoeas*. The process of pollen recognition in this plant differs from the Solanaceae and Cruciferae in that a single-allele gametophytic SI system is present despite the fact that the stigma is dry (Lawrence, 1975). As with the sporophytic SI system pollen arrest occurs at the stigma surface, although

callose deposition is inside the pollen grain as is seen in the solanaceous SI response. As such *P.rhoeas* offers a third class of SI system for investigation.

Investigations into the physiology of the *Papaver* SI response have been assisted by the development of an in vitro pollen growth bioassay system (Franklin-Tong et al., 1988). Addition of incompatible stigmatic extracts to the growth media can inhibit upto 80% of pollen developing indicating that only stigmatic components rather than an intact tissue is required for inhibition (Franklin-Tong et al., 1988). This assay was also used to demonstrate that only extracts of mature stigmas contain the active SI agent(s) (Franklin-Tong et al., 1989).

The in vitro system was also used to investigate the effect of metabolic inhibitors on compatible and incompatible pollen growth (Franklin-Tong et al., 1990). Actinomycin D, an inhibitor of transcription, blocked both compatible and incompatible pollen growth, while cyclohexamide, an inhibitor of translation, and tunicamycin, an inhibitor of protein glycosylation, could both block the SI activity of stigmatic extract implying that:

i) New protein synthesis but not new RNA synthesis is required for compatible pollen growth. This is consistent with many hypotheses that claim that proteins required for pollen growth and germination are transcribed from mRNA's deposited in the pollen during development in the anther.

ii) New RNA synthesis is required in the pollen for an SI response suggesting that new pollen genes are activated in incompatible pollinations.

iii) Some, or all, of these new pollen genes encode proteins which require glycosylation.

Further evidence for 2) and 3) came from *in vitro* transcription of mRNA's isolated from pollen in compatible and incompatible reactions. About sixteen new proteins, not seen in the compatible mixture, could be detected by SDS-PAGE of the products from the incompatible mix mRNA (Franklin-Tong et al., 1990; Franklin-Tong and Franklin, 1992).

Initial attempts to determine the female component used the *in vitro* pollen growth system as a bioassay to purify an SI-active protein (Franklin-Tong et al., 1989, Franklin-Tong and Franklin, 1992). As with the *Brassica* and *Nicotiana* S-proteins the *Papaver* protein was shown to be glycosylated by concanavalin A binding (Franklin-Tong et al., 1989). Unlike the Solanaceae no RNase activity was associated with the purified protein (Franklin-Tong et al., 1991), which correlates with the relatively low RNase activity found in the carpel of *P.rhoeas* compared to *N.alata* (Franklin-Tong et al., 1991). The protein also differs from those of the other SI systems by its size (~22kb, Franklin-Tong et al., 1989) and its abundance: the *Papaver* S-protein constitutes <0.1% of total stigma protein compared to 10% for SLG.

Sequence for the *Papaver* female S-component came from isolation of the an S1-genotype linked protein by SDS-PAGE (Foote et al., 1994). Using an oligonucleotide derived from the N-terminal protein sequence cDNA clones were isolated which could code for a glycoprotein with no homology to SLG or S-

RNase. Conformation that this was the female SI-component came from three studies: The cDNAs come from a single copy gene which cosegregates with the S-locus, transcripts homologous to the cDNA are only expressed in mature carpels and expression of the protein in *E.coli* produces a recombinant protein that can elicit the SI response in incompatible pollen in vitro. This last result also indicates that post-translational glycosylation is not required for activity.

As with the Brassica S-genes the sequence of the *Papaver* S-gene provides no information on how recognition of incompatible pollen is mediated. Some form of intracellular signalling has been proposed due to the activation of pollen genes (Franklin-Tong et al., 1990). Evidence for a role for calcium and protein phosphorylation, both of which are found in many signal transduction cascades in plants (Roberts and Harmon, 1992; Ranjeva and Boudet, 1987), has been provided by the *in vitro* bioassay:

Injection of pollen tubes with calcium green followed, by addition of incompatible stigma extracts, reveals a pulse of calcium around the nuclei suggesting that calcium may act in gene activation (Franklin-Tong et al., 1993). Injection of caged calcium and subsequent photoactivation induced pollen arrest though expression of the SI-responsive genes was not assayed (Franklin-Tong et al., 1993). Addition of recombinant S-protein also elicits an identical response to that seen with stigmatic extracts (Franklin-Tong et al., 1995) confirming that the S-protein alone elicits this calcium response.

Rapid changes in phosphorylation of pollen proteins together with changes in phosphatase activity have also been detected in *in vitro* SI reactions (Franklin-Tong and Franklin, 1992). Investigations into the relationship between these events and the calcium pulse (i.e. does artificially generating the latter induce the former), however have not been reported, so the linear progression of the signal cascade is still unknown.

In summary, while extensive isolation of different alleles from different species, biochemical characterisation of the recognition component or isolation of related gene sequences, have not been carried out as they have for the Solanaceae and Cruciferae, significant advances have been made on understanding the molecular basis of the *Papaver* SI system. the development of an in vitro bioassay has greatly enhanced this and makes the *Papaver* system, perhaps, the most attractive of the three model systems in studying the molecular basis of SI.

1.5.2 Cell wall proteins.

As well as those identified in solanaceous plants, putative cell wall proteins have been isolated in the stigmas and styles of other plants. The most extensively characterised of these is a cDNA encoding a proline-rich protein from *Antirrhinum* spp. that was isolated via differential screening of a carpel cDNA library (Baldwin et al., 1992). The predicted protein encoded by this cDNA lacks the classical SP4 extensin motif but does contain XP3 sequences which coupled with its high proline and serine content lead the authors to describe it as an extensin. The cDNA is restricted in its expression to the transmitting tissue where it occurs in a central set of cells being absent from the cell layers that from the transmitting tract/parenchymal boundary.

AGP's have also been identified in the stigmas and styles of a number of plants but not in great detail. The adhesive properties of these proteins, coupled with the availability of the specific Yariv reagent initially attracted workers who considered them potential candidates for mediating interaction between the pollen and the female tissue. AGP's have been detected in a variety of plant stigmas (Hoggart and Clarke, 1984). Detailed analysis of the proteins has been reported for *Gladiolus* (Gleeson and Clarke, 1980) and *Lilium longiflorum* (Aspinall and Rosell, 1978) which indicates that the proteins in the different species appear similar on the gross level.

In the latter species a specific AGP epitope was identified that was restricted, in the carpel, to the transmitting tissue (Jauh and Lord, 1996) The proteins carrying this epitope have been predicted to mediate interactions in this tissue, whether between cell wall glucans within the tissue or between the transmitting tract and the growing pollen tube.

1.5.3 Vitronectin-like sequences.

After Sanders and Lord (1989) demonstrated that directed movement of inert particles was possible in the style work has been carried out to identify the proteins involved with this process. The search for components of a cell movement system has taken as its starting point the similarity with animal systems in which cells are moved around, where key components are the fibronectins and vitronectins (Underwood and Bennet, 1989). Anti-human vitronectin antibodies were used by Sanders et al. (1991) to identify cross-reacting proteins that were present in *Lilium*, *Vicia* and *Lycopersicon*. Reactive proteins were found in a variety of tissues including roots, leaves and styles. Putative genes for these proteins were identified by low-stringency screening using human cDNA sequences in the genomes of *Lilium* (Wang et al., 1994).

Purification of the putative vitronectin from salt-stressed tobacco culture cells allowed antibodies to be raised (Zhu et al., 1994). These cross-reacted with proteins in a several tissues including the style where they are localised to the inner wall of the transmitting tract cells. A cDNA clone was isolated using these antibodies via expression screening. Sequencing of this clone revealed that it has no sequence homology to vitronectin but encodes a member of the EF1α superfamily of proteins (Zhu et al., 1994). This does not preclude the protein from being involved in intercellular interactions in the plant as members of the EF1α family have been implicated in actin binding and cell adhesion in animals (Yang et al., 1990).

These results, while isolating a putative adhesion protein have diverged from their original purpose which was the study of the stigma and style. The identified cDNA's and proteins are expressed in multiple tissues, not just the style, which would appear to preclude a style specific role for them. The presence of such sequences in vegetative tissue is also contradictory to their proposed role in cell movement which, as previously stated, is unique to the pollen tube in higher plants.

2
AIMS AND OBJECTIVES

2.1 Motives for studying the stigma and style of the Cruciferae

As illustrated in the introduction, all the work involved in the isolation of genes and proteins from the stigma and style of cruciferous plants has concentrated on the analysis of the SI system and its related sequences. While concentrating on one aspect of the pollination process mediated by the carpel has allowed significant progress to be made in isolating the components of this system it has lead to the neglect of other aspects of these tissues. The study of other stigma/style components is desirable as a complete description of the SI system in the Cruciferae will require characterisation of how the SI components block pollen germination and growth, which in turn will require an understanding of the molecular basis of the interaction between the stigma/style and compatible pollen. Given the anatomical and physiological differences between the cruciferous and solanaceous stigma and style it cannot be assumed that similar molecular processes underpin pollination in the two families. Therefore some analysis of the molecular genetic composition of the stigma and style in the Cruciferae is desirable.

One problem with working with the Cruciferae is that their floral organs are significantly smaller than some of those in the Solanaceae so that methods that require only small amounts of tissue, or can utilise tissue that has been stored frozen are the best approach to studying these species. The isolation of mRNA sequences, via cDNA libraries is one such approach. As the female tissues are estimated to express 10000 genes not found in the vegetative organs (Kamalay and Goldberg, 1980) the expression of novel mRNA's in the stigma and style would be predicted to form the molecular basis of its physiology, compared to the rest of the plant. Experimental support for this approach comes from the work on the Solanaceae, in which many of the stigma/style specific sequences have been identified by the isolation of cDNA's with tissue-specific expression.

2.2 The *Brassica* clones isolated in previous studies.

The isolation of cDNA's expressed in the stigma and style of *B.oleracea* var. was performed by Scutt (1990) as a sideline to his work involved in isolating an SLG gene sequence from this plant . These clones were identified by differentially screening a stigma cDNA library with probes derived from stigma and leaf mRNA's. Four different classes of clone were identified from a preliminary screen of 2000 clones. One of these clones was an SLR sequence and has since been extensively studied by other workers (Nasrallah and Nasrallah, 1993). A second class of clone, 11H4, was not available for further study as both representatives had been lost.

The third clone, 13G6(i), whose insert, estimated to be ~1100bp. long, was one of two cDNA inserts in the same clone. The second insert was found not to be differentially expressed and so was discarded. Partial sequencing of 13G6(i), after purification from the second clone, indicated that it represented a novel sequence, while preliminary Northern analysis indicated that ~1300nt. long transcripts homologous to the cDNA sequence were abundantly expressed in the stigma, less abundant in the style and immature anthers and absent from leaf tissue.

The fourth clone, 15H11, was ~800bp long. The entire clone was sequenced by Scutt who proposed that it coded for a cell wall glycine-rich protein (GRP) as over 90% of its putative protein product consisted of the motif GGGX, a motif also seen in two cell wall proteins from *Phaseolus* (Keller

et al., 1988). Transcripts homologous to the clone were detected in stigmatic tissue but not in style, immature anther or leaf. The stigmatic transcript was estimated to be ~1300nt. in length indicating that 15H11 represents only a partial sequence.

2.3 Plan for PhD project.

The initial aim of this project was to complete sequencing of the 13G6(i) insert and to perform more extensive studies on its expression, so as to determine if it might play a stigma/style specific role. As the clone was available more extensive analysis of the expression of 15H11, which, from Scutt's work, appeared to be stigma-specific was also to be carried out. As two clones represent only a small fraction of the stigma/style-specific sequences it was decided that a new cDNA library should be produced so as to isolate different expressed sequences

Whereas previous work was carried out on *B.oleracea* var. *gemifera*, stocks of this plant were not available, therefore the closely related *B.oleracea* var. *alboglabra* was used. As these two organisms are members of the same species, the probability of significant molecular differences between them was considered to be low

3.
MATERIALS AND METHODS

3.1 Materials.

3.1.1 Plant Material.

Brassica oleracea var. *alboglabra* was grown from laboratory seed stock maintained from original seeds from Dr D.Ockendon (Plant Breeding Institute, Wellesbourne, Warwickshire). *Brassica napus*, *Nicotiana tabacum*, *Hordeum vulgare* and *Triticum aestivum* plant material was provided by C.Illet of the Department of Biological Sciences, University of Durham. *Pisum sativum* and *Oryza sativa* tissue was provided by Dr R. Bernardi and Dr K.Powell, respectivly, both of the University of Durham, Department of Biological sciences.

3.1.2 Bacteria and Phages.

E.coli strains *XL1-Blue* and *DH5α* (both Sambrook et al., 1989)were used for general cloning and were taken from stocks maintained in the laboratory of Dr R. Croy. Strain *LE392*, (Sambrook et al., 1989)used to propagate the *B.oleracea* var *alboglabra* genomic library, also came from a laboratory stock. *E.coli* strains *NM522*, used to propagate the cDNA library, and *E.coli NP66*, used for *in vivo* phagemid rescue were purchased from Pharmacia Biotech Ltd.

The *B. oleracea* var *alboglabra* S29 library in Lambda FixII was provided by Dr. M.Trick of the John Innes Institute, Norwich.

3.1.3 Chemicals.

Unless otherwise stated, chemicals were purchased from Sigma Aldrich Chemicals Ltd., .Poole and BDH Ltd., Newton Aycliffe All chemicals were of "Analar' grade, or the best grade available.

Bactopeptone and Yeast extract were purchased from Oxoid Ltd, Basingstoke. Bacteriological agar was purchased from Difco, Detroit, Michigan, USA.

Agarose was purchased from Boehringer Mannheim. Low-melting point agarose was purchased from BRL, Gaithesburg, USA.

Ficoll 400 was purchased from Pharmacia Biotech Ltd.

Silica fines were provided by Dr Nigel Robinson of the University of Durham, Department of Biological Sciences.

α^{32}P-dCTP was purchased from Amersham Life Science plc., Amersham.

Nitocellulose membranes were purchased from Schleicher and Schuell, Dassel, Germany. Hybond-N nylon membranes were purchased from Amersham Life Science plc. GFC filters and 3MM chromatography wer purchased from Whatman ltd, Maidstone.

UHQ water (18MΩ purity) was purified using an Elgastat UHQ PS filtration unit.

Sephacryl S300 spin columns were provided by Pharmacia biotech Ltd.

3.1.4 Enzymes

Restriction endonucleases were purchased from Boehringer Mannheim Ltd., Lewes; MBI Fermentas Ltd., Gateshead or Northumbria Biologicals Ltd., Cramlington.

T_4 DNA ligase, RNaseA, RNaseH, and T_4 DNA polymerase were purchased from Boehringer Mannheim ltd. Klenow fragment was purchased from both Boehringer Mannheim ltd and MBI Fermentas ltd. T_4 polynucleotide kinase was purchased from Promega.

MuMLV reverse transcriptase was purchased in glass stabilised form from Pharmacia Biotech Ltd.

Biotinylated Concanalvin A (ConA), generated by the method of Bayer and Wilchek (1990), and Extravidin were provided by Dr D.Robertson of the University of Durham, Department of Biological Sciences.

3.1.5 Nucleic Acids.

pKS+Bluescript was provided from a laboratory stock by Dr. R.Croy, Department of Biological Sciences, University of Durham. The clones 13G6(i) and 15H11 (in pUC18) and pHA1 (in pACYA184, Cheung and Cohen, 1975) were also provided from laboratory stocks as above. λ ExCell DNA was purchased from Pharmacia Biotech. Ltd. EcoR1 adaptors were purchased from Promega Ltd, Southampton.

Arabidopsis thaliana genomic DNA was provided by Dr A.Fordham-Skelton, Department of Biological Sciences, University of Durham. *Zea mays* and *Homo sapiens* genomic DNA were provided by Dr.T.Daniell, Department of Biological Sciences, University of Durham.

3.2 Plant Growth and Treatment

3.2.1 Plant Growth.

Plants were initially grown in a departmental greenhouse without temperature or lighting control. RNA for the construction of the cDNA library and for the Northern blots for the 13G6(i) clone was isolated from tissue from these plants. Other RNA was isolated from plants grown in a controlled environment growth room in the Department of Biological Sciences, University of Durham. Plants were grown at 20°C on a 16 hour light/8 hour dark cycle. Tissue for protein analysis also came from these plants.

B.napus and *N..tabacum* was grown in controlled growth rooms under identical conditions to the *B.oleracea*. *P.sativum* was grown under the greenhouse conditions described above. *H.vulgare* and *T.aestivum* were grown in a 15°C growth room under continuous illumination.

3.2.2 Stress treatment of plants.

Leaf wounding was carried out on mature flowering plants by cutting the leaves at 1cm intervals at right angles to the midrib with a pair of scissors. Stem wounding was carried out by making 3mm deep incisions at 1cm deep intervals into the stem using a razor blade. Carpels were wounded by slicing the organ in two along its longest axis using a razor blade. Wounded plants were maintained at 20°C on a 16 hour light/8 hour dark cycle day for the times indicted in chapter 4.

Salicylic acid was administered by cutting leaves from the stems of mature plants at their node and placing them, cut-end first, into a solution of 5mM salicylic acid in distilled water These leaves were

incubated at 20°C on the light cycle described above for the time periods described in chapter 3. Control leaves were left in distilled water alone.

All other stress treatments were carried out on four week old seedlings:

Dark-treated plants were covered with a black PVC bin bag left loose around the base of the plant pot to allow free circulation of air. These plants were maintained at 20°C.

Cold treated plants were placed in a room maintained at 4°C under continuous illumination.

Dessicated plants were maintained at 20°C on the light regime described above. Plant pots were left on a dry tray and no water was added to the plants from day zero

3.3 General Molecular Biology

3.3.1 General Molecular Techniques

Phenol/chloroform/IAA extraction, ethanol precipitation and isopropanol precipitation were performed as described in Sambrook et al. (1989).

Spin column purification was performed using Pharmacia sephacryl S300 columns as described by the manufacturer.

3.3.2 Growth and Storage of *E.coli.*

Liquid cultures of *E.coli* were grown in Luria Broth (LB: 10g/l trypticase peptone, 10g/l NaCl, 5g/l bacto yeast) supplemented with suitable antibiotics at 37°C. Plate stocks of bacteria (except *NM522*) were grown on LB containing 15g/l agar at 37°C and stored at 4°C. *E.coli* strain *NM522* were grown on M9 salt plates containing agar as above (Sambrook et al., 1989). For long term storage bacteria were kept in LB supplemented with 30%(v/v) glycerol at -80°C.

3.3.3 Transformation of *E.coli.*

Frozen competent *E.coli* cells were prepared and transformed by the method of Hanahan (1985): Cells were grown to mid log phase in LB then chilled for 15min in an ice/water bath. The cells were pelleted (400g, 10 min, 4°C) and resuspended in 1/3 volume of ice cold FB (100mM KCl, 50mM $CaCl_2$, 10mM CH_3COOK, 10%(v/v) glycerol pH6.7). After a 15min incubation in ice water the cells were repelleted as before and resuspended in FB (1/8 of the original volume). Cells were flash frozen in liquid nitrogen and stored at -80°C.

Cells were transformed by thawing on ice and mixing in a 10 fold excess with the relevent DNA. The mixture was incubated on ice for 30min then heat shocked at 42°C for 2min. A four fold excess of LB was added and the cells grown at 37°C for 30-60min to induce expression of antibiotic resistance genes prior to plating onto LB containing suitable antibiotics to allow selection of transformants.

3.3.4 Preparation of Plasmid DNA from *E.coli.*

Plasmid DNA was isolated from *E.coli* by the alkali lysis method as described by Sambrook et al. (1989). For sequencing DNA was further purified using PEG8000 precipitation (Sambrook et al., 1989)

or by preparation using the Promega Magic Miniprep kit (J. Bartley, personal communication): 10ml of bacterial culture were pelleted (900g, 10min.) and lysed by the 'Magic minipreps' alkali lysis procedue using 300µl of each solution. 900µl of the lysate was mixed with 1000µl of 'Magic Miniprep' resin and incubated for 5 min. at room temperature. This mixture was then pushed through a 'Magic minipreps' column, washed with 3ml of the manufacturers' wash solution and then spun at top speed in a microfuge for 1min to remove residual wash buffer. The DNA was eluted by the addition of 100µl of MQ water to the column and spinning for a further 1min at top speed in a microfuge.

3.3.5 Manipulation of lambda phage.

λ phage were routinely maintained in SM media (50mM TrisHCl (pH7.2), 100mM NaCl, 10mM $MgSO_4$, 2%(w/v) gelatin) supplemented with 2%(v/v) chloroform at 4°C. For long term storage phage were stored in SM supplemented with 7% DMSO at -80°C.

Phage were grown in top (7g/l in LB) agar with host cells, at 37°C, as described in Sambrook et al. (1989). In all cases LB agar plates and top agar were used except where DNA was to be isolated when agarose was used as the gelling agent at the same concentration as agar.

To allow multiple phage stocks to be grown on one plate a spotting method was used. Phage were spotted by dipping the end of a 0.5mm diameter capiliary tube into the phage solution and touching this end onto the surface of a plate coated in top agar containing the host *E.coli.* These plates were then grown as described above.

λExCell phage DNA was converted to phagemid as described by the the supplier.

3.3.6 Preparation of lambda phage DNA

λ phage, derived from purified clones, were grown, to near confluency, on 90mm diameter petri dishes overnight. The phage were harvested by the addition of 2-3ml of SM per plate and agitation for 2-6 hours. The SM was then collected from each plate and cell debris removed by centrifugation (10000g 5min). RNaseA (30ug/ml) and DNase1 (20ug/ml) were added to the supernatant and bacterial nucleic acid digested at 37°C for 30min. The phage were then precipitated by the addition of 1/3 volume of 20%(w/v) PEG6000, 2M NaCl and incubation on ice for one hour. The phage were pelleted (14 000rpm, 10min) and resuspended in 0.5ml of SM. Protein and PEG were removed by sequential extractions with chloroform/IAA, phenol/chloroform/IAA and chloroform/ IAA. The λ DNA was precipitated by the addition of 1/10 volume of 3M NaOAc pH5.2 and 2 volumes of ethanol. DNA was spun down (13000rpm 15min 4°C), washed with 70% ethanol and resuspended in UHQ water or TE buffer.

3.3.7 PCR from lambda phage.

In order to obtain a rapid assessment of insert size, PCR was performed directly on unpurified lambda phage eluates using primers based on the T7 (TAATACGACTCACTATAGG) and SP6 (GATTTAGGTGACACTATAG) primer binding sites in lambda ExCell (Clackson et al., 1991). Phage eluent (2µl) was mixed with 10pmol of each primer, 2.0µl of Promega 'Thermo' PCR buffer, 1µl 25mM

Mg^{2+} and 0.2U of Taq polymerase in a total volume of 20µl. The reaction was overlayed with 100-200µl of wax to prevent evaporation then cycled through the following routine:

90s, 94°C

90s, 60°C

90s, 72°C.

Thermal cycling was performed on a Perkin-Elmer DNA Thermocycler for thirty repetitions. The PCR mixes were then cooled to 4°C before removal from the PCR machine and subsequent analysis.

3.3.8 Quantitation of nucleic acids.

DNA/RNA concentrations were routinely determined using a Beckmann DU7500 spectrophotometer according to the manufacturers instructions. The concentration and purity of nucleic acids was determined from the optical density at 260nm and 280nm (Sambrook et al. 1989) For gene quantitation the concentration of DNA was determined by the DABA method of Thomas and Farquhar (1978).

3.3.9 Enzymic treatment of DNA.

Restriction digest, ligation and phosphorylation were carried out using the conditions recommended by the manufacturers of the relevant enzymes. Where double digestions were carried out a buffer was chosen to maximise the activity of both enzymes. Recessed 5' termini were filled using Klenow fragment with the buffer conditions recommended by Sambrook et al. (1989). Protuding 3' termini were blunt ended by incubation with 1-2 units of T_4 DNA polymerase at 12°C using the buffer originally used for the enzyme digestion supplemented with 5mM dNTPs.

3.3.10 Sequencing of DNA

Double stranded plasmid DNA, purified by Magic miniprep or PEG precipitation (section 2.3.4) was sequenced by Dr.J.Bartley, Department of Biological Sciences, University of Durham. Sequencing reactions were carried out using Taq dye primer sequencing kits (Applied Biosystems part 901482 rev.B). The reaction products were analysed on an Applied Biosystems 373A DNA sequencer.

3.4 Isolation of Nucleic Acids from Plants

3.4.1 DNA extraction from plants.

DNA was extracted from plants using a modification of the method of Murray and Thompson (1980): Leaf tissue was ground to a fine powder in liquid nitrogen and added to 2 volumes of extraction buffer (2%(w/v) CTAB, 100mM TrisHCl pH8.0, 20mM EDTA, 1.4M NaCl) preheated to 65°C. The extraction mix was heated at 65°C for 15min then spun (900g, 10min) to remove insoluble material. The supernatant was extracted twice with chloroform/IAA (24:1) and a DNA/RNA fraction precipitated by the addition of an equal volume of 'low-salt' buffer (1%(w/v) CTAB, 50mM TrisHCl (pH8.0), 10mM EDTA)

and incubation at room temperature for 30min. The nucleic acid was pelleted (9000g, 10min) and resuspended in CTE (1M CsCl, 20mM TrisHCl(pH8.0), 1mM EDTA) containing 60ug/ml of RNaseA. Genomic DNA was precipitated by the addition of an equal volume of ice cold 13%(w/v) PEG8000 and pelleted (12000rpm, 10min). The DNA pellet was washed twice with 70%(v/v) ethanol, dried and resuspended in TE buffer.

3.4.2 RNA extraction from plants.

Tissue for RNA extraction was harvested in liquid nitrogen and stored at -80°C until required. RNA was extracted from plant tissue by a modification of the method of Logemann et al. (1986): Frozen tissue was ground to a fine powder in a pestle and mortar and covered in 2 volumes of extraction buffer (8M guanidium hydrochloride, 50mM MES, 50mM EDTA 0.015% ß-mercaptoethanol pH8.0). Once thawed this solution was extracted twice with phenol/chloroform/IAA (25:24:1) and RNA precipitated by the addition of 0.2 volumes of 1M acetic acid and 0.6 volumes of ethanol then incubated at -80°C for 30-60min. The precipitate was pelleted (13000g, 15min, 4°C) and washed with first 3M NaOAc pH5.2 then 70%(v/v) ethanol. After drying the pellet was resuspended in DEPC-treated UHQ water. RNA was stored at -80°C.

Where required, further purification of RNA was achieved by the addition of an equal volume of 4M LiCl to the RNA in water and incubated at -20°C for 60min (Walterscheid and Milburn, 1993). The precipitated RNA was pelleted (13000g, 30min, 4°C), washed with 70%(v/v) ethanol and resuspended in DEPC-treated UHQ water.

3.4.3 Isolation of poly (A) RNA.

Poly(A) RNA was isolated from total RNA using the Promega 'PolyAttract' kit by binding poly(A) transcripts to oligo dT linked to parmagnetic particles, as described by the manufacturer. RNA for library construction was prepared by two serial extractions. For probe generation only one extraction was carried out. After isolation the RNA was ethanol precipitated and resuspended in UHQ water. poly(A) RNA was stored at -80°C.

3.5 Gel Electrophoresis

3.5.1 Electrophoresis of DNA.

DNA was electrophoresed through 0.6-1.0%(w/v) agarose 1xTAE (40mMTris-acetate, 1mM EDTA pH 8.0) gels as described by Sambrook et al. (1989). DNA was stained during electrophoresis by the addition of 0.1ug/ml ethidium bromide to the gel. The DNA was visualised by placing the gel on a long wave UV transilluminator and recording the result on Kodak 'Polaroid 667' film (Sambrook et al. 1989)

3.5.2 Electrophoresis of RNA.

RNA gel electrophoresis was performed by two methods:

3.5.2.1 Sodium phosphate/SDS method (Pellé and Murphy, 1993).

RNA in UHQ water was denatured by the addition of 1/6th volume of 2%(w/v) SDS, 30%(w/v) glycerol, 0.001%(w/v) bromophenol blue and incubation at 70°C for 15min. The denatured RNA was loaded onto a 1-1.3%(w/v) agarose gel buffered by 10mM sodium phosphate pH6.8 and containing 0.1ug/ml ethidium bromide. The RNA was electrophoresed at 3V/cm with continuous buffer recirculation. After electrophoresis the migration of the major RNA species was determined by placing the gel on a long wave UV transilluminator.

3.5.2.2 Formaldehyde method (Sambrook et al., 1989)

RNA was denatured by the addition of 2-3 volumes of formaldehyde/formamide solution (2:15) and incubation at 70°C for 10min. The samples were cooled on ice and 1/9th volume of loading buffer (122%(w/v) sucrose, 1.5%(w/v)bromophenol blue) added. Samples were loaded onto a 1.3%(w/v) gel buffered by 50mM MOPS, 1mM EDTA (pH7.0) containing 6.5%(w/v) formaldehyde. The gel was run at 5-6V/cm and stained with ethidium bromide after blotting.

3.5.3 Isolation of DNA fragments from agarose gels

Silica fine extraction was carried out according to the protocol of N.Robinson (personal communication): After electrophoresis, the band of interest was cut from the gel and dissolved in 1ml of 90.8%(w/v) NaI, 1.5%(w/v) Na_2SO_3 by heating at 65°C for 10-15 min. The dissolved gel was cooled and a small aliquot of silica fines added. The DNA was bound to the fines by incubation at room temperature for 15min. then collected by centrifugation (14000rpm, 30s) in a microfuge. The fines were washed twice with 70%(v/v) ethanol, dried then resuspended in 20-50µl of UHQ water. DNA was eluted from the fines by incubation at room temperature for 15min. The fines were removed by centrifugation (14000rpm, 30sec) and the DNA stored at -20°C.

DNA was also isolated from 1.3%(w/v) low melting point agarose TAE gels. After electrophoresis the band of interest was excised from the gel using a razor blade and surplus agarose gel trimmed away. The gel slice was then melted by the addition of 3 volumes of UHQ water and boiling for 5min. The DNA was stored at -20°C and thawed at 37°C before use (Sambrook et al. 1989).

3.6 Hybridisation of Nucleic Acids

3.6.1 Blotting of DNA and RNA.

DNA and RNA were blotted from agarose gels onto 'Hybond N' nylon membranes. DNA was blotted as described by Khandijan (1987): DNA agarose gels were incubated in 0.25M HCl, until the tracking dyes changed colour, in order to partially depurinate and cleave high molecular weight DNA to enhance transfer. The blot was then washed in three changes of denaturing solution (1.5M NaCl, 0.5M NaOH) for 15min each with shaking. The DNA was transferred to a nylon membrane overnight by capillary blotting in this alkaline solution. After transfer, the nylon filter was neutralised by a 3min

incubation on 3MM chromatography paper saturated in neutralising solution (0.5M Tris HCl pH7.2, 0.5M NaCl) and washed by a similar incubation on 6xSSC saturated paper. The blot was fixed by baking under vacuum at 80°C.

RNA was transferred to nylon membranes by overnight capillary blotting in DEPC-treated 10mM sodium phosphate pH6.8. The RNA was fixed by baking as above. Markers for the gel were removed and stained in 0.5μg/ml ethidium bromide solution then destained in DEPC treated water. The position of the markers was determined by placing the membrane on a UV transilluminator and marking the position of the bands with waterproof ink.

3.6.2 Squash blotting.

Squash blotting was carried out on floral tissue as an alternative to RNA extraction and Northern blotting. Prior to squashing, carpels and anthers were sliced in two longitudinally, while petals and sepals were kept intact. Nitrocellulose membranes for squash blotting were kept under vacuum, while being saturated in 50mM TrisHCl, 100mM NaCl. Tissue to be 'squashed' was placed on the membrane and saturated with the Tris/NaCl buffer supplemented with 50mM vanadyl-ribonucleoside complexes (Berger and Birkenmeier, 1979). Nucleic acid in the tissue was transferred to the membrane by covering the tissue with polythene and applying pressure, via a spatula. After squashing the polythene was removed and the membrane removed from vacuum before drying, removal of tissue and fixing, according to the manufacturers instructions.

3.6.3 Transfer of λ phage DNA to nitrocellulose membranes.

DNA from λ phage plaques was transfered to nitrocellulose membranes by the plaque lift method (Sambrook et al. 1989)

3.6.4 Radiolabelling of DNA fragments.

DNA fragments were radiolabelled by the random priming method (Fernberg and Vogelstein 1983). α-^{32}P-dCTP of specific activities 400Ci/mmol and 3000Ci/mmol were used. The DNA probe was purified from the unincorporated nucleotides by ethanol precipitation at -20°C for 1-3 hours. DNA was recovered by centrifugation (14000rpm, 20min), washed with 70%(v/v) ethanol and resuspended in TE.

3.6.4 Hybridisation of membranes with radiolabelled DNA probes.

3.6.4.1 Calculation of hybridisation conditions.

Hybridisation and washing conditions for probe hybrisation to membrane-bound DNA were estimated using the equation $TM=81.5+16.6\log M_{Na}+0.41(\%GC)-(500/l)$, where TM is the probe:target melting temperature, M_{Na} is the molarity of sodium ions, %GC is the percentage of G+C content of the radioactive probe and l is the length of the hybridissing species (Howley et al., 1979). For hybridisation to membrane-bound RNA the equation, used, was $TM=79.5+18.5\log M_{Na}+ 0.58(\%GC)+11.8(\%GC/100)^2$ (Casey and Davison, 1977). From the TM value the hybridisation temperature was calculated as TM-

20°C. The temperature and salt conditions for the washing were determined so that the TM was 5°C higher than the final wash temperature.

3.6.4.2 Standard hybridisation procedure.

For medium to high stringency hybridisation nitrocellulose and Hybond-N membranes were prehybridised for 2-16 hours in 5xDenhardts (1%(w/v) Ficoll 400, 1% polyvinylpyrillodine, 1% BSA), 5xSSC, 0.5%(w/v)SDS, 100µg/ml herring sperm DNA at 65°C. Hybridisation was carried out for 16-24hours in a similar solution supplemented with radiolabelled DNA probe. The blots were then washed twice with 2xSSC, 0.1%(w/v)SDS and once with 1xSSC, 0.1%(w/v)SDS at 60 or 65°C For high stringencies the blot was washed in 0.1xSSC, 0.1%(w/v)SDS at 65°C (Sambrook et al. 1989).

Lower stringency hybridisations were carried out in the above solution using the temperatures and post-hybridisation washes described in the section 4.

After hybridisation, membranes were sealed in plastic bags and exposed to Fuji RX medical X-ray film with intensifying screens at -80°C.

3.6.4.3 Hybridisation of pHA1 to Northern blots.

To assess the relative loading and quality of RNA on Northern blots they were probed with the *P.sativum* ribosomal clone pHA1, which contins genomic 18S and 28S sequence (Cheung and Cohen, 1975). Hybridisation was carried out as idescribed in section 3.6.4.2 and the blot washed to 0.5xSSC, 0.1%(w/v)SDS at 65°C before exposure to X-ray film. After development the autoradiograph was checked, by eye, to assess the intactness of bands (quality of the RNA) and equivalence of signal (loading). Full densitometric scanning was not performed as the Northern analysis performed on the isolated clones was only semi-quantitative.

3.7 cDNA library construction.

3.7.1 cDNA synthesis

2µg of stigma poly(A)RNA was reverse transcribed using NotI-d(T)$_{18}$, primers in a total volume of 33µl, using glass stabilised MuMLV reverse transcriptase, supplied with its own buffer (final concentration 45mMTrisHCl (pH8.3), 68mMKCl, 15mM DTT, 9mM MgCl2, 0.08%BSA, 7.2mM dNTP), according to the manufacturers instructions. RNA primers for second strand synthesis were generated by addition of 10U of RNAseH and 25U of DNA polymerase I with 10µl of Boehringer Mannheim 2nd strand synthesis buffer in a total volume of 100µl. The reaction was carried out by incubation for 60min at 12°C then a further 60min at 22°C. The reaction mixture was then heated at 65°C for 10min to terminate all enzyme activities.

In order to estimate the efficiency of conversion of mRNA to dsDNA one tenth of the second strand reaction was mixed with 2µCi of α^{32}P-labelled dCTP. This radiolabelling reaction was split into two and each half was applied to a GFC glass fibre filter. One filter was washed with 10%(w/v) TCA at 0°C to selectivly precipitate the DNA (incorporated radionucleotide) while the second filter was retained as

a measure of total radioactivity (Sambrook et al. 1989). The specific activities of both filters were then measured in a Canberra Packard TR1600 liquid scintillation counter.

In order to polish the double stranded cDNA, 4U of T4 DNA polymerase were added and the reaction incubated at 37°C for 10min. To terminate this reaction 10µl 200mM EDTA (pH7.2) and 2µl 10%(w/v) N-laurylsarcosine were added. The DNA was purified by one extraction with phenol/chloroform/IAA (25:24:1), passage through a Sephacryl S300 spin column and ethanol precipitation at -80°C for 16 hours. The DNA was pelleted (14000rpm, 10min, 4°C), washed with 70% ethanol and resuspended in 20µl of UHQ water.

3.7.2 Cloning the cDNA into λ ExCell

10µl of the DNA was ligated to a 50 fold molar excess of EcoR1 adaptors with T4 DNA ligase in a total volume of 20µl at 15°C for 16 hours. The ligase was inactivated by incubation at 70°C for 15min and the DNA prepared for ligation into vector by the addition of 10µl of 10xNotI restriction buffer (500mM TrisHCl (pH7.5), 1M NaCl, 100mM $MgCl_2$, 10mM DTE) and 30U of NotI 2U of T_4 polynucleotide kinase in a final volume of 100µl. Digestion/phosphorylation was carried out for 90min at 37°C. The cDNA was then purified by phenol/chloroform/IAA (25:24:1) extraction, passage through a Sephacryl S300 spin column and ethanol precipitation at -80°C for 16 hours. The DNA was pelleted (14000rpm, 10min, 4°C), washed with 70% v/v) ethanol and resuspended in 8µl of UHQ water.

3.1µl of this DNA was ligated to 1.5µg of λExCell NotI/EcoRI predigested DNA with 0.8U of T4 DNA ligase in a total volume of 8µl. The ligation was carried out at 15°C for 16 hours. The ligase was then deactivated by incubation at 70°C for 15min. 3µl aliquots of this ligation were then packaged using a Stratagene 'Gigapack II' kit according to the manufacturers instructions.

3.7.3 Amplification of the cDNA library.

The cDNA libray was amplified by plating out on a $23cm^2$ plate and growing for 16 hours. The amplified library was collected by washing the plate with 15ml of SM for 24hr at 4°C. The eluted phage were collected into a 15ml falcon tube and chloroform added to 2%((v/v). The phage suspension was clarified by centrifugation (900g, 10min) and stored at 4°C after the addition of a further 2%(v/v) chloroform.

3.8 Protein extraction and analysis.

3.8.1 Protein Extraction

Total protein extractions on plant material were performed by freezing the material in liquid nitrogen then grinding in a 1.5ml eppendorf tube. Extraction buffer (50mM TrisHCl (pH6.8), 5%(w/v) SDS, 1%(v/v) β-mercaptoethanol) was added at 4µl/mg. and the tube was incubated in a boiling water bath for 5 min. Particulate material was removed by centrifugation (14000rpm, 5min).

3.8.2 SDS-PAGE

SDS-PAGE was carried using out the discontinuous method of Laemmelli (1970) in BioRad Mini-Protean II gel apparatus (Bio-Rad, York). After electrophoresis, the seperating gels were fixed and stained in methanol:acetic acid:water (5:2:5) containing 0.1%(w/v) coomasie blue stain. The gels were destained in methanol:acetic acid:water (4:1:5).

3.8.3 Glycoprotein staining with ConA and alkaline phosphatase.

Glycoproteins were detected by binding of ConA then alkaline phosphatase to proteins immoblised on nitrocellulose. The method was provided by Dr D.Robertson of the University of Durham, Department of Biological Sciences: SDS-PAGE gels were run as described in section 3.8.2 and the electroblotted onto nitrocellulose membrane using a Whatman Kem En Tec semi-dry blotter with 48mM Tris, 39mM glycine, 1.3mM SDS, 20%(v/v) methanol as transfer buffer. The membrane was then washed 3x15min in 5xTBS before blocking in 5xTBS (10xTBS is 1M TrisHCl (pH8.0), 1.5M NaCl) supplemented with 1% BSA (1 hour). The membrane was washed (3x15min in 5xTBS) then incubated for 1 hour in 5xTBS supplemented with 1µg/ml biotinylated ConA. Unbound ConA was then removed by washing prior to a 1 hour incubation in 5xTBS supplemented with Extravidin-alkaline phosphatase (0.15µg/ml). Unbound Extravidin was removed by washing (3x15min, 5xTBS) before incubation of the blot in developing buffer (100mM TrisHCl (pH9.0), 100mM NaCl, 5mM $MgCl_2$ containing 0.26mg/ml NBT and 0.016mg/ml BCIP). The developer was changed every 5min until the desired signal was obtained when the reaction was terminated by washing in water.

3.9 Computer Analysis of Data.

DNA and protein sequences were manipulated and interchanged using DNA Strider (Marck, 1988) and the GCG package (Devereux et al., 1984) at the Daresbury laboratory, Warrington. DNA Strider was used to generate hydropathy plots, using the method of Kyte and Doolittle (1982), and to calculate amino acid composition of proteins together with codon usage. Garnier structure predictions (Garnier et al., 1978) were performed using the 'Protean' program from the DNASTAR suite of programs for the Apple Macintosh.

Comparison of DNA and protein sequences with those in the Genbank and Swissprot databases was carried out using the FASTA and BLAST programs available through the European Bio-informatics Institute, Cambridge. Sequences which showed homology to those studied during this PhD project were also obtained from this site. Multiple alignment of DNA and protein sequences was achieved using the Clustal W program (Thompson et al., 1994). Preliminary analysis of the Clustal alignments was performed using DNAdraw 2.5.1 (M. Shapiro, unpublished) using exact matches or equivalence groups (defined by the amino acid subsets GAP, MILV, FYW, DE, NQ, KHR, ST, C). The final shaded alignments that appear in the results were produced manually with Microsoft Word 5.0.

Searching of the PROSITE library and screening of sequences for defined patterns was carried out using the MacPattern program (Fuchs, 1994).

4
Results and Discussion.

4.1 The clones isolated from a previous library screen.

4.1.1 The clone 13G6(i).

4.1.1.1 Sequence of the cDNA clone.

The 13G6(i) insert was initially cloned in the plasmid pUC18 as an EcoRI fragment. To facilitate its sequencing, the insert was subcloned into pKS+Bluescript as this latter vector has a larger number of restriction sites in its polylinker which increases the possible options for generating subclones for sequencing. The full sequence of the 13G6(i) clone was then obtained by subcloning and sequencing as shown in fig.5a. The cDNA clone is 1001 bases long which is shorter than the transcript size indicated by the previously reported Northern blots, so 13G6(i) only represents a partial sequence. As two putative polyadenylation signal sequences and a poly(A) tail are present in the clone, the 5' end is missing from the sequence .

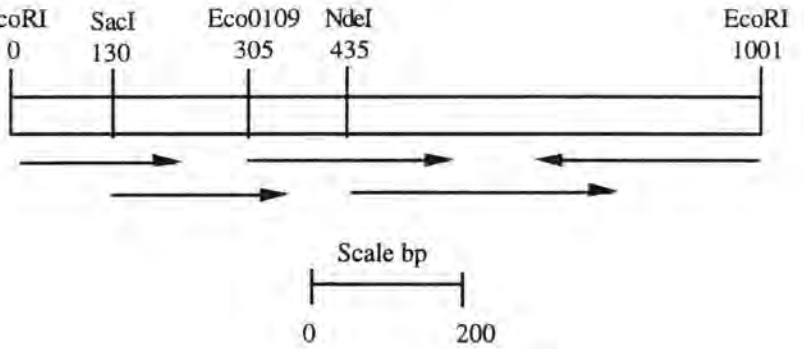
Searching for ORF's, in 13G6(i), revealed one long reading frame coding for 259 amino acids as indicated in fig.5b. The sequence around the first ATG codon (nt. 41), in this putative protein, does not conform to the consensus for translation initiation in plants (TGGCATGAA as opposed to AACAATGGC) (Lutcke et al. 1987). Therefore the N-terminal region of the protein is probably encoded by the missing sequence. From this information the sequence of 13G6(i) and its putative protein are as shown in fig.5c.

The putative 13G6(i) amino acid sequence was analysed for possible functions of the protein: The predicted 13G6(i) protein does not contain any known structural motifs as defined by the PROSITE database or by the consensus sites for glycosylation. Similarly, a hydropathy profile of the predicted 13G6(i) protein does not show any significant hydrophobic regions which might act as membrane spanning/interacting motifs (fig.6a). Plotting of charged amino acid residues reveals no acidic or basic clusters are present except at the C-terminus in which 4 of the last 6 residues are glutamates (fig.6b). The relevance of this acidic cluster is not known.

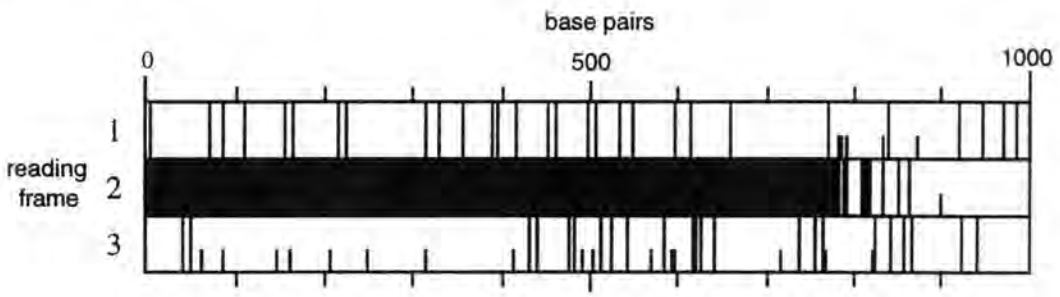
A principal feature of the 13G6(i) protein is that several fragments of sequence in the N-terminal half of the protein are repeated in the C-terminal. Alignment of the two halves of the known protein sequence indicates that 13G6(i) appears to consist of a tandomly duplicated repeat (fig.7). However the exact nature of this structure could not be determined due to the lack of the N-terminal coding region of 13G6(i).

This structure is not unique amongst plant genes. The globulin seed storage proteins are also believed to have arisen by duplication of a single ancestral gene. Evidence for this not only comes from alignment of the protein sequences (Shutov et al., 1995,Gibbs et al., 1989) but also from secondary structure predictions (Argos et al., 1985), alignment of intron/exon boundaries in genomic sequences (Shutov et al., 1995), positions of proteolysis-sensitive sites (Shutov et al., 1995) and the three dimensional crystal structures (Lawrence et al., 1994, Gibbs et al., 1989). However the level of identity between the duplicated halves are lower amongst these proteins than between those of the predicted 13G6(i) protein suggesting that 13G6(i) is not one of this class of proteins.

(a)



(b)



(c)

```
1.   CCTTGATCGCACTATCCAGTTTTACACAGAGTGCTTTGGCTTGAAGGTGTTGAGGAAAAGAGATGTACCTGAGGAGAAGT .80
       L  D  R  T  I  Q  F  Y  T  E  C  F  G  L  K  V  L  R  K  R  D  V  P  E  E  K
81.  ACTCTAATGCTTTCCTTGGTTTTGGTCCTGAAACCTCCAACTTCGTTGTCGAGCTCACTTACAATTATGGTGTTAGCTCA .160
     Y  S  N  A  F  L  G  F  G  P  E  T  S  N  F  V  V  E  L  T  Y  N  Y  G  V  S  S
161. TATGACATTGGAACTGGATTTGGGCATTTCGCTATTTCAACTCAAGATGTTTCTAAGATGGTTGAGGCGGTTCGTGCCAA .240
       Y  D  I  G  T  G  F  G  H  F  A  I  S  T  Q  D  V  S  K  M  V  E  A  V  R  A  K.
241. GGGTGGAAATGTCACCAGAGAGCCTGGTCCGGTCAAAGGTGGAGGCAGTGTCATTGCCTTTGTCAAGGACCCTGATGGTT .320
       G  G  N  V  T  R  E  P  G  P  V  K  G  G  G  S  V  I  A  F  V  K  D  P  D  G
301. ACACTTTTGAGCTTATCCAGAGAGGTCCAACTCCTGAACCTCTCTGTCAAGTCATGCTTCGTGTTGGTGATCTTGACCGC .400
     F  Y  T  E  L  I  Q  R  G  P  T  P  E  P  L  C  Q  V  M  L  R  V  G  D  L  D  R
361. GCCATCAAGTTCTATGAAAAGGCCCTTGGGATGAGACTCTTGAGAAGGATTGAGAGGCCTGAATACAAGTACACCATAGG .480
       A  I  K  F  Y  E  K  A  L  G  M  R  L  L  R  R  I  E  R  P  E  Y  K  Y  T  I  G
481. CATGATGGGATATGCTGAGGAATATGAGTCAATAGTTTTGGAGCTGACCTATAACTACGGTGTGACTGAGTACACAAAGG .560
       M  M  G  Y  A  E  E  Y  E  S  I  V  L  E  L  T  Y  N  Y  G  V  T  E  Y  T  K
561. GCAACGCATATGCACAGATTGCAATAGGCACGGATGATGTGTACAAAAGCGCTGAAGTAGTGAAGATAGCCAACCAAGAG .640
     G  N  A  Y  A  Q  I  A  I  G  T  D  D  V  Y  K  S  A  E  V  V  K  I  A  N  Q  E
641. CTAGGAGGAAAGATCACTAGAGAAGCCGGACCTCTTCCTGGACTCGGCACCAAGATTGTCTCATTCCTCGATCCAGATGG .720
       L  G  G  K  I  T  R  E  A  G  P  L  P  G  L  G  T  K  I  V  S  F  L  D  P  D  G
721. CTGGAAAACAGTTCTGGTAGACAACGAAGATTTTCTGAAGGAACTGAATGAGGAAGAGTAATGATGGTGATGATCACAGA .800
       W  K  T  V  L  V  D  N  E  D  F  L  K  E  L  N  E  E  E
801. CAAAACCATAATAATAATAAAATGTGATCTTATGAAGTAGAGTGAAGACTTAGTATTTAATCTAAATAAGATGTTGGTTC .880
881. AGTGTGCGTCGTGTTTCTATGTGTTTGTTCCTTGTCTTTGAGCTAAACTCTGTTACGTGCCTGAATAAATATATTTCCAC .960
961. CACCTGGTTTAGGTGCTTTCCGGTTAAAAAAAAAAAAAAAA                                        .1001
```

Fig. 5: Sequencing strategy and analysis of the cDNA clone 13G6(i). (a) Sequencing strategy for the clone 13G6(i). Arrows indicate the direction and sequence length obtained from each subclone. Numbers represent the restriction site position from the 5' end. (b) ORF map for the clone 13G6(i). Full bars indicate stop codons, half height bars indicate putative translation initiation codons. The shaded region indicates the predicted coding region. (c) Nucleotide and putative protein sequences of the clone 13G6(i). The consensus polyadenylation signals are underlined.

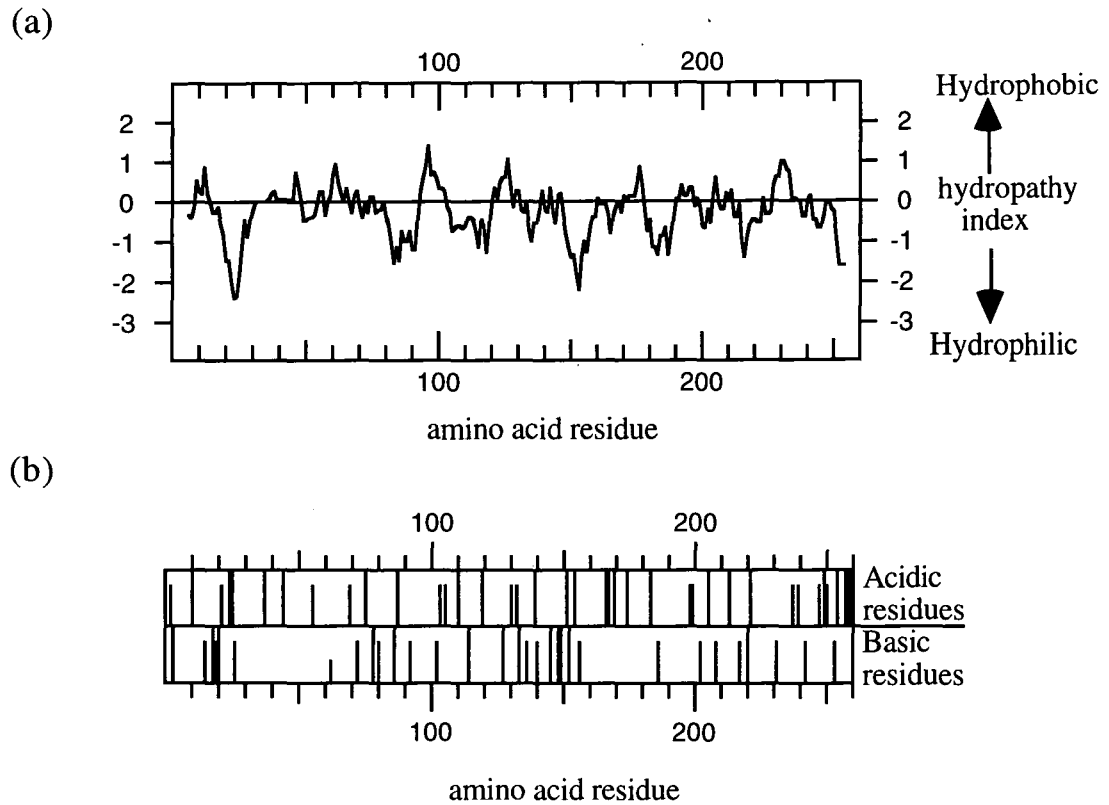


Fig.6. Hydropathy and acid/base distribution plots for the predicted protein encoded by the cDNA clone 13G6(i). (a) hydropathy plot by the method of Kyte and Doolittle. (b) acidic and basic residue plot. The length of each bar is proportional to the propensity to form a charged species at neutral pH (Marck,1988)

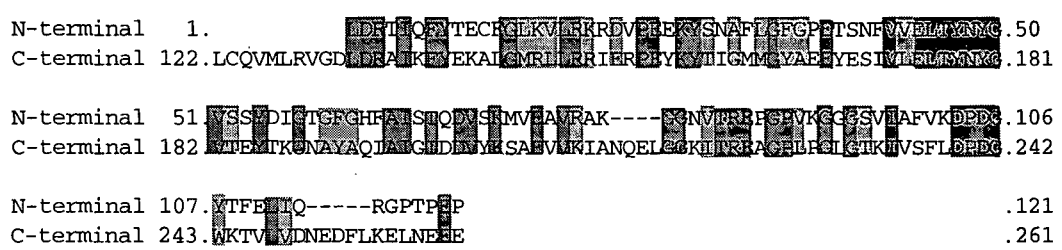


Fig.7. Clustal W 1.5 alignment of the two halves of the putative 13G6(i) protein. The point of division into two halves is arbitrary as the high homology between the two putative domains does not extend to the alignment of the C-terminal or to the centre of the protein. Identical amino-acid matches are indicated by dark grey shading. Conservative residue pairs, as defined in section 3.9, are indicated by light grey shading.

4.1.1.2 Eukaryotic homologues of 13G6(i)

13G6(i) was used to screen the EMBL database to determine whether it represents a novel plant sequence or if there are already similar sequences known. Several sequences similar to 13G6(i) were identified, most of which are plant EST's (table 4). The EST's show similarity to several different regions of 13G6(i) (fig.8) with the level of identity to 13G6(i) ranging from 95% over ~300 bases to 65.7% over 201 bases (table 4).

In order to determine whether these sequences are actually related to 13G6(i) the EST's were translated and the putative proteins aligned to that of 13G6(i) (data not shown). All of the EST's have the capacity to partially encode a protein with high identity to that of 13G6(i), although there is a requirement in some clones for frame shifts in order to maintain the protein homology and in all the clones the level of identity to 13G6(i) falls near their 3' end. As all these sequences are from single pass sequencing (references in table 4) they almost certainly contain errors (most sequences have unassigned bases) and so, despite the frame shifts and 3' differences, these EST's probably represent homologues of 13G6(i) in their respective species.

The EST's are from the related Crucifers *A.thaliana* and *B.campestris* and the monocots *O.sativa* and *Z.mays*. The pairs of clones from the latter three species are identical except for unidentified bases so probably represent the same transcripts. The overlapping regions of the *A.thaliana* clones also show high identity to each other and so may represent identical transcripts. The high level of homology between their putative proteins suggests a high degree of sequence conservation which is often indicative of a conserved function.

Further evidence for such a conserved role would come from similar patterns of expression. As the clones are all EST's a complete description of their expression is not available, although their tissue of origin can be compared with the expression reported by Scutt (1990). The only clones, whose origins are consistent with the reported 13G6(i) expression, are those from *B.campestris* which were both isolated from flower buds (Lim et al., 1994). Their known mode of expression therefore encompasses the reproductive organs in which 13G6(i) expression has been reported.

Most of the *A.thaliana* clones were isolated as part of work involved in mass sequencing of *Arabidopsis* transcripts (Newman et al., 1994). These were derived from a mixed pool of RNA and so their tissue of origin is unknown. Three sequences were from a different screening program and their tissue of origin is known: Atts4074 was derived from etiolated seedlings (Desprez et al., 1994), Atts5748 from nitrate-treated roots (Krivitzky, 1995) and Atts5265 from a cell suspension culture (Cooke et al., 1996). The origin of the monocot sequences is also known: the *O.sativa* clones were derived from root and shoot RNA (Sasaki et al., 1994) while the *Z.mays* clones were derived from seedling (Shen et al., 1994).

This expression in vegetative tissue suggests that the protein encoded by 13G6(i) is not a specialised component of the reproductive tissue. However, in the previous work the expression in these tissues was not assessed, although due to the close evolutionary relationship between *Brassica* and *Arabidopsis*, similar patterns of expression between the two species would be expected. More detailed

Plant	Clone	Identity/overlap	Reference
A.thaliana	At31012	87.1%/442bp	Newman et al., 1994
	At70414	85.6%/361bp	Newman et al., 1994
	At283	85.7%/300bp	Newman et al., 1994
	Atts4074	85.6%/292bp	Desprez et al., 1994
	At36616	83.2%/382bp	Newman et al., 1994
	At9541	79.6%/328bp	Newman et al., 1994
	Atts5265	81.2%/271bp	Cooke et al., 1996
	Atts5748	80.0%/260bp	Kriviyzky et al., 1995
	At44114	81.2%/191bp	Newman et al., 1994
	At6225	65.7%/201bp	Cooke et al., 1996
B.campestris	Bcestq	95.1%/307bp	Lim et al., 1994
	Bcestc	92.4%/238bp	Lim et al., 1994
O.sativa	Osr27361a	74.9%/219bp	Sasaki et al., 1994
	Oss1454a	77.2%/79bp	Sasaki et al., 1994
Z.mays	Zm293	72.3%/141bp	Shen et al., 1994
	Zm638	71.6%/141bp	Shen et al., 1994

Table 4. Identity and source of the eukaryotic nucleic acid sequences similar to 13G6(i) as identified by screening of the EMBL nucleic acid database.

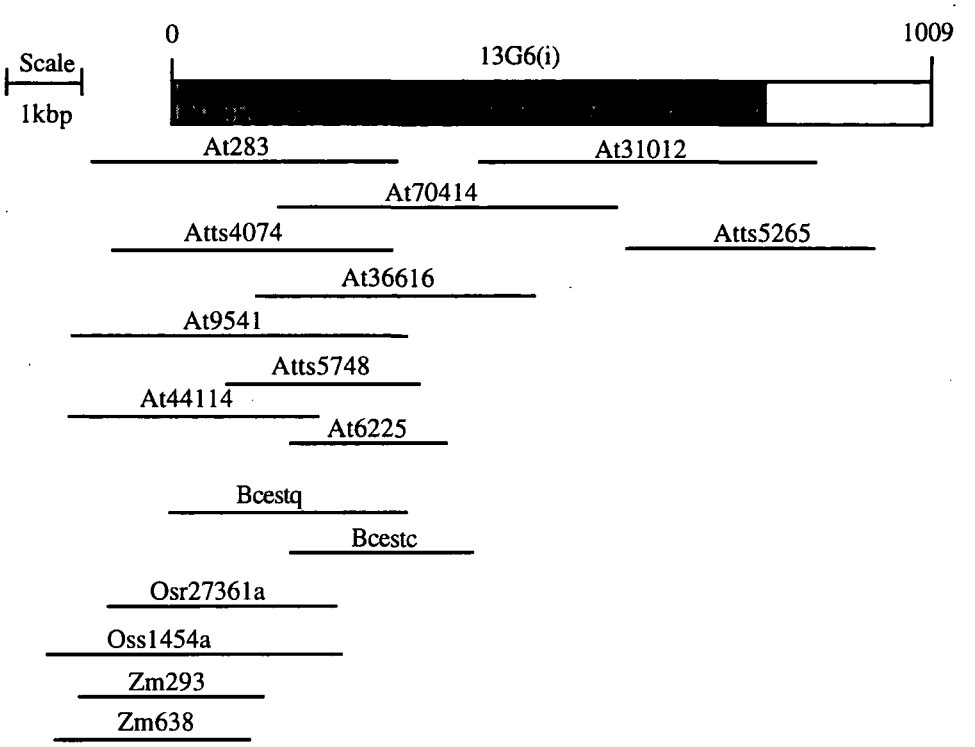


Fig.8. Regions of homology between 13G6(i) and the EST's described in table 4. The shaded region of 13G6 indicates the extent of the coding region of the cDNA.

analysis of the expression of 13G6(i) homologous sequences in *B.oleracea* is described in sections 4.1.1.7 to 4.1.1.9.

4.1.1.3 Prokaryotic homologues of 13G6(i).

As well as the eukaryotic EST's with similarity to 13G6(i) two prokaryotic DNA sequences were identified. Both of these were genomic clones: Vp06989 is an ORF identified while sequencing around the locus of a *Vibrio parahaemolyticus* flagellar protein (McCarter, 1994). Hihi0323 is a putative gene identified during the sequencing of the genome of *Haemophilus influenzae* (Fleischmann et al., 1995). These sequences are 61% over 349bp and 58% over 341bp identical to 13G6(i) respectively. The relevance of this similarity was shown by screening the SWISSPROT protein sequence library with the putative protein sequence of 13G6(i), from which the only significant matches were the two proteins (YRN1 of *Vibrio* and HI0323 of *Haemophilus*) encoded by these ORFs.

Each hypothetical protein is under half the length of the known sequence of 13G6(i) and both proteins are homologous to both putative halves of the 13G6(i) protein (fig.9) so providing more evidence for the double-domained structure of the putative 13G6(i) protein. Further evidence that the proteins might be related comes from plotting the hydropathy profiles which appear similar for each half of the putative 13G6(i) protein and the bacterial sequences (fig.10)

Given the relatively high homology between the nucleic acid and predicted protein sequences of 13G6(i) and these bacterial genes suggest that the former may be related from the latter and therefore have similar activities/common functions. If the sequences are related then 13G6(i) must have arisen by gene duplication rather than fusion between two similar genes as only one locus with similarity to 13G6(i) occurs in *Haemophilius* (Fleishman et al., 1995). The two halves would then have diverged by point mutation and small insertions/deletions. Mutation is not restricted to specific domains but differs between the two 13G6(i) repeats and the bacterial sequences. This is most apparent in the regions conserved between the two halves of the predicted 13G6(i) protein and the bacterial sequences: The N-terminal half of 13G6(i) shows highest identity to the C-terminal region of the bacterial proteins while the 13G6(i) C-terminal half has its highest identity to the N-terminal half of the bacterial proteins.
Once again this is analogous to the seed globulins. Alignment of the protein structures of these proteins allowed the identification of a potential conserved motif which was also identified in plant germins and fungal sperulins (Bäumalein et al., 1995). Alignment of the genomic sequences encoding these plant and fungal proteins revealed conservation of intron/exon boundaries which was interpreted as indicating an actual evolutionary relationship, as opposed to being a result of convergent evolution. This led to the hypothesis that the globulins, germins and spherulins all share a common ancestor, or the former two, plant, protein families evolved from the latter. The latter is supported by functional considerations in that the spherulins are believed to play a role in dessication during spherulation with which analogies to seed maturation can be drawn.

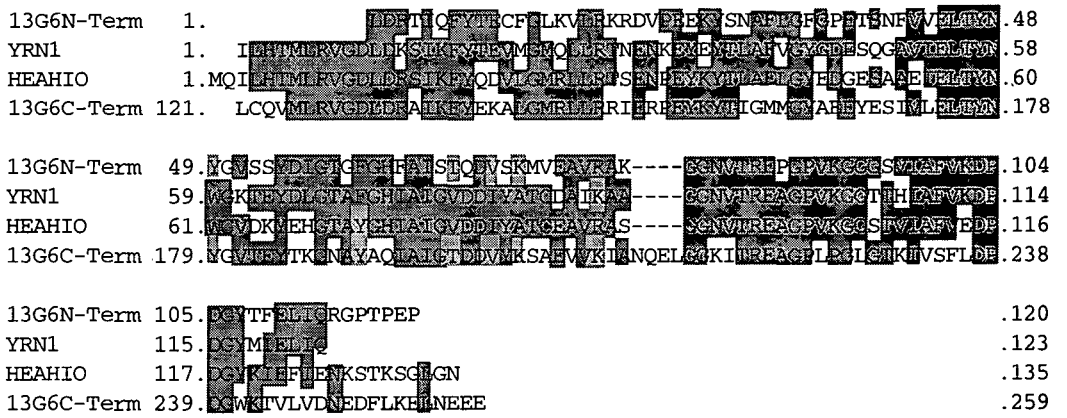


Fig.9. Alignment of the predicted 13G6(i) protein with the putative bacterial proteins YRN1 and HEAHIO. The division of the 13G6(i) protein is different from that in fig.7 as the optimal alignment is achieved by splitting the protein at a point 2 residues away from the previous division. Identically matching residues are indicated by dark grey shading. Where two different pairs of matching residues occur at the same position one pair is shaded in light grey to differentiate it from the other

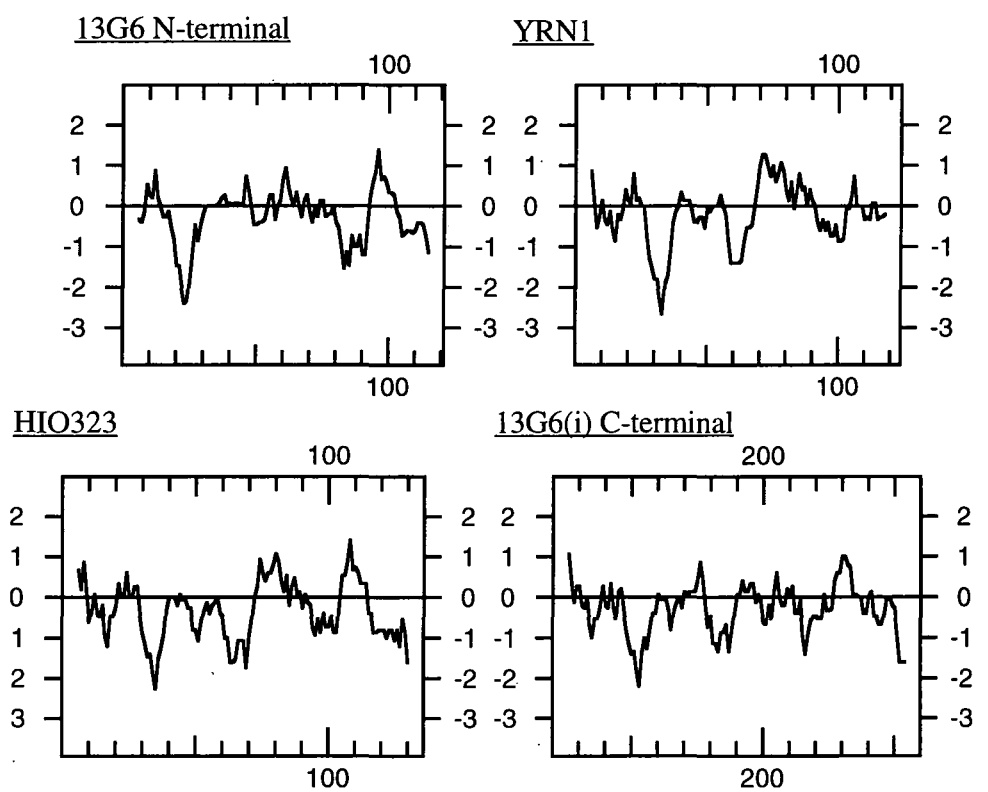


Fig.10. Comparison of the hydropathy profiles of the bacterial proteins YRN1 and HIO323 with the two halves of the predicted 13G6(i) protein. Horizontal scales are the amino acid residue numbers while the vertical scales represent hydropathy as calculated by the Kyte-Doolittle algorithm with positive and negative values representing hydropathic and hydrophilic regions of protein respectively.

The relationship between 13G6(i) and the bacterial proteins is more difficult to analyse, given the lower amount of sequence data available. As with the seed storage globulins and spherulins convergent evolution cannot be discounted as reason for the similarity. Unlike the spherulin/storage protein work intron/exon analysis would not be possible between 13G6(i) and bacterial genomic sequences as the prokaryotic genes lack introns (McCarter, 1994, Fleischmann et al., 1995).

To determine the exact relationship between the sequences the gene homologues from other species, such as animals, lower plants and lower eukaryotes, could be isolated and sequenced. These could provide information on the, possible,intermediate stages of 13G6(i) evolution between prokaryotes and higher plants. Data from this work could be used to pinpoint the most highly conserved residues in the amino acid sequence which would be suitable targets for mutagenesis studies which could be used to investigate sequence/activity relationships if the function of 13G6(i) or the bacterial genes is determined in the future.

While the presence of these homologues suggests that 13G6(i) is a member of a conserved gene family, the lack of function/biochemical activity available for these sequences means that they provide no information on what the role of the predicted 13G6(i) protein may be in *B.oleracea*. Therefore further work was carried out to determine further sequence information and the pattern of expression of the *Brassica* gene.

4.1.1.4 Genomic organisation of 13G6(i).

It had previously been reported that a probe for 13G6(i) only hybridises to a small number of bands on *B.oleracea* genomic Southern blots, suggesting that it may represent a member of a small gene family (Scutt, 1990). In order to confirm this genomic organisation, gene reconstructions were carried out using the pKS+Bluescript/ 13G6(i) construct as a standard for quantitation. By comparing the signal intensity of the genomic bands with the controls ait is estimated that two sequences homologous to 13G6(i) are present in the *B.oleracea* genome (fig.11). Identical hybridisation patterns were obtained at both low (80%, 1xSSC, 0.1%(w/v)SDS, 65°C) and high (98%, 0.1xSSC, 0.1%(w/v)SDS, 65°C) stringency. This suggests that 13G6(i) in *B. oleracea* is part of a small conserved gene family.

The level of homology between the 13G6(i)-like EST's is sufficiently high that these sequences should also be detectable by low stringency hybridisation on genomic blots using 13G6(i) as a probe. Therefore the presence of 13G6(i)-homologous sequences in other plant genomic DNA's was assessed on a Southern blot. The high similarity of the bacterial genes to the *Brassica* sequence suggests that the latter may be ancestral genes so that an animal homologue would also be predicted to be present. Therefore human DNA was also included in the study to determine if such an animal counterpart could be detected. As the genomic DNA's chosen for study came from organisms with C values ranging from came from 0.07 to 2.5pg per genome (Arumuganathan and Earle, 1991), the quantity of each genomic DNA loaded on the gel was proportional to the C-value in order to produce similar levels of hybridisation signal from the different DNA's.

Positive signals were obtained from all DNA's under conditions allowing up to 35% mismatch between probe and target (fig.12a). The number of copies of sequences which cross-hybridise with 13G6(i) varied between species with the extent of hybridisation tending to be proportional to the C-value of the

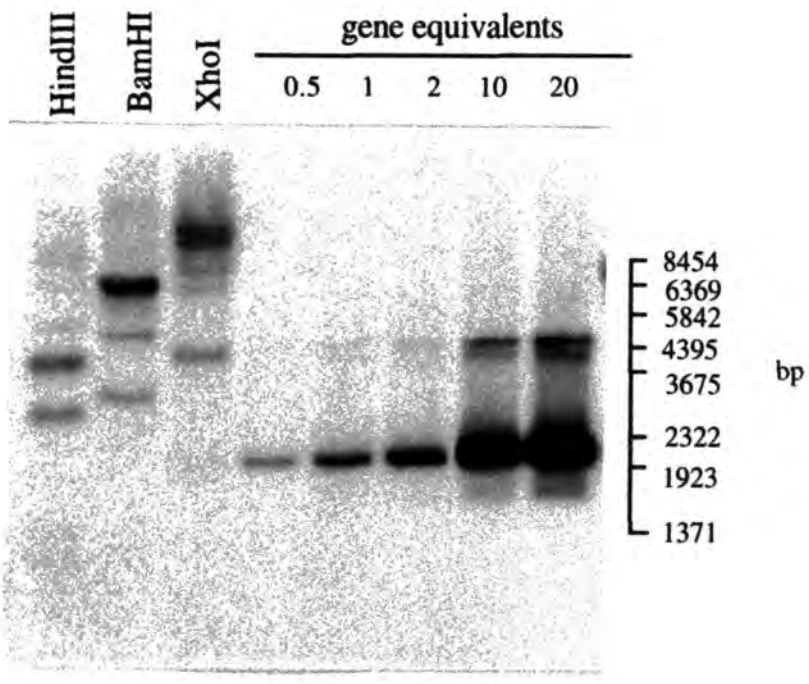


Fig.11.Quantitation of the gene copy number for 13G6(i)-like sequences in the genome of *Brassica oleracea* var.*alboglabra*. 10µg of genomic DNA was used for each digestion. The autoradiograph is of a blot washed to ~85% stringency (1xSSC, 0.1%(w/v)SDS, 65°C). No change in this pattern of hybridisation was seen by rewashing the blot to ~98% stringency (0.1xSSC, 0.1%(w/v)SDS, 65°C).The sizes were derived from a lambda BstEII marker co-run on the gel with the genomic DNA.

relevant plant. This suggests that in some species the 13G6(i) family may be considerably more complex than in the Cruciferae. However the DNA in many of the high C-value plants is smeared, probably due to overloading or non-specific hybridisation. The latter is possible as more sequences will be available to non-specifically bind the 13G6(i) probe at low stringency in plants with relatively complex genomes. The effect of high concentrations of target DNA on hybridisation is illustrated by the λ DNA size marker which also bound probe at low stringency, even though alignment of lambda DNA and 13G6(i) reveals no regions of significant identity (data not shown). This pattern does not hold with human DNA, which has a C-value similar to *B.napus*, where multiple smeared bands occur suggesting that a large number of sequences similar to 13G6(i) occur in man.

In *A.thaliana* and *O.sativa* only one major high molecular weight hybridising band is visible, of similar or weaker intensity to the *Brassica* bands, which could indicate a single copy gene. Unfortunately on this blot, band intensity is not directly related to copy number, when comparing hybridisation between two species as probe/target mismatching can effect the intensity of hybridisation. However if single copy genes are present then the promoter of each gene must be capable of driving a complex pattern of expression in different tissues, given the isolation of the relevant EST's from different tissues.

The blot was then rewashed at higher (85%) stringency (fig.12b). This did not affect the pattern of the principal bands in the Cruciferae, although the single band in *Arabidopsis* DNA decreased in intensity relative to the *Brassica* bands. This would be expected as most of the *Arabidopsis* EST's are 80-90% identical to 13G6(i) so some stripping of probe, from the blot, would be expected at the bottom of this range.

Much of the signal in the monocots and large-C-value dicots is lost, which would also be predicted from the similarities between the *B.oleracea*, *Z.mays* and *O.sativa* sequences. This also suggested that non-specific hybridisation was occurring at low stringency giving the smeared signal. A single, ~4000bp, band still remains in *P.sativum*, *T.aestivum*, *H.vulgare*, *Z.mays* and *H.sapiens*. The presence of such a highly conserved sequence on the same size restriction fragment between these species is unlikely, particularly as different restriction fragment patterns occur in the three Cruciferae assayed. The size of the band is equivalent to the linearised pBluescript 13G6(i) clone (2968bp plasmid + 1001bp insert). As such a DNA fragment would be produced by HindIII digestion, contamination by this nucleic acid species is the probable cause of the observed hybridisation pattern.

4.1.1.5 Sequencing of a genomic homologue of 13G6(i).

In order to obtain more sequence information it was decided to isolate a genomic clone from a *B.oleracea* var. *alboglabra* strain S29 genomic library in λFixII. Genomic clones, homologous to 13G6(i) were isolated from an initial screen of 10^6 plaques (~6 genome equivalents). This produced 13 initial clones, which is consistent with the predicted gene copy number. Two of these clones were purified and mapped using SacI, SalI, XbaI, KpnI and AccI. A complete restriction map was deduced for one clone (λ10) but for the second clone (λ8) only the SacI and XbaI sites were fully mapped (fig.13). The different restriction maps for these two clones suggests that they represent both copies of the 13G6(i) sequences in

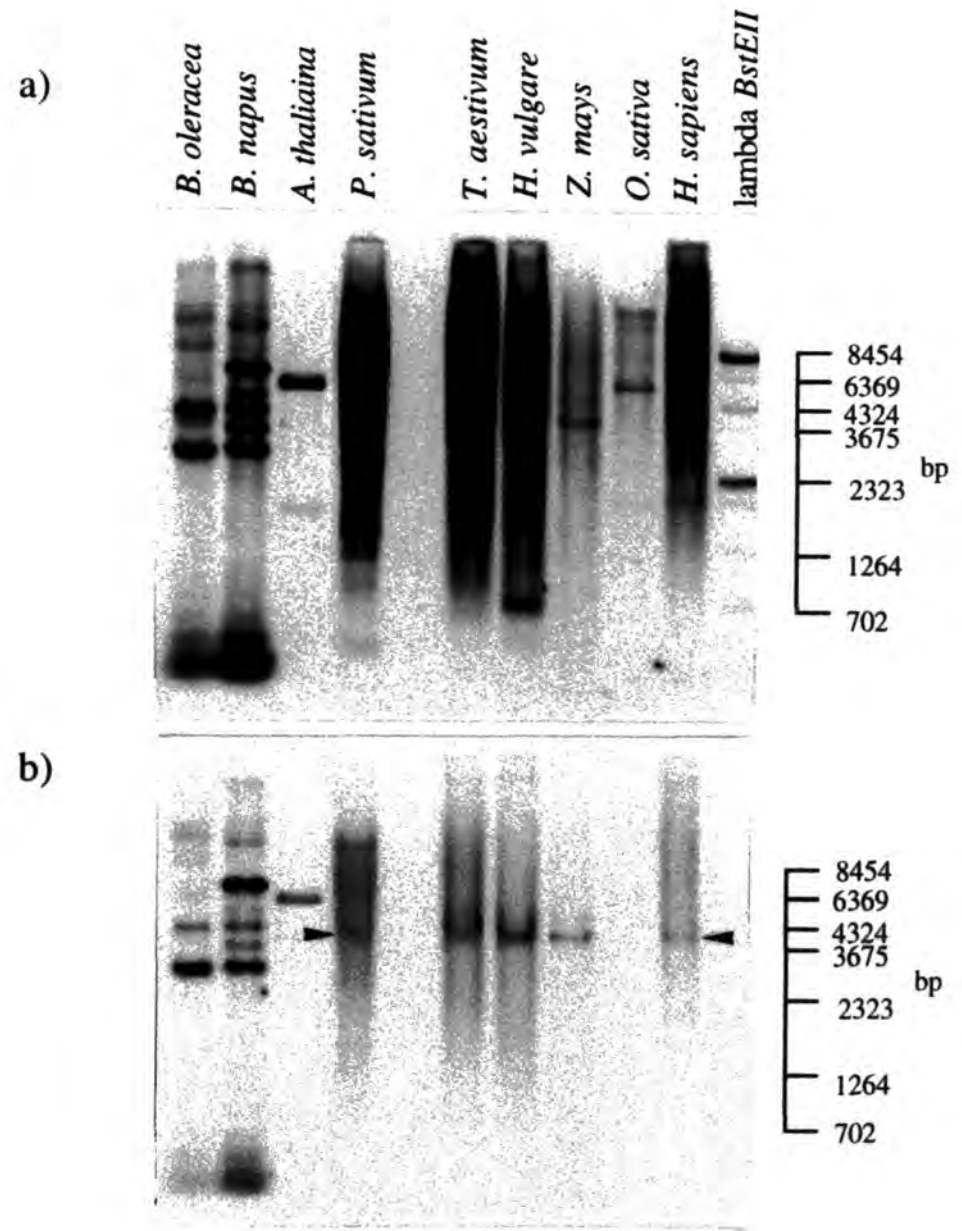


Fig.12. Hybridisation of the cDNA 13G6(i) to HindIII digested genomic DNA from a variety of plant species and man. The quantity of DNA loaded is proportional to the C-value of the relevant organism according to Arumuganathan and Earle (1991) and Alberts et al. (1989). Actual amounts of DNA loaded were *Brassica oleracea* 5.0µg, *Brassica napus* 9.8µg, *Arabidopsis thaliana* 1.2µg, *Pisum sativum* 31.4µg, *Triticum aestivum* 42.4µg, *Hordeum vulgare* 38.8µg, *Zea Mays* 18.2µg *Oryza sativa* 3.8µg and *Homo sapiens* 18.2µg. The blot was hybridised at 45°C. (a) Autoradiograph of the blot washed in 1xSSC, 0.1%(w/v)SDS at 52°C to give a stringency of ~65%. (b) Autoradiograph of the blot after rewashing in 1xSSC, 0.1%(w/v)SDS at 65°C. The apparently conserved 4000bp band is indicated by arrows.

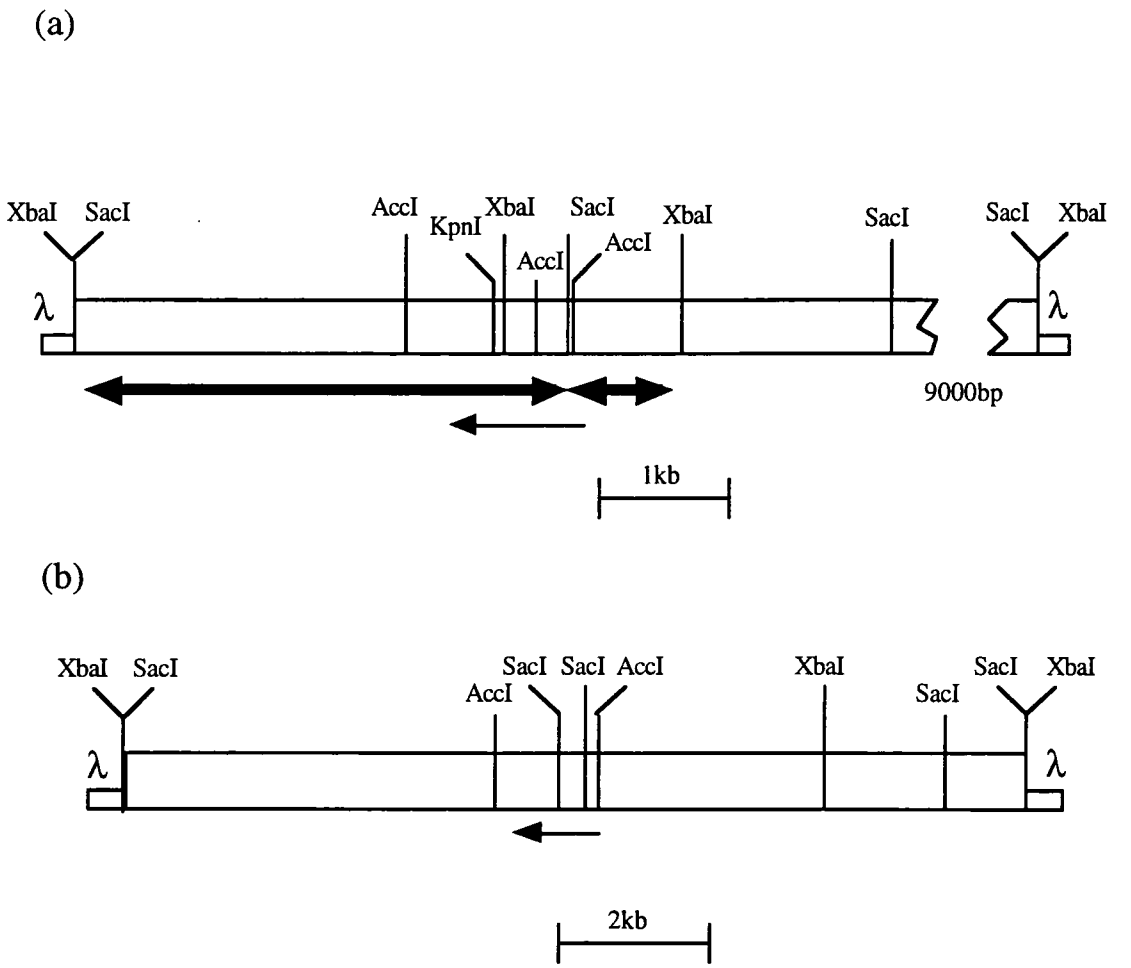


Fig.13. Maps of the two isolated λ genomic clones homologous to 13G6(i). (a) λ10, (b) λ8. Thin arrowed lines indicate the regions with which 13G6(i) cross-hybridises and their relative orientation (3' at point of arrow). The thick double arrowed bars indicate cloned regions.

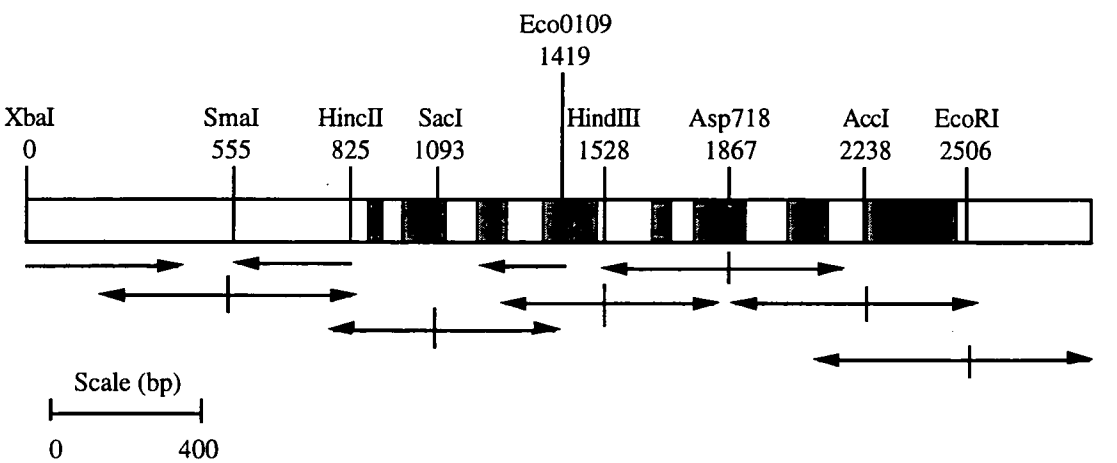


Fig.14. Schematic diagram indicating the approach used to sequence the λ10 13G6(i) genomic homologue in order to obtain the majority of the sequence from both DNA strands. Arrowed lines indicate the direction and extent of sequencing. Shaded blocks indicate regions equivalent to the cDNA clone. The restriction sites indicated are not necessarily unique.

B.oleracea. As a greater number of restriction sites were mapped in λ10, this was chosen to be sequenced. Initially two fragments (one SacI, one SacI/XbaI) were subcloned into pKS+Bluescript. Further subcloning and sequencing were then carried out on these clones as described in fig.14. A total of 2843bp were sequenced, encompassing the entire region that hybridises to the cDNA clone together with ~1kb of 5' and ~300bp of 3' sequence although the latter two regions were not sequenced entirely along both strands (fig.14).

The genomic sequence does not align directly with that of the cDNA as it contains seven large interruptions interspersed between regions equivalent to the cDNA (fig.14). To determine whether these could be functional introns they were compared to the splice site consensus structure for dicotyledonous plants (Brown, 1986; Goodall and Filipowicz, 1989; Goodall and Filipowicz, 1990). All of the interruptions show some identity to the consensus though none show an exact match (fig.15). The invariant nucleotides and the pyrimidine-rich 3' tract are, however, present in all the interrupting sequences. Therefore it is most likely that these are all spliceable inserts, such that an unbroken ORF could be derived from this sequence. Confirmation of splicing could be obtained by ligation of fragments of the cDNA-homologous region of the genomic clone to a suitable promoter and translation in vitro. Splicing could be assayed by determination of primary transcript size and comparison with the mRNA size via Northern analysis.

As well as the introns, other differences constitute 4.4% of the overlapping sequences of the 13G6(i) cDNA and genomic clones. These consist of small (<7bp)insertions/deletions and substitutions and are catalogued in table 5. The effect on the putative protein sequence is to produce five substitutions and to remove three amino acid residues from the C-terminal end. The substitutions are all by conservative amino acids with similar properties so significant effects on the protein's structure/function are unlikely.

consensus aag:	|:GTGag.....(pyrimidine rich)TAG:|	:gt	AT>59%	Size(>70bp)
	A C			
895-980 CCA:	|:GTATC.....GTTTTCTTGAATAG:|	:GT	63%	86bp
1106-1179 ACA:	|:GTATA.....ATGAGAAATCTGGTTTACAG:|	:AC	65%	74bp
1245-1323 GAT:	|:GTAAG..... ....AGAGAGTTTTGTTTCAG:|	:GT	62%	79bp
1536-1624 AAG:	|:GTAAT.....CACTTTTAGAATCCATGCAG:|	:GC	63%	89bp
1673-1761 AAG:	|:GTGAA.....TTTTTGTTTAATGAAACAG:|	:TA	65%	89bp
1870-1967 CAG:	|:GTACC.....TGGTCTTAAATGTAATTCTCAG:|	:AT	69%	98bp
2121-2231 ACA:	|:GTAAG.....TTGTTTGGTGGTTAAAATGACAG:|	:GT	68%	110bp

Fig.15. Comparison of the putative intron sequences in the genomic 13G6(i) clone with the consensus sequence/structure. Invariant residues in the consensus sequence (Brown, 1986) are indicated by capital letters. The percentage of AT pairs which is a defining factor in dicot intron splicing (Goodall and Filipowicz, 1989) and intron size (Goodall and Filipowicz, 1990) are also indicated.

cDNA clone (pos'n) →genomic clone (pos'n)	Effect on protein sequence
C(1)→T(887)	none
T(145)→C(1181)	none
TA(160/161)→AT(1196/1197)	SN→SY
T(214)→C(1328)	none
G(217)→A(1331)	none
TT**C**TATGAA(370)→TTTATG**G**AA(1524)	FYE→FME
A(484)→G(1518)	I→V
C(486)→G(1520)	I→V
T(527)→C(1654)	none
A(533)→G(1660)	none
C(589)→G(1804)	none
C(633)→T(2020)	A→V
A(780)→G(282)	N/C
G(787)→A(289)	N/C
ATGA(790)→AT**GGTGAAA**GA(2282)	N/C
G(799)→T(2308)	N/C
T(843)→G(2349)	N/C
A(852)→T(2358)	N/C
T(856)→A(2362)	N/C
TG**GTT**CA(875)→TGCA(2381)	N/C
C(887)→T(2390)	N/C
G(901)→A(2404)	N/C
C(911)→T(2411)	N/C
G(914)→T(2414)	N/C
C(916)→A(2416)	N/C
A(925)→G(2432)	N/C

Table 5. Small scale differences between the cDNA and genomic 13G6(i) sequences and their effect on the putative protein sequences. Numbers indicate the first nucleotide in the altered region. Inserted sequences are denoted by bold text. Where the change is in a non-coding region then this is denoted by N/C in the effect on protein sequence in denoted by N/C.

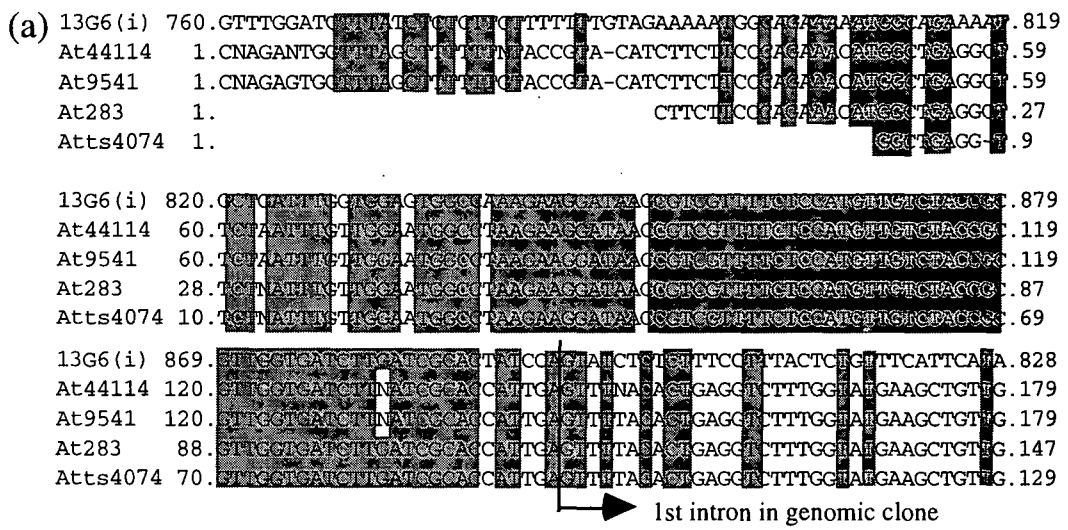
The distribution of these differences is not even: the coding regions of the cDNA and genomic clone show 97.9% nucleotide sequence identity whereas the two 3' non-coding regions share only 87.1% identity. The difference in conservation between the two parts of the sequence can be explained by greater pressure for sequence conservation on the coding part of the expressed sequence so that these differences are unlikely to be due to random mutation during cloning and sequencing. The genomic clone may therefore represent a different gene or be the same gene that has mutated between *B.oleracea* varieties
(cDNA is from *gemifera*, the genomic clone from *alboglabra*). The latter is unlikely as the high level of difference between the non-coding regions is greater than seen between the 3' regions of 13G6(i) and the *Arabidopsis* EST At31012, so the genomic clone probably represents a different gene to that represented by the cDNA clone.

4.1.1.6 Analysis of the 5' region of the genomic clone

The 5' region of the genomic clone was analysed to determine if an entire coding sequence could be determined. One problem with the analysis was that, as intron sequences were present throughout the sequence equivalent to the cDNA, it was possible that the genomic sequence equivalent to the missing 5'

portion of the cDNA was also interrupted. However by comparison with the *Arabidopsis* EST sequences some sequence could be putatively assigned (fig.16) This region of the genomic sequence contains an ATG codon which has high homology to the Lutcke consensus (AAAAATGGC against AACAATGGC). The one difference between the two sequences (C at position 3) is the second (21% C vs. 61% for A which is found in the *A.thaliana* clones) most common nucleotide in its position, and so this ATG has a high probability of being the start codon in the transcribed mRNA.

Further evidence that this is the start codon comes from a 1bp insertion 18 bases upstream from the ATG in the genomic clone when compared to the two *A.thaliana* clones. This would completely disrupt translation, via a frameshift, if it occurred within the coding region of the expressed cDNA. Furthermore 5' to the putative initiation codon the level of identity between 13G6(i) and the *Arabidopsis* sequences is less than that seen 3' to the ATG. This could be due to the presence of another intron in the genomic clone although no 3' consensus site can be identified and no sequence similar to the EST's is present in the remaining 800 upstream bp of the genomic clone. Therefore the full nucleic acid and predicted protein sequences, of the genomic 13G6(i) clone, are as described in fig.17.



(b)

782. TTTTTTTGTAGAAAAATGGCAGAAAATGCTGATTTGGTGGAGTGGCCAAAGAAGGATAAG .841

M A E N A D L V E W P K K D K

842. CGTCGTTTTCTCCATGTTGTCTACCGCGTTGGTGATCTTGATCGCACTATCCA⇒1st intron.894

R R F L H V V Y R V G D L D R T I

Fig 16. Analysis of the promoter region of the genomic clone homologous to 13G6(i). (a) Alignment of the region 5' to the first intron of the genomic clone with three *Arabidopsis* EST's which overlap with the 5' region of the 13G6(i) cDNA clone. Shading indicates positions where the EST's have nucleotides identical to those in the genomic clone. (b) predicted 5' region of the 13G6(i) cDNA together with the putative encoded N-terminal protein. The start codon consensus region is underlined.

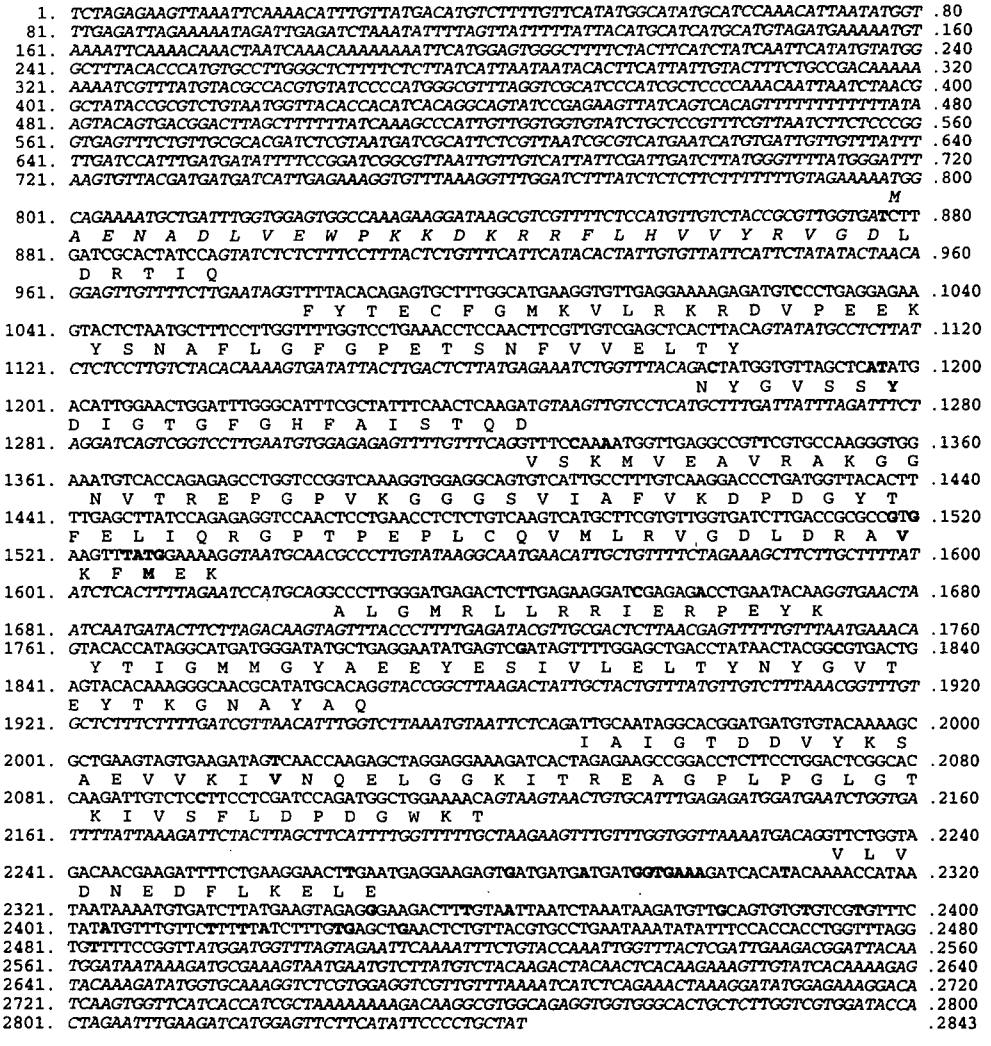


Fig.17. Sequence of a genomic clone homologous to the 13G6(i) cDNA clone. The DNA sequence is shown together with the predicted protein encoded by this clone. Sequence of the genomic clone (nucleic acid and protein) not present in the cDNA are indicated by italic type Point mutations and small insertions/deletions between the cDNA and genomic clones (nucleic acid and protein) are indicated by bold text.

The region of the genomic clone, 5' to this predicted cDNA sequence, was analysed for consensus promoter sequences. This was complicated by the fact that the initiation start site was not determined, as the high variability between plant initiation sites and their distance from the translation initiation site means that they cannot be directly screened for (Joshi, 1987). Similarly no consensus TATA box sequence (Zhu et al., 1995) could be identified. However, this consensus sequence has been implicated as a basic promoter element, its absence has been reported in other expressed plant sequences (Joshi, 1987) so does not indicate an inactive gene. Screening with the program MacSignalscan similarly failed to detect any previously identified promoter/enhancer motifs.

One explanation for the lack of motifs could be that the promoter region lies 5' to the sequenced regions of the clone. An intron in the genomic region that codes for the 5'UTR of the cDNA could be

present but as the identified introns are all small (consistent with the sizes of other plant introns as described by Goodall and Filipowicz (1990)) such an interruption would be unlikely to constitute the majority of the genomic sequence 5' to the predicted translation start.

4.1.1.7 Expression of 13G6(i) in the mature flower.

Scutt (1990) reported strong expression of 13G6(i)-like transcripts in the stigma while weaker expression occurred in the style and immature anther. To provide a more extensive analysis, it was decided to study the expression of 13G6(i) within the whole flower. Tissues from greenhouse-grown plants were initially assayed. Transcripts similar to 13G6(i) were detected in all the floral whorls but not in 4 week old seedling (fig.18). In contrast to the previous data obtained by Scutt (1990), roughly equivalent levels of expression in the stigma, style and ovary were detected (fig.19). From these blots the size of the transcript was estimated to be 1200nt, 100nt smaller than originally proposed. The lack of expression in seedling suggested that 13G6(i)-like genes have different patterns of expression between the *B.oleracea* and *O.sativa* given the isolation of 13G6(i)-like EST's from vegetive tissue in the latter.

Although these Northern blots allowed the description of the expression of 13G6(i) in the flower, they were labour intensive to produce as the relatively small size of Brassica flowers meant that a large number had to be harvested to obtain suitable quantities of RNA. Therefore the use of 'squash blots' was assessed to see if they could provide a means by which expression across the flower and during development could be investigated without such extensive tissue harvesting. This technique, whereby nucleic acid is transferred onto nylon/nitrocellulose by directly squashing tissue onto the membrane, in the presence of a vacuum, has been previously reported as a rapid method for assessing the expression of different S-alleles in the stigma (Cappadocia et al., 1993). Squashing of various floral organs onto nitrocellulose and hybridising to a cDNA probe gave the result illustrated in fig.20. From this result two problems with the technique were identified:

(i) The process of squashing causes the carpel and stamen to lose their shape so that identification of their different organs was not possible.

(ii)Transfer is variable between different organs as illustrated by the petal and sepal which gave no hybridisation signal after squash-blotting. This contrasted with the Northern blot result which suggested that these organs express 13G6(i) at a similar level to the carpel and anther.

Due to these problems it was decided that squash-blotting was not a suitable method for the localisation or semi-quantitative analysis of the expression of RNA transcripts in the flower. Therefore all other work concentrated on the use of Northern blots.

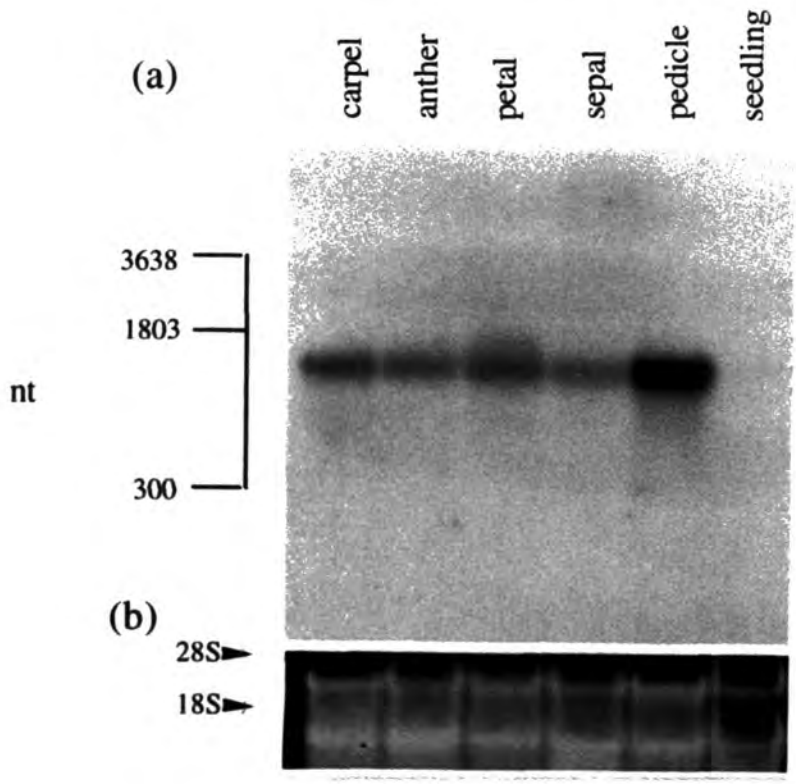
4.1.1.8 13G6(i) expression is influenced by stress

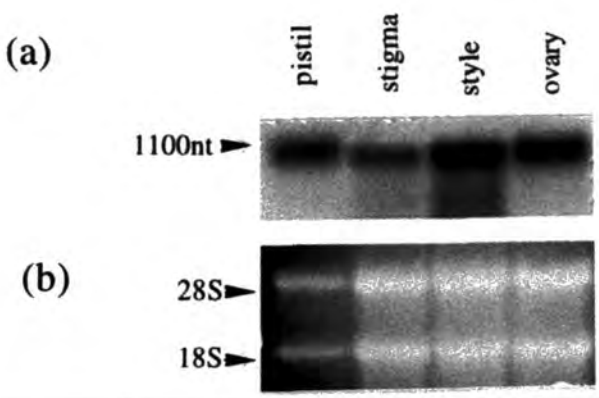
The global pattern of expression in the flower suggests that 13G6(i) does not provide a specific function in the stigma. To further assess this, the developmental pattern of expression in the carpel was

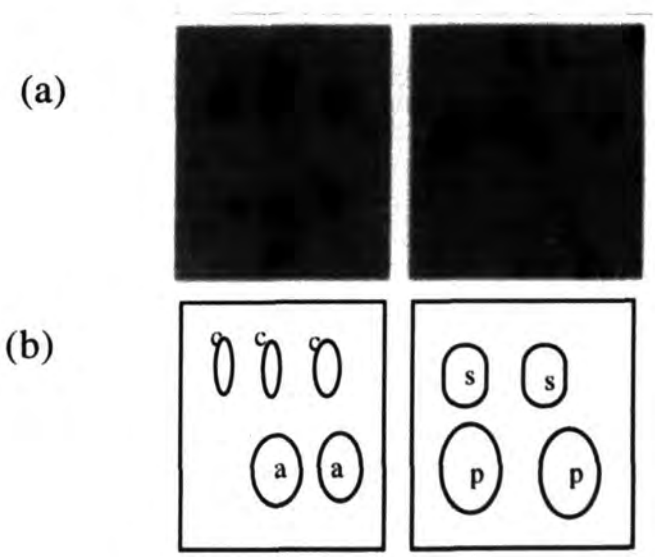
Fig18. Expression of 13G6(i)-like transcripts in the flowers of *B.oleracea* var *alboglabra*. (a) Autoradiograph of a blot of RNA from the major floral whorls (20ug per lane) after hybridisation with a 13G6(i) cDNA probe. The blot was washed in 1xSSC, 0.1%(w/v)SDS at 65°c. (b)ethidium bromide stain of the RNA to show rRNA on the same blot to confirm equal loading of total RNA. Transcript lengths were estimated from the migration relative to the rRNA species on the blot. Seedling RNA was derived from whole 3 week old plants.

Fig.19. Expression of 13G6(i)-like transcripts in the carpel (a) autoradiograph of a blot of RNA from the principal parts of the pistil (20ug per lane) after hybridisation with a 13G6(i) cDNA probe. The blot was washed as in fig.18. (b) ethidium bromide stain of the RNA on the blot in (c) to show rRNA so as to confirm equal loading. 20ug of RNA was loaded in each lane. Transcript lengths were estimated from the migration relative to the rRNA species on the blot.

Fig.20. Assessment of the squash blot technique for studying the expression of 13G6(I), and other cDNA's, in the flowers of *B.oleracea*. (a) Autoradiograph of squash blot after hybridisation to a 13G6(i) cDNA probe. The blot was washed as in fig.18. (b) Map indicating the approximate positions of squashed organs. Abbreviations are c=carpel, a=anther, p=petal, s=sepal.


(a)
carpel
anther
petal
sepal
pedicle
seedling
3638
1803
nt
300
(b)
28S
18S


(a)
pistil
stigma
style
ovary
1100nt
(b)
28S
18S


(a)
(b)
a
a
s
s
p
p

analysed. Two differing results were obtained from two separates batches of RNA samples, harvested at different times during the PhD project, from greenhouse-grown plants: Expression levels appeared to be roughly equivalent, throughout development or, expression was lower in the earlier stages of development.

This discrepancy suggested that expression in the carpel is subject to regulation other than imposed by development. Similarly the differences between the data presented here and those of Scutt could be due to external effects on expression. One factor could be stress: the greenhouse plants were subject to a variety of temperatures and lighting regimes, due to the non-controlled nature of the facilities and were also open to insect and fungal attack. The detailed condition of the plants, from which the two pools of RNA were isolated, was not recorded. Therefore the possible stress induction of 13G6(i) genes was investigated experimentally.

Initially physical wounding of greenhouse-grown plants, as described in section 3.2.2, was carried out. At low stringency (1xSSC, 0.1%SDS, 65°C) this increased expression of 13G6(i)-homologous transcripts in the stem to levels similar to that seen in floral organs, while wounding of leaves elicited a weaker response (fig.21a,b). No change in expression was apparent in wounded carpels, although changes in expression in the stigma or style could be masked by their small physical size relative to the ovary.

The wound-induced transcripts could also be detected at high stringency (0.2xSSC, 0.1%(w/v)SDS, 68°C), suggesting that the protein products of the floral and wound-induced classes of transcript are identical or nearly identical and so may share a similar function.

Further investigations were carried out to investigate the effect of cold, dark and drought as stress inducers of 13G6(i) in growth room-grown young plants as effects by these factors could not be ruled out under greenhouse conditions. Lack of heat and water did not affect 13G6(i) expression in RNA extracted from mixtures of leaf and stem tissue in these experiments but lack of light did induce 13G6(i) transcripts (fig.21c,d). This correlates with the homologous *Arabidopsis* EST isolated by Desprez et al. (1994) which was derived from etiolated seedlings so suggesting that expression of 13G6(i) homologous transcripts is a general response to incubation in the dark. The 13G6(i) transcript only appeared after 96 hours in the dark suggesting that a prolonged lack of light is required for transcription, ruling out the possibility that 13G6(i) might be a circadian gene.

Wounding and light-starvation cannot, however, explain the inconsistent patterns of expression recorded for 13G6(i) in flowers. None of the plants were exposed to continuous periods of darkness and only healthy undamaged buds were used for tissue isolation. Therefore some other factors must have caused the observed differences. The plants were not deliberately exposed to pathogens and insects, however these pests had access to the plants and were found in the greenhouse used. While floral shoots displaying symptoms of such attacks were not used these stresses could act via acquired resistance mechanisms whereby unaffected parts of a plant activate defence genes in response to signals generated by other parts of the plant under attack (Chester, 1933). Unfortunately no facilities for directly exposing plants to insect or pathogen attack were available, so that experiments on these factors could not be carried out.

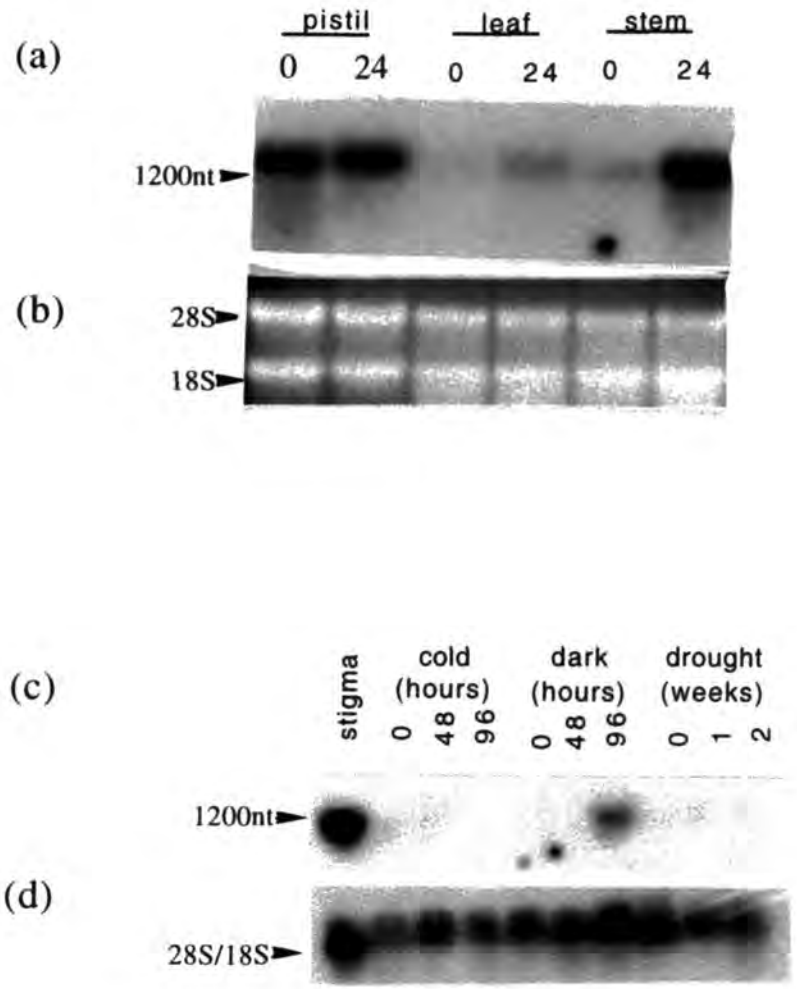


Fig.21. Expression of transcripts homologous to 13G6(i) under stress conditions in *B.oleracea*. (a) Autoradiograph showing the effect of wounding on floral and vegetative tissues. Numbers indicate hours after wounding. The blot was washed to 1xSSC, 0.1%(w/v)SDS which gave a stringency of 76%. An identical pattern was obtained by rewashing the blot at 0.2xSSC, 0.1(w/v)%SDS,65°C equivalent to 90% stringency (data not shown). (b) ethidium bromide stain of the RNA on the blot in (a) to show the rRNA's in order to show equal loading of all lanes (20ug were loaded per lane). (c) Autoradiograph showing expression of 13G6(i)-like transcripts in seedlings under various stresses.The blot was washed in similar conditions to that in (a). (d) Autoradiograph of the blot in (c) hybridised to the pHA1 ribosomal DNA probe to show equal loading of lanes (10ug per lane).

These experiments show that some stress treatments can induce 13G6(i) expression in vegetative tissue, although in untreated greenhouse-grown plants, significant levels of expression were not detected. However, in a similar manner to Wemmer's model of carpel protection, the entire flower may be more sensitive to stress as the plant protects its reproductive capacity. Therefore the observed mode of floral expression described in section 4.1.1.7 may be induced as part of a hypersensitive stress response rather than being constitutive. Growth in the controlled growth facility is believed to minimise the effect of stress factors, as temperature, light and humidity are standardised and, as the facility is sealed, pathogens should not have access. Therefore the expression of 13G6(i)-like transcripts was assessed in plants grown in these facilities

4.1.1.9 Expression of 13G6(i) in plants grown in controlled growth conditions

The expression of 13G6(i)-homologous transcripts was assayed in flowers from plants grown under controlled conditions. A similar pattern of expression in the floral organs was observed to that seen in greenhouse grown plants (fig.22a,b), although vegetative expression was also detected in the stem at an equivalent level to that seen in the flower. Expression is high in the reproductive organs, though weaker in the sepal. Carpel RNA from plants grown in this facility give the developmental pattern of expression shown in fig.22c,d. Expression is weaker during early stages of development, when compared to later stages. The greater reliability of the tissue used for these experiments leads us to believe that this represents the uninfluenced developmental pattern of expression.

However growth in a controlled environment cannot guarantee that the plants are stress free. *Brassica* produce several floral meristems and multiple flowers per meristem. It is impractical to use single flowers from each plant or to strip whole plants in a single experiment. Therefore some flower buds may be induced to express stress genes by systemic (acquired) resistance, caused by the harvesting, a process which is technically wounding the plant.

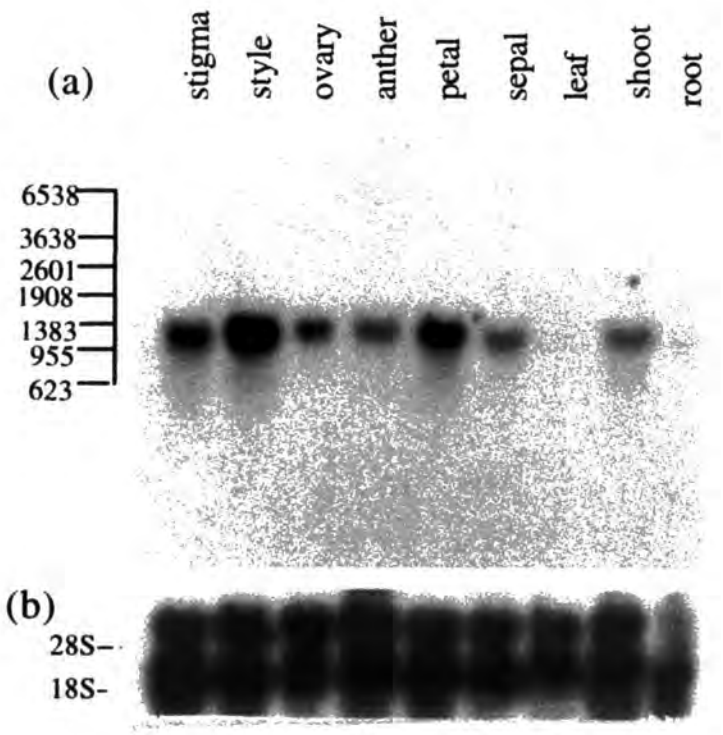
Therefore as a final experiment to determine the possibility of stress-induced expression the effect of bud picking was assayed. Half of the buds (early stages of development) on an inflorescence on a previously untouched plant were removed and the mature buds rapidly frozen in liquid nitrogen for RNA extraction. The mature buds on the plant after one week were then harvested and RNA extracted. Comparison of the level of 13G6(i) expression in each pool of buds indicates that some increase in expression in the post-picking buds was apparent although significant expression was already evident in the pre-picking buds (fig.23a,g). Therefore the action of harvesting buds influences the expression of 13G6(i)-like transcripts. This could be due to upregulation of the transcript over the entire flower or by activation in some organs of the flower that do not normally express 13g6(i) or express it at a relatively low level. Due to the difficulty in obtaining pre-picked tissue these models were not investigated further.

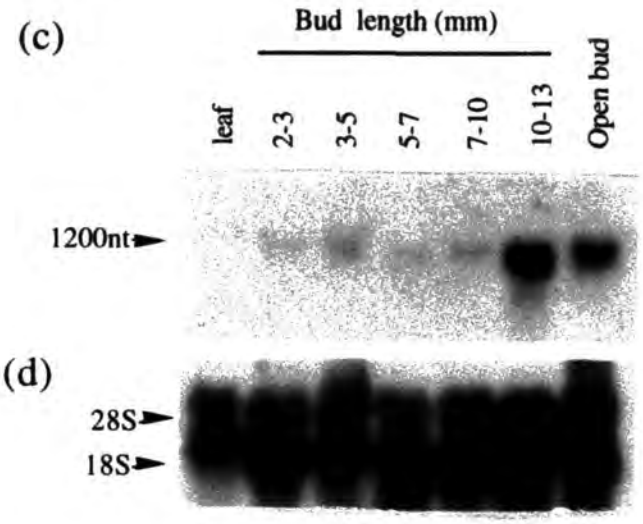
4.1.1.10 Role of 13G6(i).

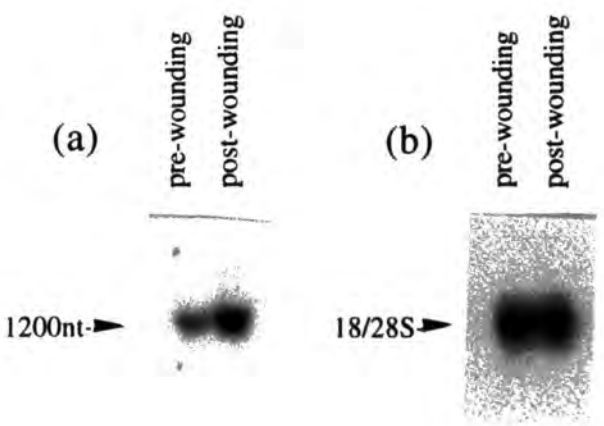
The data presented here provides detailed description of the sequence of 13G6(i), its homologues, the genomic organisation in plants and expression in *Brassica oleracea*. Unfortunately, none of this

Fig.22. Expression of 13G6(i)-like transcripts in the flowers of growth room-grown plants. (a)Autoradiograph showing pattern of expression in the mature floral and vegetative organs (10μg RNA per lane). (c) Autoradiograph showing expression during carpel development (20μg RNA per lane). These two blots were washed at 65°C in 1xSSC, 0.1%(w/v)SDS (stringency ~76%). In each case a similar pattern was obtained at higher (0.2xSSC, 0.1%(w/v)SDS) stringency. (b) and (d) are the autoradiographs of the blots in (a and (c) after hybridisation to a pHA1 probe to demonstrate equivalent loading of total RNA's.

Fig.23. The effect of bud-picking on the expression of 13G6(i) in the flower. (a) Autoradiograph showing expression in floral buds pre- and post- bud picking (10μg RNA per lane). (b) blot in (a) reprobed with pHA1.


(a)
stigma
style
ovary
anther
petal
sepal
leaf
shoot
root
6538
3638
2601
1908
1383
955
623
(b)
28S–
18S-


(c)
Bud length (mm)
leaf
2-3
3-5
5-7
7-10
10-13
Open bud
1200nt
(d)
28S
18S


(a)
pre-wounding
post-wounding
1200nt-
(b)
pre-wounding
post-wounding
18/28S

information provides a clear indication of the activity/function of 13G6(i). The sequence information for 13G6(i) suggests that it is one member of a highly conserved family of sequences found throughout the angiosperms and probably also in animals and bacteria. While the protein family members in the latter consist of only one repeat unit compared to the tandem duplication structure seen in *B.oleracea* the high homology between the protein primary sequences from the two species suggests that they are related. This homology would suggest a conserved functional role for the proteins. The presence of genes in bacterial and, probably, human cells predicts a non-cell specific, possibly fundamental, housekeeping role.

Such a role, however, is contradictory to the observed expression of 13G6(i) in *B.oleracea*, where organ-specific expression occurs, implying that it is not essential for cell function in this organism. Also the *Brassica* transcript is not detected in roots or whole seedlings, as the *Oryza* transcript is, suggesting that the pattern of expression is not conserved amongst the angiosperms. Therefore the function of the 13G6(i) protein may be context-specific, playing different roles in different cells. The discussion of potential roles for 13G6(i) will therefore concentrate on *Brassica.*

Much of the problem in studying 13G6(i) has come from its varied expression. Three patterns of expression can be described: floral, constitutive-vegetative (stem) and stress-induced-vegetative (leaf and stem). The latter two modes of expression have been easy to determine, although an exhaustive analysis of the latter has not been carried out. This pattern, however, is difficult to explain: why 13G6(i) should be constitutively expressed in stem but stress-induced in leaf is not known.

13G6(i) was originally isolated due to its floral expression which has been the most difficult to accurately define. While it is clear that 13G6(i) represents an abundantly expressed floral sequence, the floral expression is affected by the harvesting of buds from the plant, although the precise effect of this could not be determined. The effect of inducing factors on the flower, such as possible release of ethylene from nearby plants affecting 'virgin' plants, cannot be discounted either. However the relevance of these factors *in vivo* has to be considered. If 13G6(i) expression is induced by an extremely sensitive stress-response mechanism, then any plant growing 'in the wild' would be expected to have this mechanism triggered. Therefore when considering the biological function of 13G6(i) its highest level of expression in the flower could be regarded as constitutive.

The pattern of expression seen for 13G6(i) can be used to draw parallels with the solanaceous stress-gene family sequences that have been found in the stigma and style. As with the solanaceous sequences roles in floral defence or physiology can be proposed for the protein encoded by 13G6(i). Such roles in relation to the biology of the female tissue and the relevance of 13G6(i) to the Solanaceous sequences is discussed in section 5.2.

Of note is the fact that only two genomic sequences cross-hybridise with the 13G6(i) cDNA despite the complex pattern of expression of cross-hybridising RNA transcripts. This means that one or both of the genomic copies must have a promoter that is multifunctional in that it can direct tissue-specific and/or stress-induced expression in a variety of organs. If the sequence represented by the genomic clone is non-functional, as could be interpreted from the lack of detectable promoter sequences in the 5' region, then only one gene must be able to direct the observed patterns of expression. Such a complex

promoter could occur as other promoters which can direct both floral-organ and stress-responsive expression have been reported such as the *Solanum* PR10a (Constabel and Brisson, 1995), *Oryza* chitinase (Zhu et al., 1993) and *Phaseolus* HRGP promoters (Wycoff et al., 1990).

4.1.1.11 Proposals for the further study of 13G6(i)

Data on the sequence and pattern of expression provide insufficient information on which to formulate a model for the role of 13G6(i) in the flower. The only definitive way to determine the role of 13G6(i) in the *Brassica* flower is to determine its biochemical function/activity. The former could be analysed by reducing/eliminating expression of the transcript via the introduction of antisense constructs into transgenic plants (Goring et al., 1991). Although *Brassica* species can be transformed (Fry et al., 1987), a simpler system would be provided by the closely related crucifer *Arabidopsis*. However a similar pattern of expression would have to be first demonstrated in this plant (which may have a different pattern of expression as it appears to only have one gene copy), while analysis of the flower may be more difficult to achieve due to the small size of the plant. Functional analysis, however, may not provide any data on 13G6(i): If the gene is essential for development of the plant no tissue may be generated to analyse, whereas if the gene is involved in a defensive capacity, particularly in the flower, an experimental approach to studying the effect of elimination of expression is not obvious.

An alternative approach to functional analysis would be to use bacterial or lower eukaryote models, via gene deactivation by homologous recombination, and try to extrapolate the data to plants. This approach would be technically simpler than transformation of *Brassica* or *Arabidopsis* and there would be no problems with obtaining sufficient quantities of tissue for characterisation. The use of a bacterial system would also allow analysis of function even if the 13G6(i) bacterial equivalent is essential for growth. Expression could be controlled by co-transforming with the coding region of the 13G6(i) homologue coupled to an inducible promoter such as *lacZ* (Sambrook et al., 1989) which would allow expression to be maintained while the bacterial transformants are obtained and cultured for analysis. Then removal of inducer could be used to deactivate expression when the gene's function was to be assessed.

There are, however, disadvantages in using this approach. Firstly expression of 13G6(i) homologues would first have to demonstrated in these organisms (such expression has not been determined in *Vibrio* and *Haemophilius*). Secondly, any data generated by this approach may not be directly applicable to *Brassica* or any other higher plants as the bacterial monomeric form may have a different function to the angiosperm tandem-duplication form. The potential for context specific function, even between different tissues of the same plant, may also prevent application of any data generated by this approach. The biological/biochemical activity of the proteins is more difficult to study. Some hint about activity could be provided by the functional analysis. However if this is not so then characterisation of the activity of this class of proteins may only be determined by classical methods whereby work on isolating the gene/protein involved with a certain activity uncovers a 13G6(i) homologue.

Whatever the role of 13G6(i) it is clear that it is not restricted in expression to the reproductive organs and therefore, even if a context-specific function is considered probably does not play a direct role in pollination.

4.1.2 The clone 15H11.

The clone 15H11 was completely sequenced previously (Scutt, 1990). The nucleic acid sequence is rich in guanidine (49%) and was proposed to encode a putative glycine-rich, cell-wall protein (Scutt, 1990) due to its homology to previously characterised sequences (fig.24). The presence of such proteins in plants has been demonstrated in a variety of tissues (Cassab and Varner, 1988) by a high glycine content in the cell walls. In relation to the current study, acid hydrolysis of a protein homogenate from the stigma of *B.campestris* indicated that glycine was one of the three major amino acids occurring in the stigma, together with serine and glutamine (Knox, 1984).

As the pattern of expression for the clone 13G6(i) had been shown to be more complex than that predicted previously by Scutt (1990), more extensive expression studies were also carried out on 15H11 to confirm the pattern reported by Scutt. As the number of reported GRP sequences had also increased, since 15H11 was initially reported, additional analysis of the relationship between 15H11 and such sequences was also undertaken

4.1.2.1 Expression of 15H11 in the mature plant.

Northern analysis identified transcripts homologous to 15H11, of estimated length ~1300nt, in the stigma but not in the style, immature anther or leaf (Scutt, 1990), suggesting a specific role in stigma function. To expand on these results a more thorough analysis of expression was carried out using the cDNA as a probe against a Northern blot containing RNA from all the major organs of the mature plant. The blot was initially washed in 1xSSC, 0.2%(w/v) SDS at 65_C which gives a predicted stringency of 73%. However the 15H11 probe is so GC-rich (65% over coding region) that the equations used to calculate stringency may not be accurate (Sambrook et al., 1989). Washing under these conditions gave the pattern of expression in fig.25a. As with the work of Scutt, the strongest expression is seen in the stigma. The main transcript in this tissue is estimated to be 1200nt long, smaller than proposed previously (Scutt, 1990), but similar enough that the difference could be attributed to experimental error. Hybridisation also occurs to multiple, higher molecular weight stigma transcripts which form a broad smear. This smear may be due to non-specific hybridisation to other GC-rich sequences as has been suggested for similar patterns obtained after probing blots with extensin sequence probes (Bown et al., 1993), or may be an artefact of electrophoresis which is visible due to the strong exposure of this blot.

```
  1. GGAGGAGCTGGTGGAGGTTTTGGCGGTGGTGCAGGAGGAGGCCATGCTGGCGGTGGAGCTGGTGGAGGTCACGGTGGTGG .80
       G  G  A  G  G  G  F  G  G  G  A  G  G  G  H  A  G  G  G  A  G  G  G  H  G  G  G
 81. TGCTGGAGGAGGATTTGGCGGTGGAGCTGGTGGAGGTCACGGTGGTGGTGCTGGAGGAGGATTCGGCGGTGGTGCTGGTG .160
       A  G  G  G  F  G  G  G  A  G  G  G  H  G  G  G  A  G  G  G  F  G  G  G  A  G
161. GAGGTAAAGGTGGTGGTCTTGGTGGTGGTGGTGGCGTTGGAGGAGGCCACGGTGGTGGTGTTGGAGGAGGGTTCGGTGGT .240
       G  G  K  G  G  G  L  G  G  G  G  G  V  G  G  G  H  G  G  G  V  G  G  G  F  G  G
241. GGTGCTGGTGGAGGCAAAGGTGGTGGTGTTGGAGGAGGGTTCGGTGGTGGTGCTGGTGGAGGCAAAGGTGGTGGTTTAGG .321
       G  A  G  G  G  K  G  G  G  V  G  G  G  F  G  G  G  A  G  G  G  K  G  G  G  L  G
321. TGGAGGTCACGGTGGTGGTGTTGGAGGAGGATTTGGTGGTGGTGCTGGTGGAGGCAAAGGTGGTGGCGTCGGAGGAGGCC .400
       G  G  H  G  G  G  V  G  G  G  F  G  G  G  A  G  G  G  K  G  G  G  V  G  G  G
401. ATGGTGGTGGTGCTGGTGGAGGGTTCGGTGGTGGTGCTGGTGGAGGACATGGTGGTGGTGCTGGAGGAGGATTTGGCGGT .480
     H  G  G  G  A  G  G  G  F  G  G  G  A  G  G  G  H  G  G  G  A  G  G  G  F  G  G
481. GGAGCTGGTGGAGGTCACGGTGGTGGTGCTGGAGGAGGATTCGGCGGTGGTGCTGGTGGAGGCCACGGTGGAGGAGCTGG .540
       G  A  G  G  G  H  G  G  G  A  G  G  G  F  G  G  G  A  G  G  G  H  G  G  G  A  G
561. TGGAGGTTTTGGCGGTGGTGCAGGAGGAGGCCATGCTGGCGGTGGTGGTGCTGGTGGAGGGTTTGGTGGTGGTGCCGGAG .620
       G  G  F  G  G  G  A  G  G  G  H  A  G  G  G  G  A  G  G  G  F  G  G  G  A  G
641. GAGGCGGTGGTGGTGGATTTTGAATCTAGCTTTTTAGGGCTACTTACTTGTGTACCACGCCAAAAAAGAATGAGATTATA .660
     G  G  G  G  G  G  F
721. AATGCTGGGATTTTATTATGAATGAAATAAATGATACTTAACAAAAAAAAAAAAAA                        .776
```

Fig.24. Sequence of the cDNA clone 15H11 and its putative encoded protein as described by C.Scutt, 1989. The possible AATAAA polyadenylation signal motif is underlined.

As well as the strong expression in the stigma, bands hybridising to the cDNA are seen in other tissues: In the style a transcript, of the same size as the main stigma species, hybridises with the 15H11 cDNA. Expression in this organ was not detected previously although earlier autographs were developed after a lower level of exposure, which may have been insufficient to detect the transcript. However, due to the small physical size of the two organs some cross-contamination of the style RNA with stigma mRNA cannot be ruled out.

In the ovary, petal, sepal and stem tissue longer transcripts (~1650nt, ~1650nt, ~1400nt and ~1400nt respectively) occur. The hybridisation signal with stigma RNA is considerably higher than in any of these other tissues while the style and petal RNA hybridise more strongly than the ovary, anther, sepal and stem. The weaker hybridisation in these latter six tissues could be due to low expression of the relevant transcript(s) or once again non-specific hybridisation of the 15H11 probe to other GC-rich sequences. In order to differentiate between these possibilities, the blot was washed at a higher (85%) predicted stringency. The signal in the petal is greatly reduced when compared to that of the stigma and style (fig.25b), suggesting that the transcript in this tissue is not identical to 15H11. The signals in the ovary, anther, sepal and stem disappear suggesting that these too may be related, but not identical, sequences, although due to the generally lower level of signal on this autoradiograph these transcripts may still be hybridising but may not be detected.

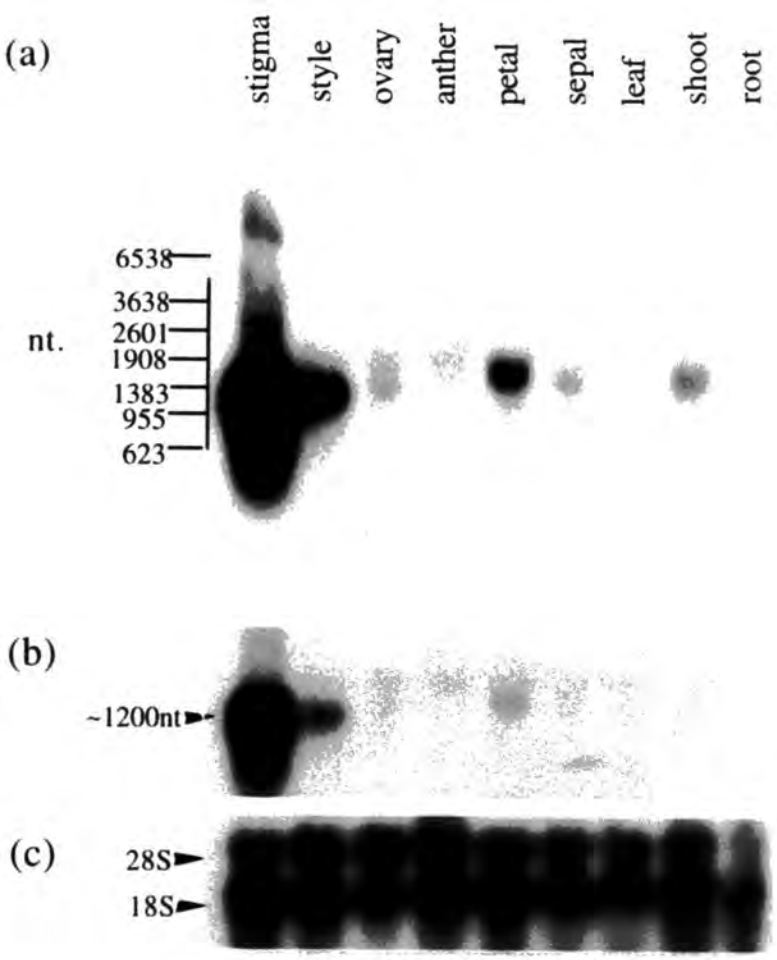
A much weaker hybridising band is also observed in all the tissue except the stigma where its possible presence is masked by the high signal from the major transcript smear. This band could be due to hybridisation to a GRP transcript, though the relative intensity of the signal is so weak when compared to the major hybridising species of the stigma, style and petal that expression in these other tissues must be negligible compared with the three strongly expressing organs. The transcript length does however correspond to that of plant 18S rRNA so the observed signal may be due to non-specific binding (Sambrook et al., 1989).

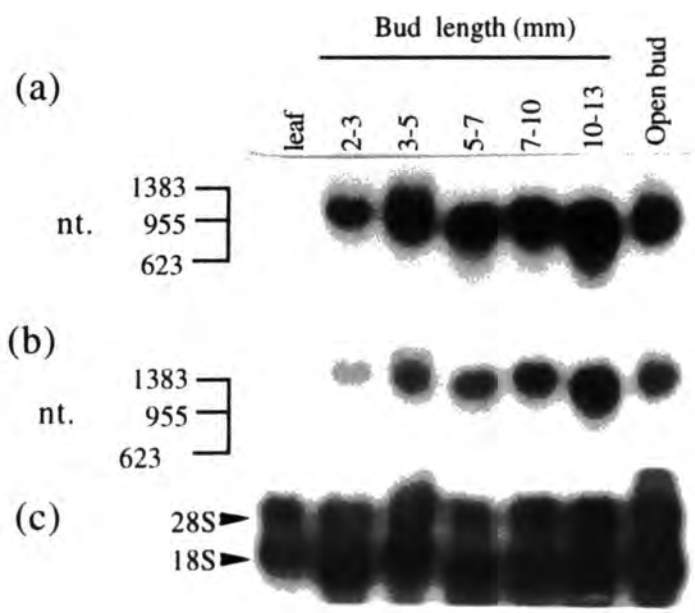
The developmental expression of 15H11 was studied in whole carpels. At low stringency, transcripts, homologous to 15H11, can be detected in all stages studied at approximately the same level of expression (fig.26a). The size of the transcripts appears to be irregular but mirrors the relative migrations of the rRNA's (fig.26c), so is assumed to be an artefact of electrophoresis.

At high stringency the hybridisation to the earlier developmental stages decreases relative to the late stages (fig.26b). This could be due to one of two factors. The change in expression could be due to different developmental patterns of expression in different sub-organs of the carpel (fig.27a), or the observed pattern could be due to different expression patterns of homologous, but different, sequences during development (fig.27b). One approach to differentiating between these two models would be to study expression in the stigma and style at different stages of development. This was not done as the small size of the organs during early development made harvesting for RNA extraction impractical. Micro extraction of RNA and PCR was also considered. This was also impractical, as the organ size means that contamination of stigma with style tissue and vice-versa could not be ruled out and so any results obtained by this approach would be open to suspicion.

Fig.25. Expression of 15H11-like transcripts in *B.oleracea*.as indicated by probing a Northern blot hybridised to a probe derived from the 15H11 cDNA (a) Autoradiograph of the blot washed to low stringency (1xSSC, 0.1%(w/v)SDS at 65°C). (b) Autoradiograph of the blot in (a) rewashed to high stringency (0.2xSSC, 0.1%(w/v)SDS at 65°C). (c) The blot in (a) and (b) reprobed with pHA1 to confirm equivalent loading of RNA (10μg per lane). The RNA sizes in (a) and (b) were derived from a Promega RNA size marker.

Fig.26. Northern blot of RNA from carpels, from *B.oleracea,* at different stages of development hybridised with a probe derived from 15H11. (a) autoradiograph of blot washed to low stringency (1xSSC, 0.1%(w/v)SDS at 65°C). (b) Autoradiograph of the blot in (a) rewashed to high stringency (0.2xSSC, 0.1%(w/v)SDS at 65°C). (c) Autoradiograph of the blot in (a) and (b) reprobed with pHA1 to confirm equivalent loading of RNA (20μg per lane). The RNA sizes in (a) and (b) were derived from a Promega RNA size marker.


(a)
stigma
style
ovary
anther
petal
sepal
leaf
shoot
root
nt.
6538
3638
2601
1908
1383
955
623
(b)
~1200nt
(c)
28S
18S


Bud length (mm)
(a)
leaf
2-3
3-5
5-7
7-10
10-13
Open bud
nt.
1383
955
623
(b)
nt.
1383
955
623
(c)
28S
18S

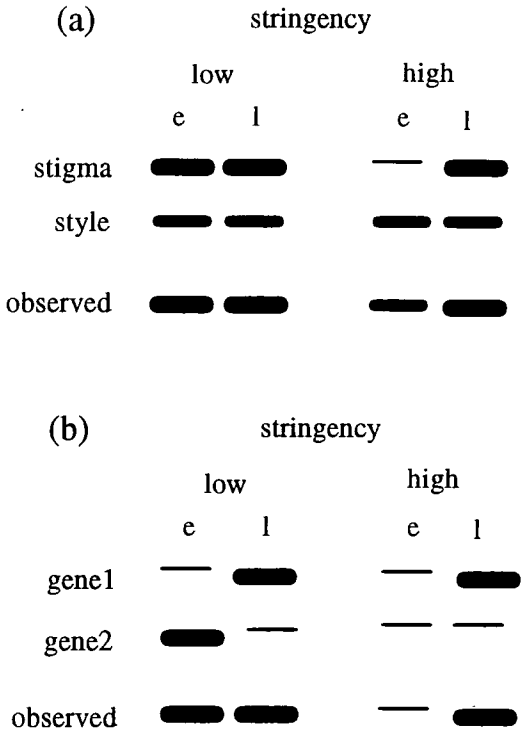


Fig.27. Models to describe the expression of 15H11-like sequences in the carpel during development. (a)differential expression of the same gene(s) between tissues. (b) differential expression of different genes during development in all tissues. The thickness of each bar indicates the level of expression: the thicker the bar the higher the expression. e=early stage of development, l=late stage of development.

4.1.2.2 15H11 is part of a complex gene family in *B.oleracea*.

The restricted mode of expression of 15H11 is in contrast with its genomic organisation. On Southern blots ~5-15 of bands per genomic digestion cross-hybridised with a 15H11 probe at both low (73%) and high (85%) predicted stringencies (fig.28), indicating that the 15H11 sequence is a member of a complex family of related sequences. Washing at high stringency eliminated several bands indicating that many of these sequences do not show complete homology to 15H11. These sequences probably contribute to the observed expression in the petal and may be expressed sequences restricted to the early stages of carpel development. The high stringency wash pattern of hybridisation was still complex suggesting multiple gene copies. It is possible that a number of these sequences may be simultaneously expressed in the stigma and style. Some of these sequences could also be pseudo-genes, may be expressed under circumstances not yet ascertained or may be the product of hybridisation to other GC-rich sequences unrelated to 15H11.

4.1.2.3 Expression of 15H11 under stress.

Many of the GRP genes, identified by other workers, have been found to have their expression regulated, both up and down, by various stress treatments. Wounding can induce *Phaseolus* (Keller et al., 1989), *Petunia* (Condit and Meagher, 1987), *Hordeum* (Rohde et al., 1990), and *Lycopersicon* (Showalter et al., 1991. Salicylic acid induces a *Nicotiana* sequence (van Kan et al., 1988)as does viral infection

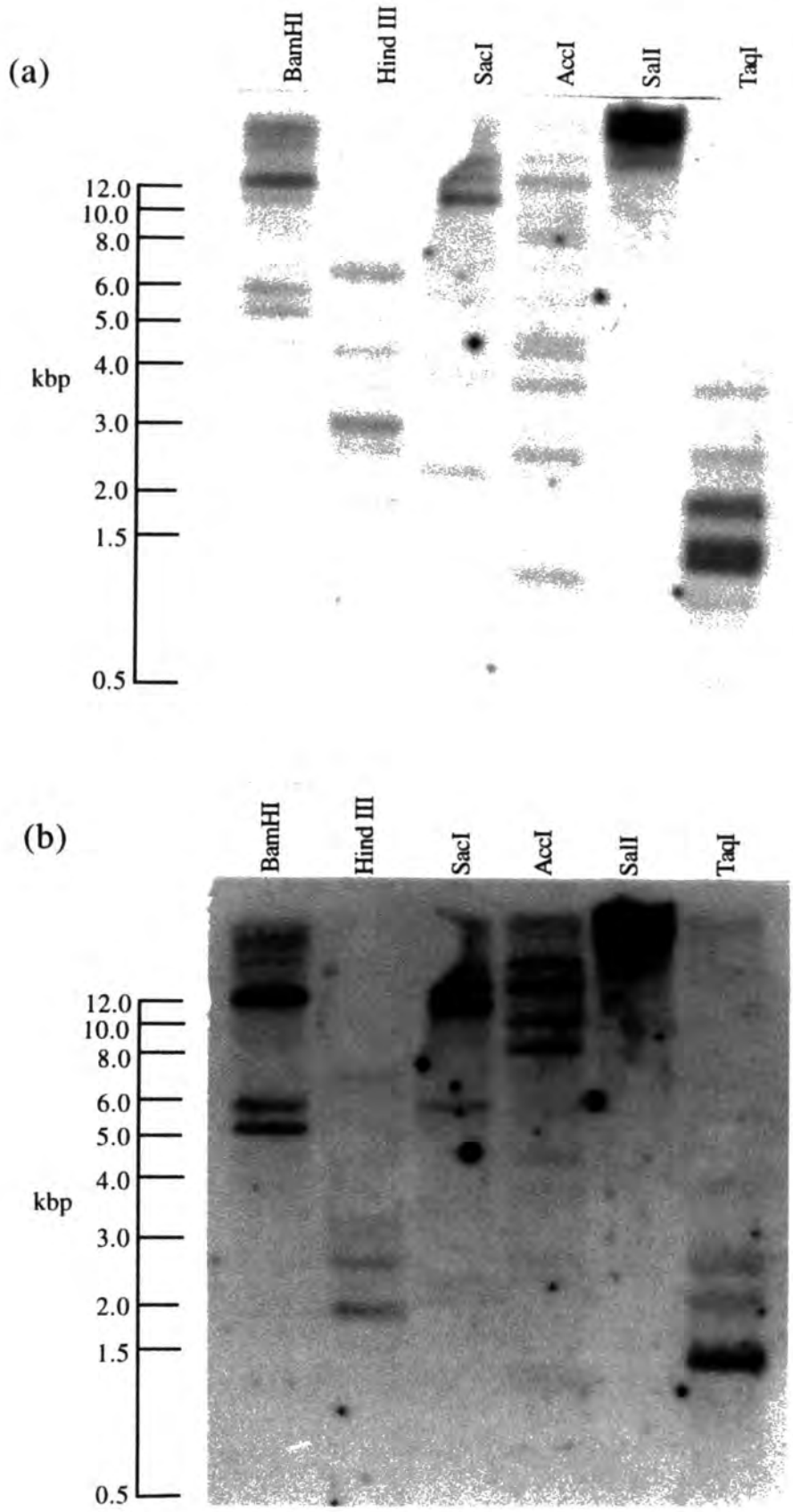


Fig.28. Genomic organisation of sequences homologous to the clone 15H11. 10µg of DNA was digested and loaded into each lane. The blots were hybridised in standard hybridisation mixture at 65°C as described in materials and methods. Figure shows autoradiographs of blots washed at (a) low stringency (1xSSC, 0.1%SDS at 65°C, ~85% stringency) or (b) high stringency (0.2xSSC, 0.1%(w/v)SDS at 65°C, ~94%stringency). Sizes are derived from a 1kbp marker co-run with the genomic DNA.

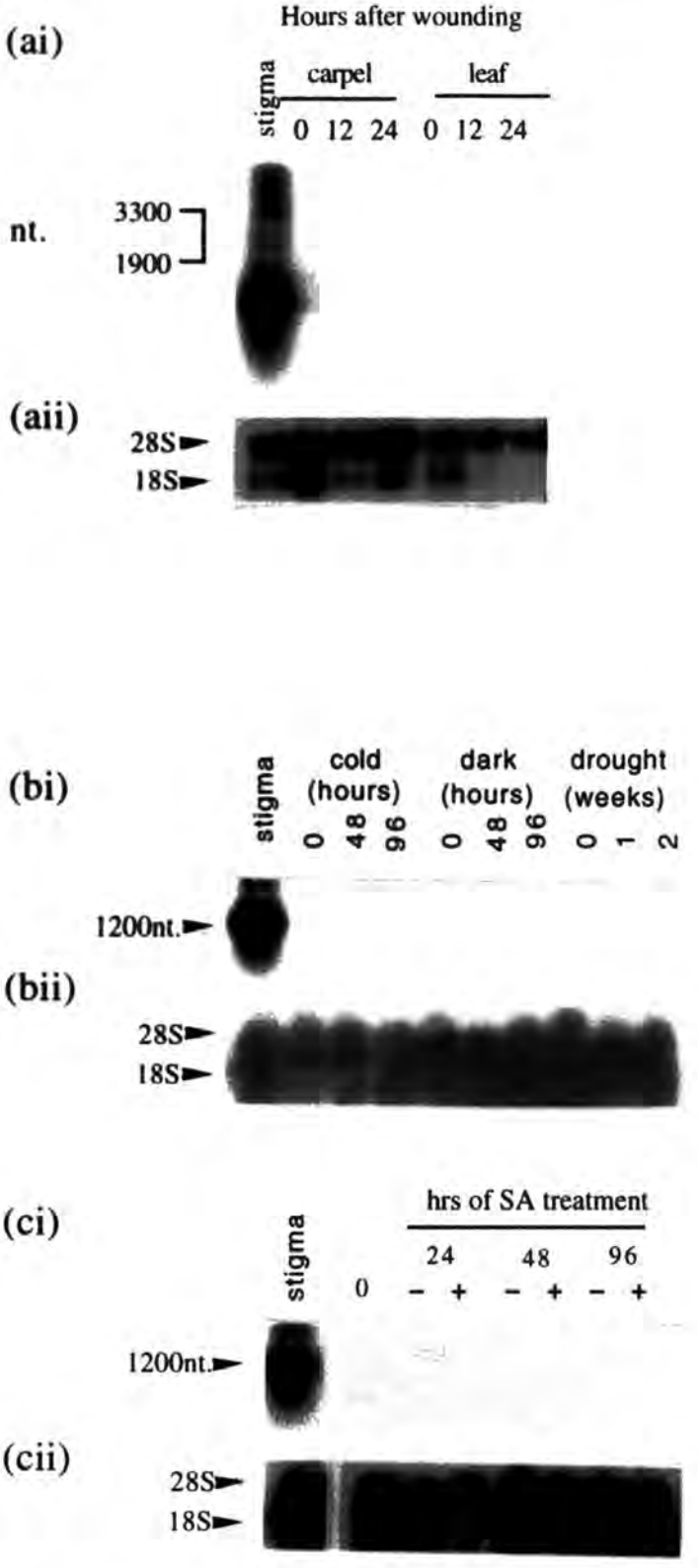


Fig.29. Northern blot analysis of the effect of various stresses on the expression of 15H11 in *B.oleracea*. (a)Autoradiographs showing the effect of wounding on carpels and leaves. (b)Autoradiographs showing the effect of environmental stresses on expression in seedlings. (c)Autoradiographs showing the effect of SA on leaves: +=leaf fed SA, -=control leaf fed water. Blots (ai), (bi) and (ci) were probed with the 15H11 cDNA insert and washed as described in fig.26. Blots (aii), (bii) and (cii) are the blots in (ai), (bi) and (ci) after hybridisation with pHA1.

which also induces an *Oryza* transcript (Fang et al., 1991). Water stress has also been reported to be induce changes in the level of expression of *Arabidopsis* sequences (De Oliviera et al., 1990). Given the relative complexity of the 15H11 gene family, in contrast to the pattern of expression in the mature plant, it was decided to investigate if some of the 15H11-like genomic sequences could be stress-induced genes.

The expression of 15H11 homologous transcripts was investigated under the conditions used in the study of 13G6(i) (4.1.1.8) together with salicylic acid treatment. Expression similar to that seen in the stigma could not be detected in vegetative tissues under any of the conditions tested, even at low stringency, although relatively weak expression, induced by wounding, was detected in the leaf (fig.29). However, the intensity of hybridisation relative to the stigma is such that this represents a relatively negligible increase in the level of expression or a transcript showing only weak similarity to 15H11. The latter is more probable as the size of the induced transcript was longer (>3300nt.) than that of 15H11 (1200nt.).

Therefore these results suggest that 15H11 or other members of its gene family are not directly induced by stress treatments although the performed experiments were not exhaustive.

4.1.2.4 The relationship of 15H11 to other GRP sequences.

The highly specific expression of transcripts similar to 15H11 indicates that they may play a specific role in the stigma and possibly also in the style. To determine what such a role may be the data available on 15H11, derived from this work and that of Scutt (1990), was compared with that of other GRP's. The EMBL and SWISSPROT databases were screened for glycine-rich plant protein sequences. These sequences are classified into two families: RNA binding proteins and other proteins which, by default, are usually classified as cell-wall proteins. The former are characterised by the presence of N-terminal RNA binding motifs (RGFGFVTF) and C-terminal glycine-rich domains which are also tyrosine-rich and are relatively small being 148-176 amino acids in length (Guiltinan and Xiping, 1996). Although the N-terminal coding sequence is lacking from 15H11 its size and amino acid composition preclude it from the RNA-binding family. Therefore these sequences were ignored and only the latter sequences selected. To minimise the analysis involved no EST sequences were analysed as these only represent partial fragments of sequence and are often inaccurate as they were produced by single pass sequencing leaving them prone to both substitution and frame-shift error.

Overall 16 sequences were analysed (table 6). Only three of these, GRP1.8, GRP1.0 (Keller et al., 1989b) and the *Petunia* GRP (Condit and Meagher, 1990) have been confirmed as extracellular by immunolocalisation. Of the other sequences, all but the atGRP4, atGRP5 and the *L.esculentum* GRP's contain potential N-terminal putative protein targeting sequences (see references in table 6) and therefore could also be secreted.

A restricted pattern of expression as seen with 15H11 has not been reported for any of these other sequences. Floral expression has only been detected on Northern blots for four sequences: *Phaseolus* GRP1.8 in the ovary (Keller et al., 1988) and *Arabidopsis* atGRP1, atGRP2 and atGRP4 in whole

GRP: Species and name		%Glycine*	Dominant motif	%15H11 motif*	%Tyrosine*	Reference
15H11		75	G(3)[A/V/L]G(3)(F/H/K]	95	0	C.Scutt, 1990; this work
Arabidopsis thaliana	atGRP1	76	G(3)AG(3)FG(3)[A/V]G(5)H	47	0	de Oliveira et al., 1990
Arabidopsis thaliana	atGRP5	76	G(1/3/5)[A/L/I/S/F/H]	0	0	de Oliveira et al., 1990
Arabidopsis thaliana	genclone	75	G(3)[A/I/L/F/H]	33	0	Quigly et al., 1991
Petunia hybrida		72	G(3)[G/A/V/L]	2	0	Condit & Meaghler, 1989
Oryza sativa	Osgrp1	70	G[G/A/S/L/Q/Y]	0	7	Lei & Wu, 1991
Phaseolus vulgaris	GRP1.0	69	G[G/A/V/I/S/H/D/E/Y]	0	7	Keller et al., 1989
Phaseolus vulgaris	GRP1.8	67	G[G/A/V/I/S/D/E]	0	6	Keller et al., 1989
Oryza sativa	Osgrp2	67	G[G/S/Y]	0	10	Fang et al., 1991
Lycopersicon esculentum		65	G(2-6)[P/R]	0	5	Showalter et al., 1991
Hordeum vulgare		62	G(4-6) Y	0	8	Rohde et al., 1990
Arabidopsis thaliana	atGRP2	59	G2[F/R/Y]	0	7	de Oliveira et al., 1990
Nicotiana sylvestris		57	GX	0	6	Obokota et al., 1991
Arabidopsis thaliana	atGRP3	44	G(4)[N/R]YQ	0	13	de Oliveira et al., 1990
Arabidopsis thaliana	atGRP4	36	G[G/A/L/P]	0	0	de Oliveira et al., 1990
Nicotiana tabacum		30	GX	0	11	van Kan et al., 1988

Table 6. Main properties of the primary sequence of the reported cell wall glycine rich proteins. All sequences are derived from cDNA or genomic nucleic acid sequences. Where square brackets occur in the motifs they contain the options for the amino acid residue at that position. The references given are those that first report the nucleic acid sequence and are not an exhaustive list of the references on that clone. * all composition figure are expressed as percentages of the glycine-rich regions as opposed to the whole protein where the figure is affected in some sequences by the presence of a glycine-poor signal sequence. Accession numbers for all the clones are catalogued in appendix. B.

flowers (De Oliviera et al., 1990). None of these are floral specific: GRP1.8 has been detected in seedlings and roots while the latter three clones were detected in the vegetative organs of *Arabidopsis*. Correlation of expression of some of these sequences with physiological processes has also been proposed: GRP1.8 and GRP1.0 expression has been linked to cells that are destined to be lignified, especially in the vascular bundles (Ye and Varner, 1991, Keller et al., 1989b). Such a function is not universal for the GRP's as reporter gene expression has indicted that the rice Osgrp1 promoter is active in the regions of cell differentiation and elongation (Xu et al., 1995).

The primary sequence of these proteins vary in a number of ways . Primarily the glycine content varies from 75% to 31%. This variation forms a continuum in that the sequences studied do not fall into discrete groups with similar glycine contents. 15H11 lies at the top of the glycine content range although its 75% composition may be influenced by the fact it is only a partial sequence and so sequence with a lower glycine content could occur in the N-terminal regions not present in the clone.

As well as the glycine content the major repetitive motif varies between the different protein sequences. This suggests that the GRP class may actually contain proteins with different properties and therefore varying physiological roles. Scutt (1990) proposed that the dominant motif in 15H11 is G(3)X. However, analysis of the sequence reveals that over 90% of the sequence consists of the octamer G(3)[A/V/L]G(3)[F/H/K]. Only three other sequences contain this motif and in only two (from *Arabidopsis*) does it form a high (>30%) proportion of the primary structure. In the *A.thaliana* atGRP1 this motif is intercalated with the motif G(3)[A/V]G(5)H which only differs from the 15H11 motif by the addition of two glycines in the second half. This high similarity between the putative 15H11 protein and the GRP's of a closely related species is unlikely to be phylogenetic as other *Arabidopsis* GRP's, with no examples of the 15H11 motif, have also been isolated. Repeats of the primary structure of these *Arabidopsis* clones can form an extended structure when analysed for secondary structure by the method of Garnier (de Oliviera et al., 1990), a structure also predicted for the putative 15H11 protein (fig.30). Such a secondary conformation appears to contradict that proposed for other GRP's which is an extended ß-sheet (Showalter, 1993), although de Oliviera et al. (1990) have suggested that their *Arabidopsis* clones can also adopt a predominantly ß-sheet conformation. Therefore that the 15H11 protein may have a similar overall general structure to all the other members of this class of protein.

One other difference between the GRP sequences is the amount of tyrosine they contain in their glycine-rich domains. This amino-acid residue is common in extensins and hydroxyproline-rich proteins

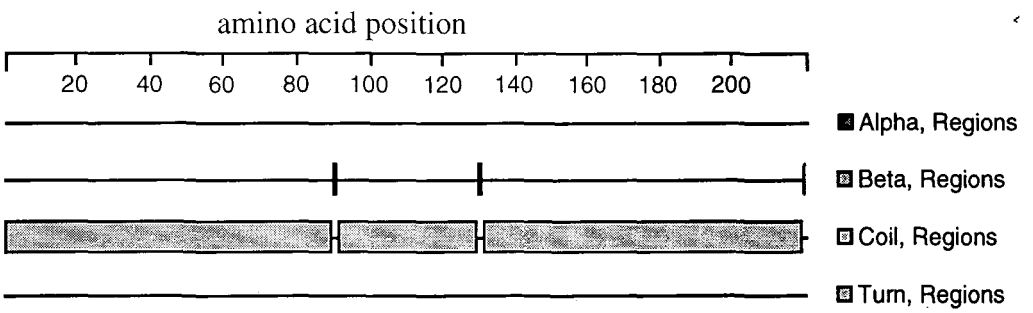


Fig.30. Garnier predictions for the secondary structure of the predicted 15H11 protein. The thickness of the bars represents the propensity to form that structure. Generated using the 'Protean" program of DNASTAR.

(Fry, 1986). This generates a rigid framework around which other cell wall components can assemble. A similar activity has been predicted for the *Phaseolus* GRP's and *Arabidopsis* atGRP3 which have regular arrays of tyrosines (De Oliviera et al., 1990). The correlation between the expression of the former proteins with pre-lignified tissue (Ye and Varner, 1991) has lead to the hypothesis that these proteins may act as substrate for the initiation of lignification.

No tyrosine is present in the predicted 15H11 protein suggesting that if the cross-linking hypothesis is right then this protein cannot covalently cross-link by known mechanisms. This would correlate with expression in the stigma which must contain many flexible cell walls to allow pollen tube penetration and growth, particularly in the papillae. A similar relationship with the lack of induction by salicylic acid and wounding where the induction of peroxidases and subsequent cross-linking of the wall has been documented (Bowles, 1990) can be proposed. However the level of tyrosine in the GRP sequence does not necessarily correlate with expression under stress: The *Petunia* sequence which is wound-induced contains no tyrosine (Condit and Meagher, 1987) while the *Oryza* Osgrp2 sequence has a relativley high tyrosine content but has not been reported to be stress induced (Lei and Wu, 1991).

4.1.2.5 Possible Functions of 15H11.

Scutt (1990) suggested that 15H11 expression may be in the vascular system as localisation to similar tissue had been shown in *Phaseolus* (Keller et al., 1989a). However, from the data presented here such a role need not be the case. Although they share a high glycine content the *Phaseolus* proteins and the predicted 15H11 protein have different primary structures with 15H11 being significantly more repetitive in its sequence and lacking tyrosine for cross-linking. Similarly, the much higher level of 15H11 transcripts in the stigma tissues when compared to the rest of the carpel is contradictory to this localisation. The vascular tissue of the carpel is considerably more well developed in the style and ovary than in the stigma (Hill and Lord, 1987) and so the opposite pattern of expression would be expected if 15H11 represents a vascularly expressed transcript Expression in a non-stigma-specialised tissue however, cannot be discounted as biochemical zoning in apparently homogeneous tissue, as with *Lycopersicon* pectate lyase (Budelier et al., 1990) cannot be ruled out.

The patterns of expression obtained in this work suggests that 15H11 is expressed in one or more of the specialised tissues in the stigma and style, as the nature of these tissues would explain the restricted expression of transcripts homologous to 15H11. The observed petal expression could be used as an argument against this hypothesis. However, as the transcripts detected in the petal could be as much as 35% different to 15H11 they may encode a different class of GRP altogether.

The high expression in the stigma and weaker expression in the style would be consistent with the distribution of the transmitting tract in these organs although the difference in expression could also be produced by biochemical zoning. A lack of detectable expression in the ovary can be explained by dilution by other RNA's, since even less of the ovary is transmitting tissue than the style and stigma. The precise role of 15H11 in this tissue, however, cannot be predicted from this work alone. The lack of any tyrosine residues means that it is unlikely to be involved in cross-linking and strengthening cell walls, as has been predicted for other GRP's. Relating this property to the transmitting tissue is difficult due to a lack of

knowledge about this environment. However, as the cells are thin walled, strengthening components would be expected to be present in these cells which is at odds with this hypothesis for the function of the putative 15H11 protein. However 15H11 may not exist in the cell wall, it may be present in the apoplast as has been demonstrated for some of the solanaceous, style cell-wall proteins like the TTS proteins (Cheung et al., 1993). It is also possible that tyrosines may be present in the protein sequence not represented on this clone.

The other principal specialise tissue in the upper portion of the stigma is the stigma surface. Expression in the secretary or papillar cells of the stigma surface could be proposed, although restriction to these tissues is inconsistent with the detection of 15H11-like transcripts in the style at high stringency. If it cannot cross-link presence of the putative 15H11 protein could contribute to the papillar cell walls elasticity that allows pollen tubes to penetrate and grow between the two layers (Elleman et al., 1990). As such a property is believed to be unique to the cell walls of the stigma surface this could explain the high level of expression of 15H11 in the stigma.

To address these questions more data on the localisation of either the 15H11 transcript or predicted protein is required. Of these two, protein localisation would provide the most information as the it could confirm whether 15H11 does encode an extracellular protein as well as identifying the expressing tissue(s). In order to achieve this a synthetic peptide, based on the predicted sequence of the 15H11 protein was synthesised and used to raise antibodies. As this was carried out by other workers and is currently still in progress a description is provided in appendix E.

4.2 Library construction and screening.

4.2.1 Strategy for library construction and screening.

The two clones isolated by Scutt (1990) revealed little about the biology of the stigma. Thus it was decided to generate a new cDNA library and rescreen this for other stigma-specific clones. As with the previous work the stigma was chosen as the source of the RNA for the library as this organ contains greater concentrations of transmitting, papillar and secretary zone tissue than the style or the stigma and style combined, thus maximising the chance of isolating clones expressed in these tissues. The one disadvantage of using only stigma tissue was that more extensive tissue harvesting was required than for a stigma/style library.

A lambda-based library system was chosen as it would allow plating at higher densities of clones than plasmid-based libraries, and so would be less labour intensive to screen. A differential screening method (Sambrook et al., 1994) was chosen as a method for identifying stigmatic clones as it is an established, technically simple method. The principal drawback of this method is that only transcript classes which constitute greater than 0.05% of the positive mRNA population can be detected. More recently developed methods such as cold-plaque screening (Hodge et al., 1992), subtraction library construction (Zimmerman et al., 1980) and differential display (Liang et al., 1993) have been reported to overcome this drawback, although these methods are technically more complex than differential screening. A differential screening approach, however, was still chosen as the inherent limitations were predicted not

A differential screening approach, however, was still chosen as the inherent limitations were predicted not to be a problem for this project as the biochemical/physiological character of the stigma tissues would be predicted to be predominantly dictated by the most abundantly expressed sequences in them. Experimental justification for this decision also comes from the work on the Solanaceae, from which many of stigma/style specific cDNA's have been isolated by differential screening approaches.

The two probes for differential screening were derived from stigma and leaf poly(A) RNA. The latter was chosen as the negative probe as it was successfully used by Scutt (1990) and required minimal effort to harvest.

4.2.2 Library construction

A stigma cDNA library was constructed as described in section 3.7 from 2μg of poly(A) RNA. Packaging of the library was done in two separate reactions. The first of these gave a library of titre $1.4x10^6$ pfu. The second reaction gave $3.7x10^5$ pfu so that, on pooling, a library of approximately $1.8x10^6$ pfu was created.

To assess the quality of the library insert lengths, ten random, discrete clones were picked and converted to phagemid. However, only five bacterial clones were produced from the equivalent λ clones. This poor efficiency meant that later inserts were subcloned into pKS+Bluescript from the λ DNA.

Plasmid DNA was isolated from the five successfully generated phagemids and digested with EcoRI and NotI. This indicated that the inserts were between 300 and 2200bp in length, suggesting that the library did not solely consist of cloned small fragments (fig 31) and so should give clones with large amounts of sequence information. The library was subsequently amplified by plating on a $23cm^2$ plate and eluting the phage into SM.

4.2.3 Differential screening.

Differential screening was carried out using probes derived from stigma and leaf poly(A) RNA. Each RNA was converted to double stranded cDNA and approximately 50ng of this was used as a template for labelling by random priming reaction.

Initially 20000pfu were plated out in a single 13.6cm diameter dish. After overnight growth, duplicate filters were lifted, from this plate, and hybridised with the two cDNA probes. Hybridisation was carried out overnight to maximise binding of the complex probe to its targets. The filters were washed in 1xSSC, 0.1%(w/v)SDS at 60°C which gave a predicted stringency of 70-85% assuming a GC content of 60-30% for the clones.

Autoradiographs, exposed to the two filters for three days, indicated a large number of differently hybridising plaques (fig.32a and b). From these, 86 different plaques were picked using blue Gilson tips cut to a 7mm diameter. These consisted of :

1 plaque that hybridised strongly to both probes (A2),

1 plaque that hybridised strongly to the leaf but not to the stigma probe (A7),

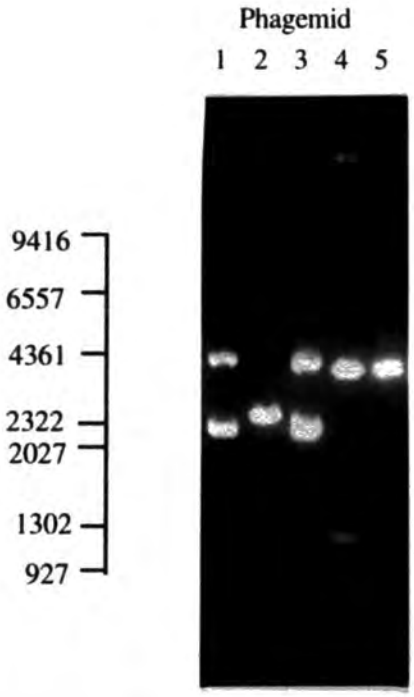


Fig.31. Analysis of the insert size of five random clones from the stigma cDNA library converted to phagemid *in vivo*. Size markers were derived from a mixture of HindIII digested lambda DNA and Sau3AI digested phage 174 DNA.

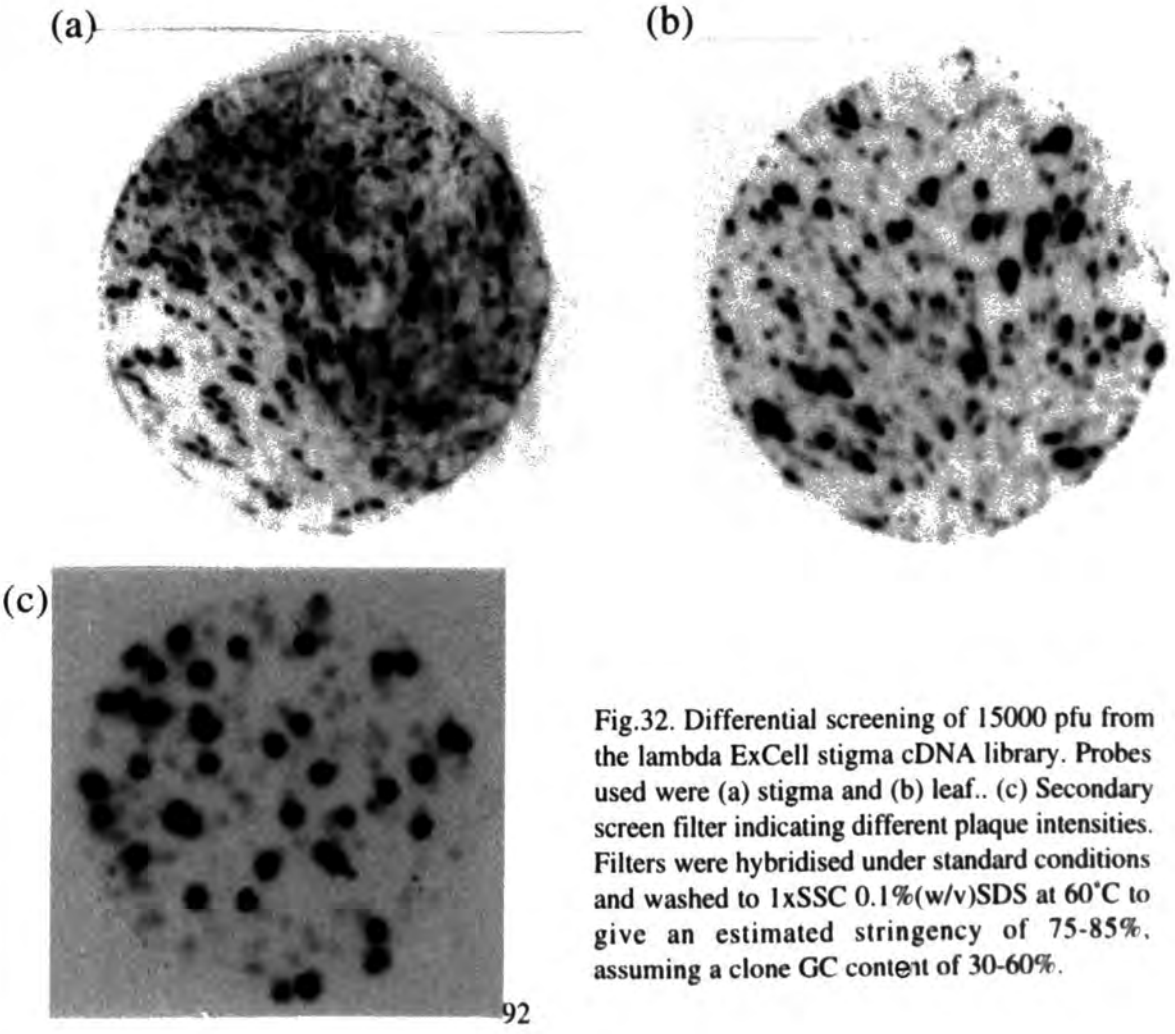


Fig.32. Differential screening of 15000 pfu from the lambda ExCell stigma cDNA library. Probes used were (a) stigma and (b) leaf.. (c) Secondary screen filter indicating different plaque intensities. Filters were hybridised under standard conditions and washed to 1xSSC 0.1%(w/v)SDS at 60°C to give an estimated stringency of 75-85%, assuming a clone GC content of 30-60%.

84 plaques that hybridised strongly to the stigma but not to the leaf probe (A3-6, B1-10, C1-10, D1-10, E1-10, F1-10, G1-10, H1-10, K1-10).

Samples from these plaque picks were spotted out and screened with the two previously characterised cDNA's (13G6(i), 15H11). Secondary screening was then carried out on an initial 12 selected plugs, plating at a density of 1-300pfu per plate. Screening was carried out only with the stigma probe in order to minimise the number of filters required.

In many cases the secondary screen filters displayed more than one hybridising species of plaque (fig.32(c)). Individual plaques for all these species were picked and spotted out for a final screen with both stigma and leaf cDNA probes. This allowed the characterisation of 30 different clones for further analysis. For preliminary analysis nine of these clones were originally selected for further study. DNA was isolated from the lambda phage, digested with EcoRI and NotI and the inserts subcloned into pKS+Bluescript to allow the generation of sufficient quantities of DNA for sequencing and radiolabelled probe generation.

4.3 Analysis of the clones generated from the λ library

4.3.1 B6E, C10B, K2B and K8E2 encode previously characterised S-genes

Four clones (B6E, C10B, K2B and K8E2) isolated from the differential screen did not display any detectable hybridisation with the leaf mRNA probe but hybridised strongly with the stigma probe. These clones when isolated were shown to contain inserts of approximately 500, 400, 400 and 1000bp respectively (data not shown). Preliminary sequencing was carried out on these clones as described in fig.33. The sequence data obtained from this work indicated that the clones represented partial sequences of two previously characterised S-genes.

C10B K2B and K8E2 represented partial sequences of the *B.oleracea* SLG29 gene (fig.34) which was originally isolated as the cDNA clone BoS29-2 (Trick and Flavell, 1989). Sequence was obtained from both ends of K2B and K8E2 while C10B was sequenced from the 5' end only. This gave the full sequence of K2B and C10B (on both strands for the former) and ~65% of the sequence of the K8E2. The sequences of these clones are 99%, 98% and 97% identical to the respective parts of SLG29.

The principal differences between these clones lies in their polyadenylation site. C10B, K2B and K8E2 are polyadenylated 35, 107 and 26bp. earlier than BoS29-2, respectively, indicating that there is variability in the 3'end processing of the SLG29 mRNA. Related to this feature is the presence of two consensus AATAAA signals in the 3' UTR of SLG29 (fig.35). Such polyadenylation signals are usually located 11-30bp from the 3' polyadenylation site (Hunt, 1994) so that if these motifs are active in the SLG29 gene, then the first signal could act to trigger 3' cleavage of K2B, while the second signal could be responsible for termination of C10B and K8E2. No other poly(A) signals occur 3' to the second signal, although a 5/6 match does occur 20-25bp upstream from the reported BoS29-2 poly(A) site. Such

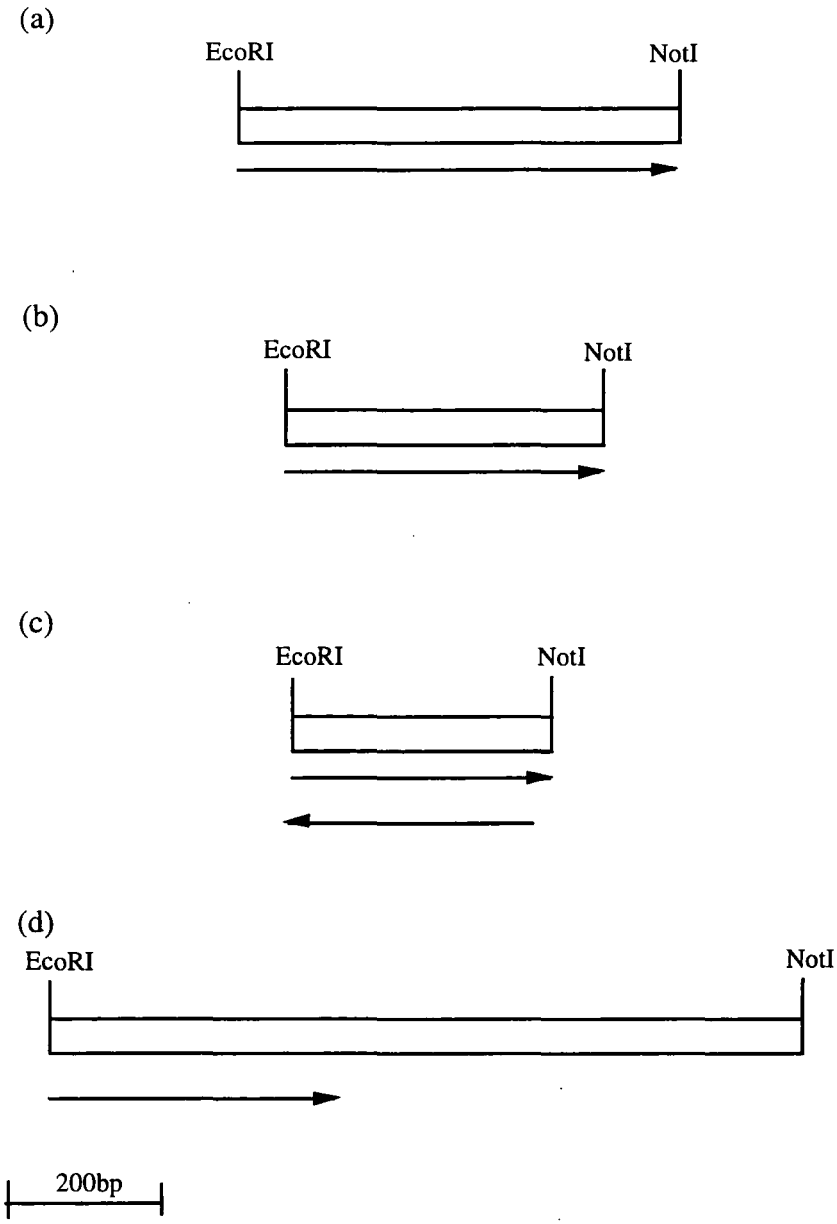


Fig.33. Sequencing performed on the four stigma clones that represent previously characterised SLG or SLR genes. (a)B6E, (b)C10B, (c)K2B, (d)K8E2. Arrows indicate the direction and extent of each sequencing reaction.

```
1200. GGTGGGACGGGTTGTGTGATTTGGACCGGACAGCTTGAGGATATCCGAACTTACTTTGCT .1259
1260. AACGGTCAAGATCTTTATGTCAGATTGGCTCCTGCTGACCTTGTTTAGCTCTTTCTCTTA .1319
1320. AAATAAAACACGGATCCAACTACGTTATCGAGAAATCCGAATATGTGACGCAACTATCTC .1379
1380. AATATGTCGACCCTTAATGTATTTCTGAATCAATAAAGATATACAGATCGTATCCAAATC .1439
1440. CAATAATATATGAAAACATGCCATAT                                    .1465
```

Fig.35. Position of putative polyadenylation signal site in the 3' UTR of the SLG29 cDNA (Trick and Flavell, 1989). The two AATAAA consensus sites are underlined with solid lines while the 5/6 match is underlined with a dotted line. The three polyadenylation sites of the clones C10B, K2B and K8E2 are indicated by bold lettering. As reference to the coding region of the transcripts represented by the cDNA inserts the predicted stop codon for the SLG29 protein (nt. 1315 in the BoSLG29 cDNA) is indicated by italic lettering.

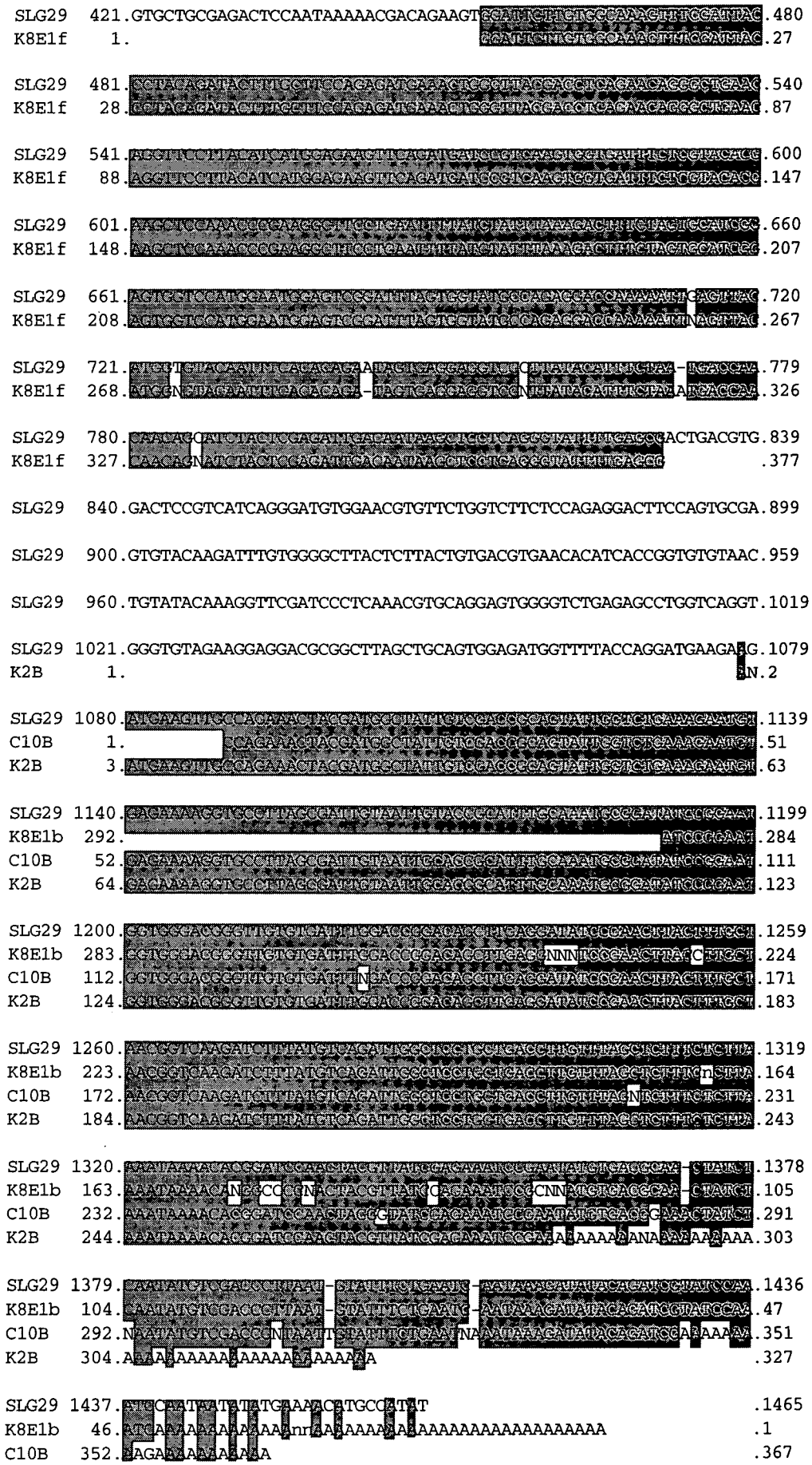


Fig.34 (opposite page). Alignment of the sequence of the stigma clones C10B, K2B and K8E1 with the SLG29 cDNA (Trick and Flavell, 1989) using the Clustal W1.5 multiple alignment program. The two fragments of K8E1 sequence are distinguished from each other by f (sequence from 5' end) and b (sequence from 3' end). Numbering for K8E1 is backwards with respect to the other sequences as it is the complement of the actual sequence obtained. Shading is used to indicate those nucleotides that are identical to the equivalent base in SLG29.

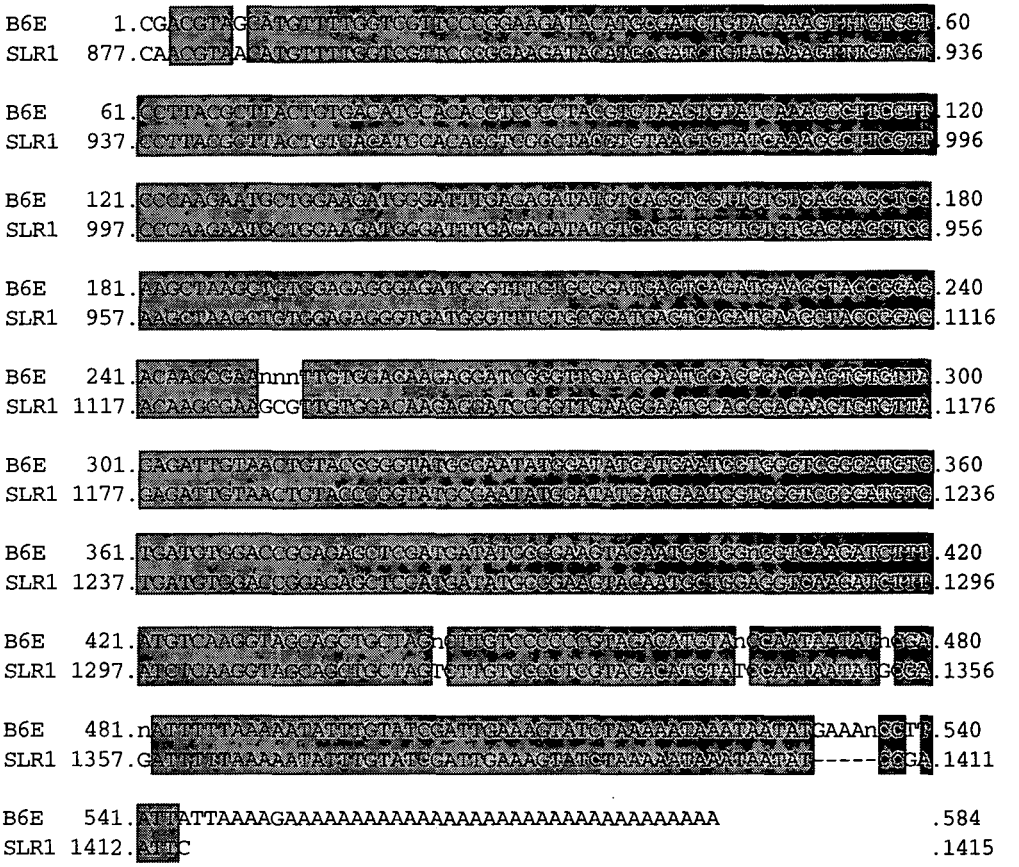


Fig.36. Alignment of the nucleotide sequence of the clone B6E with the SLR1 cDNA (Trick and Flavell, 1989) sequence using the algorithm of Clustal W1.5. Shading indicates identical residues.

functional in this transcript but, as the plant termination signals are still poorly defined (Hunt, 1994), no firm predictions can be made on any functional role for this motif.

Variable termination is not a novel phenomenon, having been previously reported in many plant genes (Dean et al., 1986). The differences reported here are unlikely to have any relevance for the biology of the SLG29 gene as the different poly(A) sites all lie in the 3' UTR of the gene (fig.35).

Other differences between the clones and SLG29 are principally due to the presence of unassigned bases or single base insertions. These can be accounted for by the fact that with the exception of K2B all of the S29 homologues have been sequenced with single passes so that several unassigned or incorrectly sequenced bases would be predicted.

The entire sequence of B6E was obtained by reading from the 5' end of the insert and shows greater than 97% similarity to the *Brassica* SLR1 gene (Trick and Flavell, 1989) (fig.36). Once again, the bulk of the differences are due to undetermined nucleotides in the B6E clone while the last few nucleotides before the poly(A) tail are divergent between the two sequences. This difference could be due to the two clones being transcribed from different genes as the precise copy number of SLR1 has not been determined.

The isolation of SLG and its related sequences was not unexpected due to the high abundance of these sequences in the stigma of *B.oleracea* (Nasrallah et al., 1985). No further work was carried out on these sequences as they have been exhaustively studied by several groups already (reviewed in Nasrallah and Nasrallah, 1993).

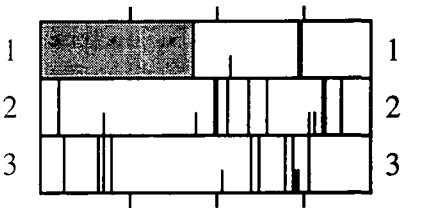
4.3.2 The Clone C10C

4.3.2.1 Sequence of the clone C10C

The C10C clone displayed only weak hybridisation to the stigma probe, when compared to the other isolated clones but also showed no detectable hybridisation to the leaf probe. The full clone was sequenced by reading from the 5' end which gave an insert length of 373bp. As only a single sequencing pass was made three undefined bases are present, at nucleotide positions 41, 327 and 333.

The clone contains a poly(A) tail allowing it to be orientated so that the ORF search was only performed for three of the six possible translation frames. Analysis of the translation products indicated that one possible reading frame is present (fig.37a), though it does not contain an in frame ATG. This coupled with the small size of the insert suggests that the clone only represents a partial mRNA sequence. As with the previously described sequences, the presence of a poly(A) tail indicates that the 5' end of the clone is missing. Therefore the available nucleic acid and predicted sequence of the clone is as described in fig.37b.

(a)



(b)

```
  1.TTCTTGCTCACGATCCAGGTGACCCCTAAGCACGGTGGGGNAGGGAGTATTGTTCACTGGCACCTTGAGTATGAGAAGAT.80
     F  L  L  T  I  Q  V  T  P  K  H  G  G  x  G  S  I  V  H  W  H  L  E  Y  E  K  I
 81.TAGCGACGAGGTTGCTCATCCCGAGACTCTCCTCCAGTTCTGTGTGGAAGTCTCCAAAGAGATCGACGAACATCTCTTGG.160
     S  D  E  V  A  H  P  E  T  L  L  Q  F  C  V  E  V  S  K  E  I  D  E  H  L.L
161.CCAAGGAATAAAGGAATGTGTGTGTGTGTCTGTTTTTAATAAAGAAATGTGTGATGTGCCTTTCTTTCTGTATTTTAATT.240
     A  K  E
241.AAACGCAATAAGGGATTAAGGAGCCAAGTCCGCGTTTGTGAGAGATATAAATGATGAGTGTGCCATGATTATGTGTGTAT.320
321.AATAAGNATCAANTTACGTTTGAAGTTGTATCAAAAAAAAAAAAAAAAAAAAA                          .373
```

Fig.37. Sequence of the cDNA clone C10C. (a) ORF map of the C10C insert. Half-height bars indicate ATG codons while full height bars indicate potential codons. The predicted reading frame is indicated by shading. (b) Sequence of C10C together with its predicted, partial protein. The 5/6 match to the animal poly(A) signal consensus sequence is double underlined while the partial FUE consensus is singly underlined.

Unlike 13G6(i) and 15H11 there are no putative AATAAA polyadenylation signals present in the clone, although one 5/6 match to this motif is present at nucleotide position 321-326. Such an absence is not unusual in plants where only 30-40% of currently identified cDNA's contain a copy of this motif (Hunt, 1994). A second sequence which is similar to an upstream element implicated in polyadenylation of CaMV RNA (Sanfacon et al., 1991) and also present in the 3' UTR of pea *rbcS*-E9 gene (Mogen et al., 1992) is present at nucleotide positions 346-350 (fig.37b). What, if any, role these motifs play in the termination of the C10C transcript is not known.

4.3.2.2 The predicted C10C protein is homologous to a poppy latex sequence.

In order to investigate the possible function of the putative C10C protein, the nucleotide and protein sequences were used to screen the EMBL and SWISSPROT databases. Screening of the SWISSPROT library revealed only one match to the predicted C10C protein sequence. This was to the major latex protein (MLP) of *Papaver somniferum* (Nessler et al., 1990) which has 36% identity with the putative C10C protein, over the length of the latter, while another 15% of amino acids pairs are equivalent (fig.38). The *P.somniferum* MLP is a major component of latex, a viscous sticky liquid that is stored in the cytoplasm of specialised cells (lactifers). In the angiosperms lactifers occur in fifteen families but appear to have multiple evolutionary origins (Nessler and Vonder Haar, 1990). In *P.somniferum* the

lactifer cells are associated with the vascular system throughout the plant (Nessler and Mahlberg, 1976) and release their contents on wounding of the plant.

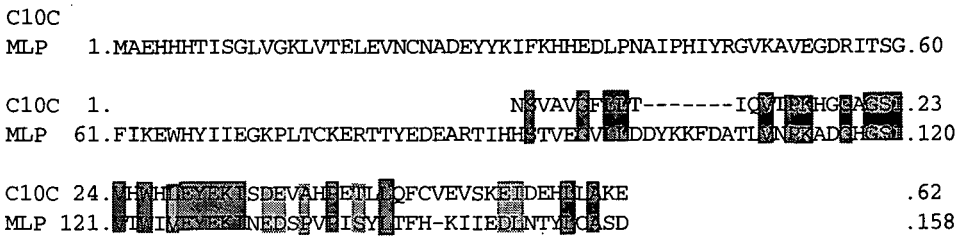


Fig.38. Alignment of the putative C10C protein with the protein sequence of *P.somniferum* major latex protein. Dark shading indicates exact amino acid matches. Light shading indicates conservative amino acid changes as defined in section 3.9.

In order to assess the relevance of this result the EMBL and SWISSPROT databases were searched for other latex proteins and/or their genes. 5 sequences were available of which 1 was ignored as it was defined as a 'latex gene' only on the basis of its homology to the *Papaver* sequence. The remaining sequences fell into two families: *P.somniferum* MLP's and *A.thaliana* MLP's containing 2 sequences each. Unlike the *Papaver* sequences, the *Arabidopsis* sequences are tissue-specific, originating from

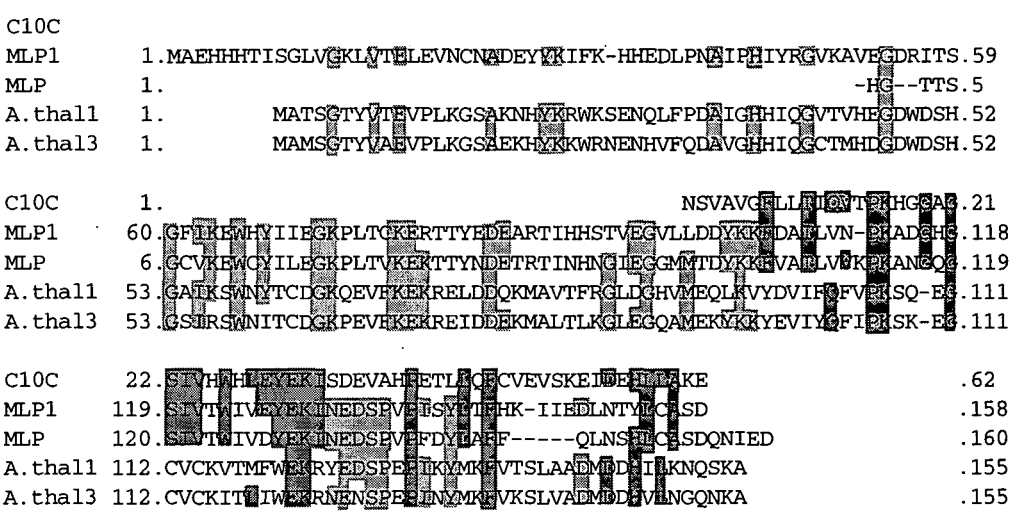


Fig.39. Alignment of the predicted protein sequence of C10C with the reported MLP protein sequences using the Clustal W 1.5 program. Dark grey shading indicates identical matches between C10C protein and the MLP's. Light grey shading indicates consensus sequence residues between the MLP's that do not match with the sequence of C10C.

immature siliques, the organ which the carpel develops into after floral maturation. Given this mode of expression and the close relationship between *Arabidopsis* and *Brassica* these latter two sequences might be predicted to be more homologous to C10C than the *Papaver* sequences, if C10C is an MLP.

Alignment of the two families and C10C indicates that the regions of high identity between the two families occur in the middle and at the C terminal end of the sequences (fig.39). In the N-terminal region the two *A.thaliana* sequences show high (74%) identity to each other but much lower (23%) identity to the one *P.somniferum* sequence that encompasses this domain. This compares to the central and C-terminal regions of the proteins which show ~50% identity between the four MLP's.

From its alignment to these sequences two factors suggest that C10C does not encode an exact homologue of them (fig.39): Firstly C10C does not align completely with the other proteins and over the 62 amino acids of C10C only 18 (29%) align with all the MLP's which compares with a 50% identity over this region between the MLP's. Secondly the degree of identity between C10C and the Arabidopsis sequences is slightly lower (32%) than the Papaver sequences (35%). This latter factor suggests that C10C does not represent a fragment of the *Brassica* equivalent of the *Arabidopsis* MLP sequences. However this data does not preclude the transcript from which the clone was derived coding for a protein evolutionarily related to the MLP's.

4.3.2.3 Nucleic acid homologues of C10C

As the predicted protein sequence available for C10C is not sufficient to precisely determine its function, the EMBL database was screened to determine if any previously characterised nucleic acid sequences show similarity to it. This screen did not pick up any MLP genes, but a series of EST's and one complete cDNA sequence were detected, all of which were derived from *A.thaliana* tissue and showed greater than 60% identity to C10C over greater than 75nt. (table 7). All but two of these sequences were derived from the EST programme of Newman et al. (1994) and so their pattern of expression is not known. The remaining two sequences were derived from dry seeds (Atts04406) and green siliques (Atts0356). As well as being related to each other (siliques develop into seed pods) these expressing tissues correlate with the silique expression for the two full length MLP sequences discussed in section 4.3.2.2.

Alignment of these sequences with C10C reveals that they do not all represent the same class of transcript. None of the sequences align completely with C10C but only show identity to part of the sequence. On the basis of these differing regions of identity the clones can be divided into four classes (fig.40).

The first three classes align in the same orientation as C10C but the fourth class align in the opposite orientation, i.e. to the anti parallel strand of the predicted C10C mRNA (fig.41). The relevance of this is not known although it may suggest that the two transcripts are related by gene duplication and inversion such as has been predicted for the GRP and PRP protein genes (Showalter et al., 1991). Another explanation is that the members of this class represent mRNA species transcribed from the same region of DNA as a C10C-similar mRNA. Transcription from both strands of the same locus has been reported for a putative male component of the *Brassica* S-locus (Boyes and Nasrallah, 1995) where one of

Sequence	Length (bp)	identity/overlap	Reference
At0564	562	87.6%/170bp	Newman et al., 1994
At7042	509	87.1%/171bp	Newman et al., 1994
At3578	502	87.1%/171bp	Newman et al., 1994
At5205	386	77.9%/204bp	Newman et al., 1994
At0683	467	87.6%/161bp	Newman et al., 1994
At13934	377	76.6%/175bp	Newman et al., 1994
At82813	392	89.4%/113bp	Newman et al., 1994
Atts0356	301	81.3%/75bp	Raynal et al., unpublished
At21018	502	70.5%/166bp	Newman et al., 1994
At37014	527	74.7%/182bp	Newman et al., 1994
Atts4406	711	64.9%/342	Raynal et al., unpublished
At4953	362	69.8%/139	Newman et al., 1994
At52221	507	75.8%/252bp	Newman et al., 1994
At41020	457	74.3%/253bp	Newman et al., 1994

Table 7. Nucleic acid sequences that show similarity to the clone C10C. EMBL database accession numbers are listed in appendix B

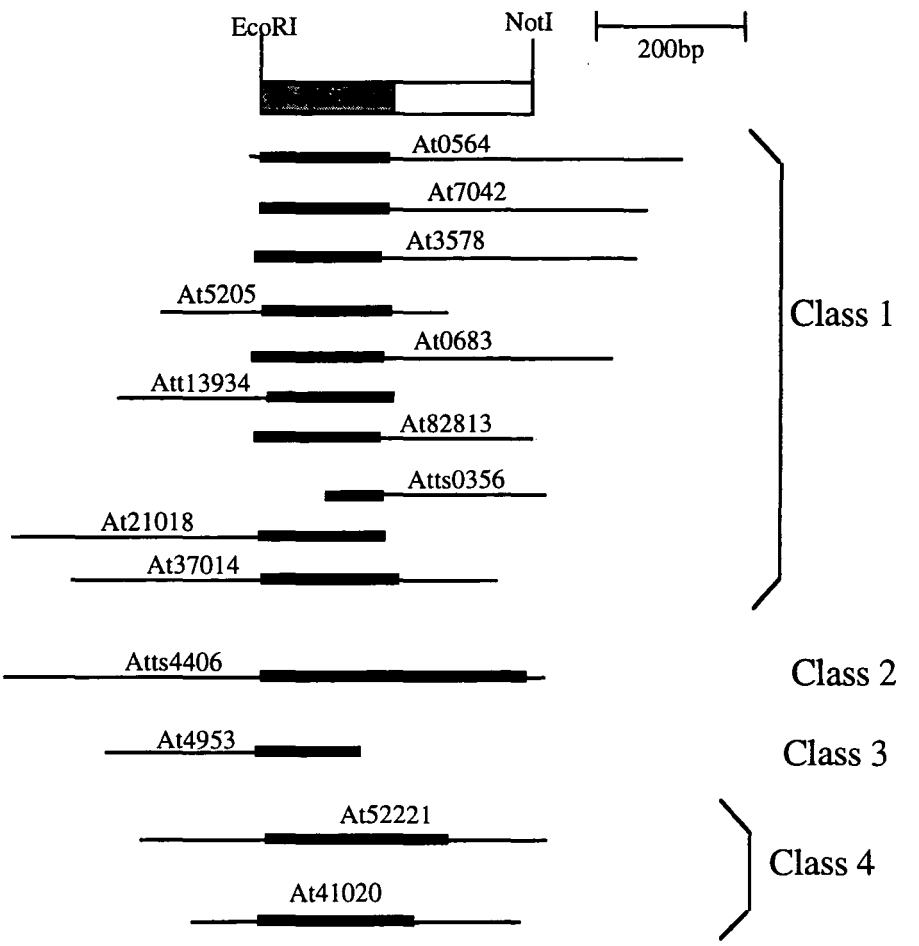


Fig.40. Relationship between C10C and the sequences that show similarity to it. Thickened bars indicate regions that are similar (>60% identity) to C10C, The shaded region indicates the extent of the predicted coding region of C10C.

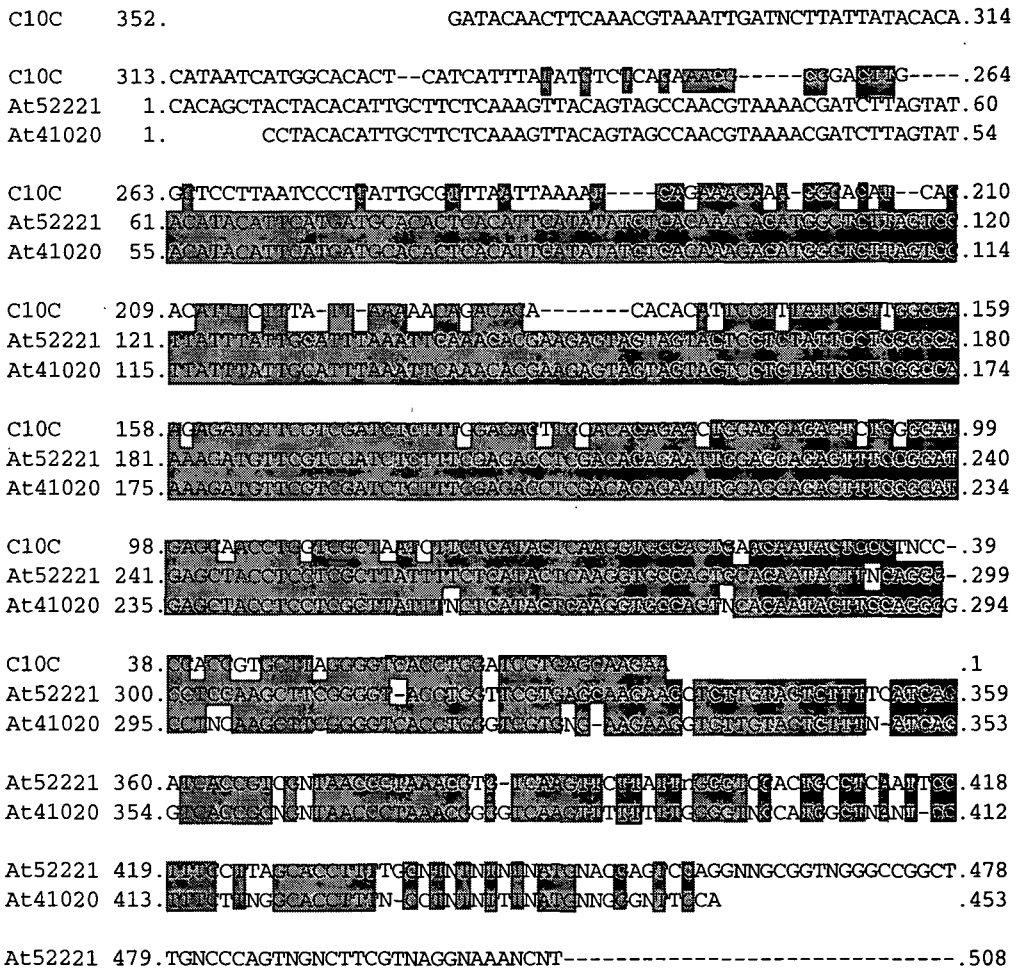


Fig.41. Alignment of the antiparallel strand sequence of C10C with the two class IV EST sequences using the Clustal W 1.5 program. Shading indicates identical nucleotides.

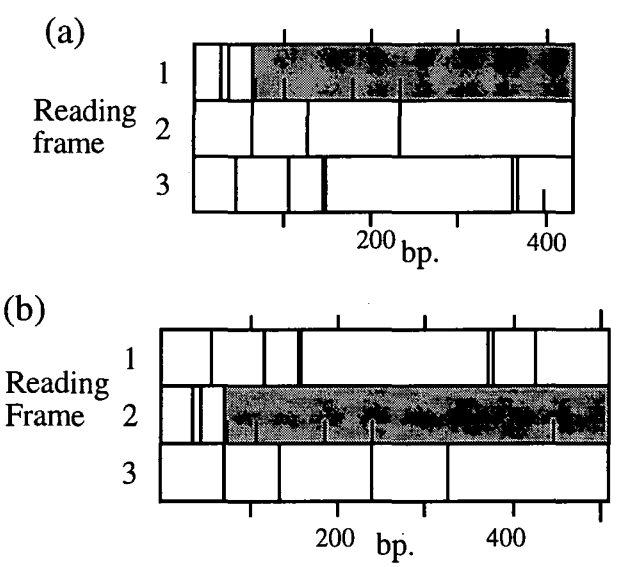


Fig.42. ORF maps for the EST clones At41020 (a) and At52221 (b). The predicted open reading frames are indicated by shading. Half height bars indicate ATG codons while full height bars indicate stop codons.

the transcripts has no coding capacity and has been proposed to regulate the coding transcript by an antisense mechanism. As both members of class IV do contain large reading frames (fig.42) this is unlikely to be the case with C10C-similar mRNA's and these transcripts.

All the other classes align in the same orientation as the predicted mRNA represented by C10C. The majority of the clones are in class I, in which all the members share identity with the first 170-180bp of C10C. None of the sequences show high identity to the 3' half of C10C, although for two of the clones (At21018 and At13934) this can only be inferred as the sequence for this region is not available (fig.40). This pattern of high-identity regions is notable in that the point at which divergence occurs is near to the end of the predicted C10C translation frame (43a). Therefore the point of divergence could be due to the junction between coding and non-coding regions of the transcripts.

The alignment of the members of this class is complex as the EST's were all produced by single pass sequencing and so are predicted to contain errors, particularly around their 3' ends (fig.43b). This is apparent from the large number of unassigned bases and small insertions/deletions in the overlapping regions containing 3' sequence from one of the clones. Some differences also occur between the clones in their N-terminal regions, suggesting that they may represent closely related, but different transcripts. However, these differences constitute less than 5% of the aligned regions so could be sequencing errors.

Two large scale difference also occur. Firstly the clone Atts0356 diverges significantly from the other EST's at the same point as the sequence of C10C. This suggests that this clone represents a distinct class of transcript from the other class I sequences . Secondly the sequence At0564 has an 18nt. insertion at position with respect to the other EST's (fig.43b). The maximum effect of this insert on any encoded protein would be a six amino acid insertion as no stop codons occur in the 18nt. Therefore the class I transcripts can be subclassified into class Ic (Atts0356), class Ib (Atts0564) and class Ia (all other sequences)

The one class II transcript is a near full length sequence that displays similarity to C10C across the first 90% of the length of C10C (fig.44a). However, the level of identity is lower (~65%) than the class I sequences (75-90%) which suggests that Atts4406 is more evolutionary distant to C10C than the latter sequences. The authors of the class II sequence indicate claim that it may encode a latex protein but do not provide any evidence for this other than nucleic acid similarity (Raynal et al., 1994).

As with class II, class III contains only one transcript, At4953. This shows ~70% identity to C10C for the first 130 nucleotides of the latter sequence (fig.44b). Only 15-20 nucleotides occur 3' to the region of overlap so a full comparison with the class I transcripts which also show this partial identity cannot be made.

4.3.2.4 Predicted proteins from the C10C-like nucleic acid sequences.

Many of the sequences which show identity to C10C contain sequence information from outside the C10C clone and so could provide additional information on whether C10C could encode a MLP-like

(a)

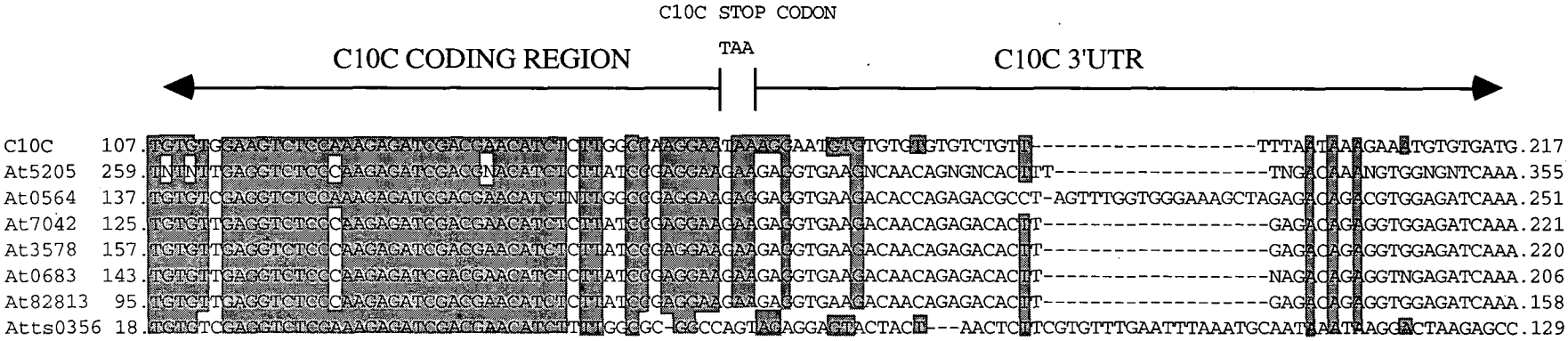
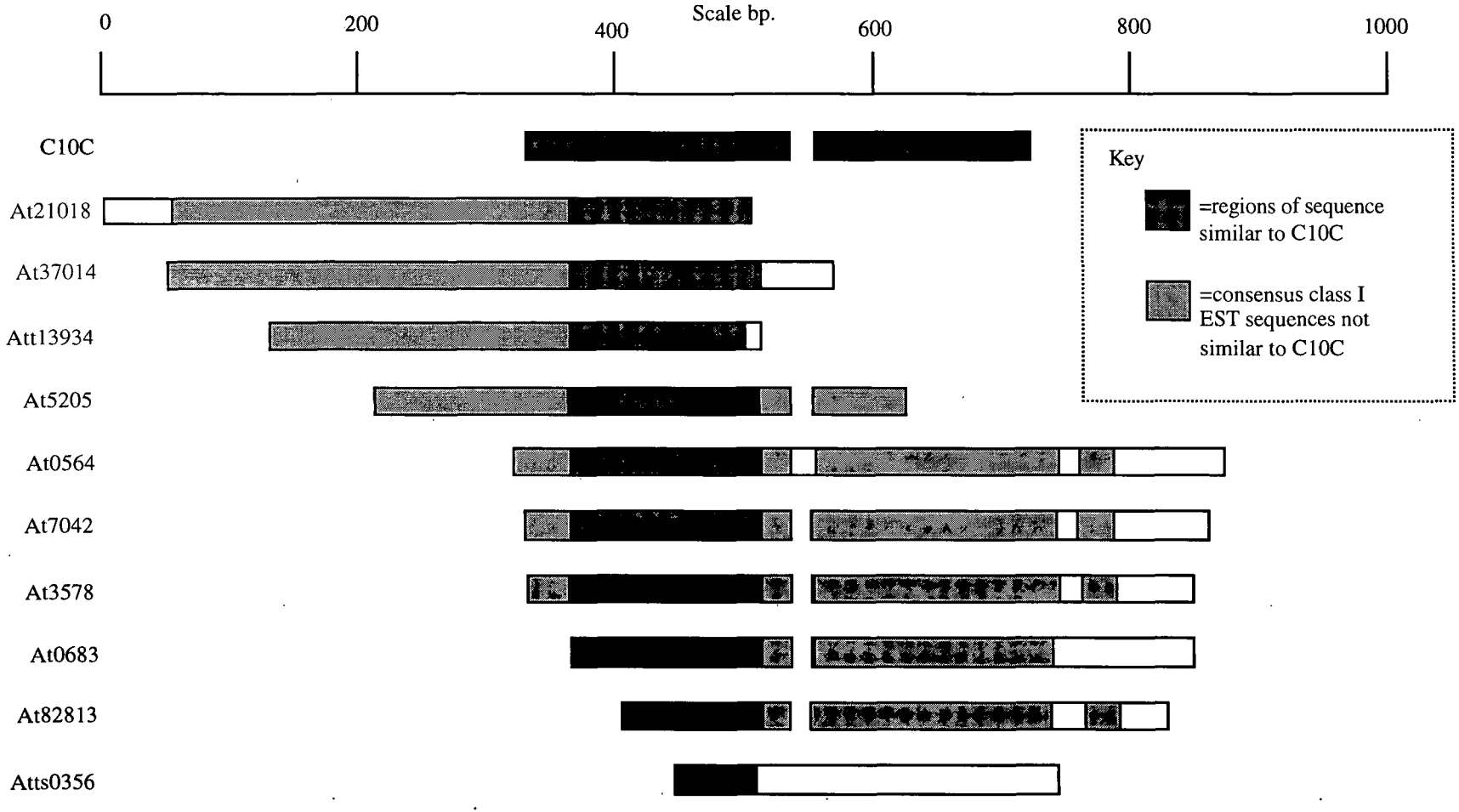
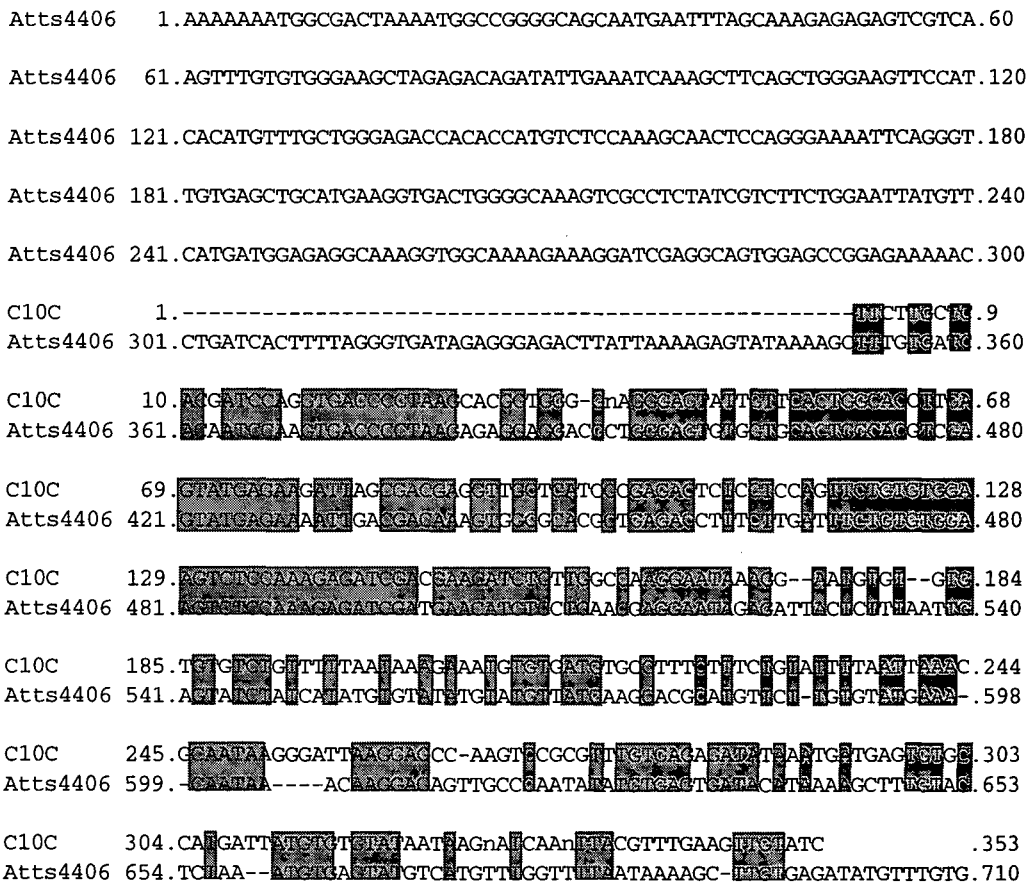


Fig.43 The relationship between C10C and the class I homologous EST's. (a)Alignment of C10C and EST's illustrating the change in levels identity either side of the predicted translation stop codon in C10C. Shading indicates positions where identical nucleotides occur between C10C and the EST sequences. The EST's At37014 and At13934 also overlap this region but where omitted from the alignment as their 3' ends, which are rich in insertions/deletions and unassigned nucleotides as they were derived from the end of the sequencing pass, which cover the above region of sequence. (b)Block diagram illustrating the alignment between the nucleic acid sequences of C10C and the class I EST after a combination of Clustal W 1.5 and manual alignment. Regions of similarity are defined as blocks of 10 nucleotides in which 7 or more show identity. The consensus sequences of the class I EST's were taken as blocks of identity shared by more than three quarters of the clones that overlap that region. The differences in the 3' ends of the class I sequences may be due to actual differences between the cDNA clones or may be due to sequencing errors as the clones were only sequenced with a single pass (Newman et al., 1994).

(b)


Scale bp.
0
200
400
600
800
1000
C10C
At21018
At37014
Att13934
At5205
At0564
At7042
At3578
At0683
At82813
Atts0356
Key
=regions of sequence similar to C10C
=consensus class I EST sequences not similar to C10C

(a) Class 2



(b) Class 3

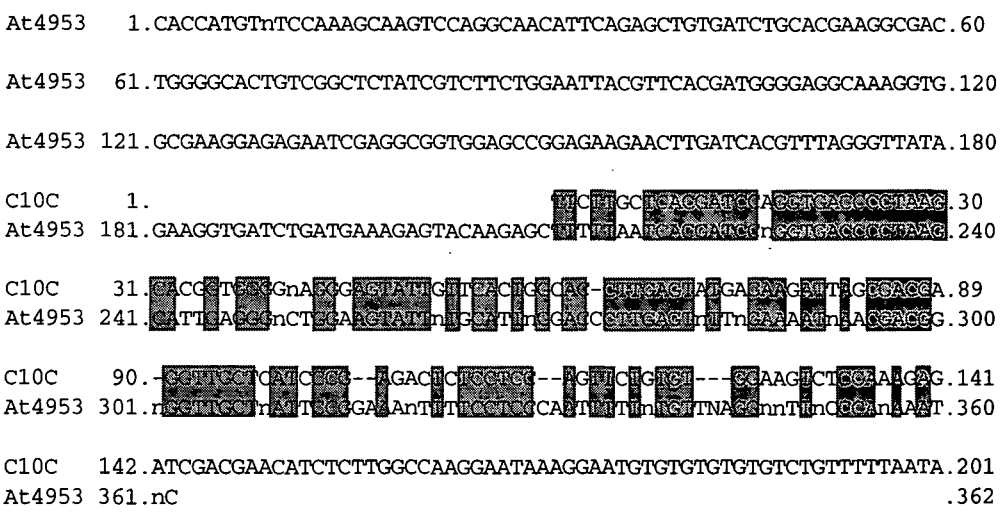


Fig.44. Alignment of the nucleotide sequence C10C with that of the class 2 EST Atts4406 using the Clustal W 1.5 program. Shading indicates identically matching nucleotides.

protein. However, the restricted regions of identity shown by some of the transcript classes make direct comparison between themselves, the MLP genes and C10C difficult. Therefore representatives of each class and subclass were translated and aligned with the MLP and C10C proteins. Clones chosen were MLP1 (*P.somniferum* MLP), Athal1 (*A.thaliana* MLP), At1394 (class Ia 5' sequence), At7042 (class Ia 3' sequence), At0564 (class Ib), Atts0356 (class Ic), Atts4406 (class II) and At4953 (class III). As many of the sequences were single-pass sequenced EST's 50-100nt were removed from their 3' ends before translation as this sequence is likely to be rich in base substitutions/deletions which produce false translation products.

Alignment of these proteins indicates that the sequence of the putative C10C protein most resembles the predicted class I proteins with greater than 95% identity until the C-terminus of the C10C protein, as predicted by the nucleic acid sequences (fig.45). The class III protein is also nearly identical to the class I proteins so may not represent a distinct class at all.

The major difference between the class I proteins and the putative C10C protein is that predicted coding domains terminate at different sites. Atts0356 terminates 5 amino acids upstream of C10C so probably encodes a similar protein but At7042 and At0564 have extended reading frames. These extended reading frames are not frame shifts as they occur at the point where the respective nucleic acid sequences diverge (fig.43a) Both extended reading frames encode nearly identical proteins although At0564 has a 6 amino acid insertion, corresponding to the 18nt. insert in the nucleic acid sequence and the two predicted protein sequences diverge at the point of a frame shift (amino acid 121 in At0564). The predicted At0564 protein terminates just downstream of this frame shift while the predicted At7042 polypeptide continues such that the C-terminal end for that polypeptide is not known. However as these are EST's and the relevant nucleotide sequence is near the 3' end of the sequencing pass this frame shift may be an experimental error. As At7042, At3578 and At82813 have the same reading frame as At0564 over this region (data not shown), then the single base insertion in At0683 relative to At0564 which gives the frame shift is probably the mutation.

The implication of this extended protein is not known but suggests that C10C encodes one member of a protein family with a highly conserved domain linked to absent/present C-terminal domains. In order to investigate this domain it was used to screen the SWISSPROT and PROSITE databases. No matches were found suggesting that the putative domain has no previously been characterised.

The class I EST's also show identity to the *P.somniferum* and *A.thaliana* MLP sequences as well as Atts4406. The region of homology extends into the middle third of the proteins as well as the region which overlaps the putative C10C protein (fig.44). This suggests that, given the high homology of C10C to the class I EST's, that the full length C10C protein may also show homology over this part of the MLP's. However, while this data could be used as further evidence that C10C encodes a protein related to the MLP's, it is not conclusive as the C10C gene sequence may diverge from the class I EST's over this range. Therefore more sequence from the transcript represented by C10C was required. Before work to isolate such a sequence was carried out, some estimate of the length of the transcript was required. Estimation of this was coupled to analysis of the expression of C10C and study of its genomic organisation

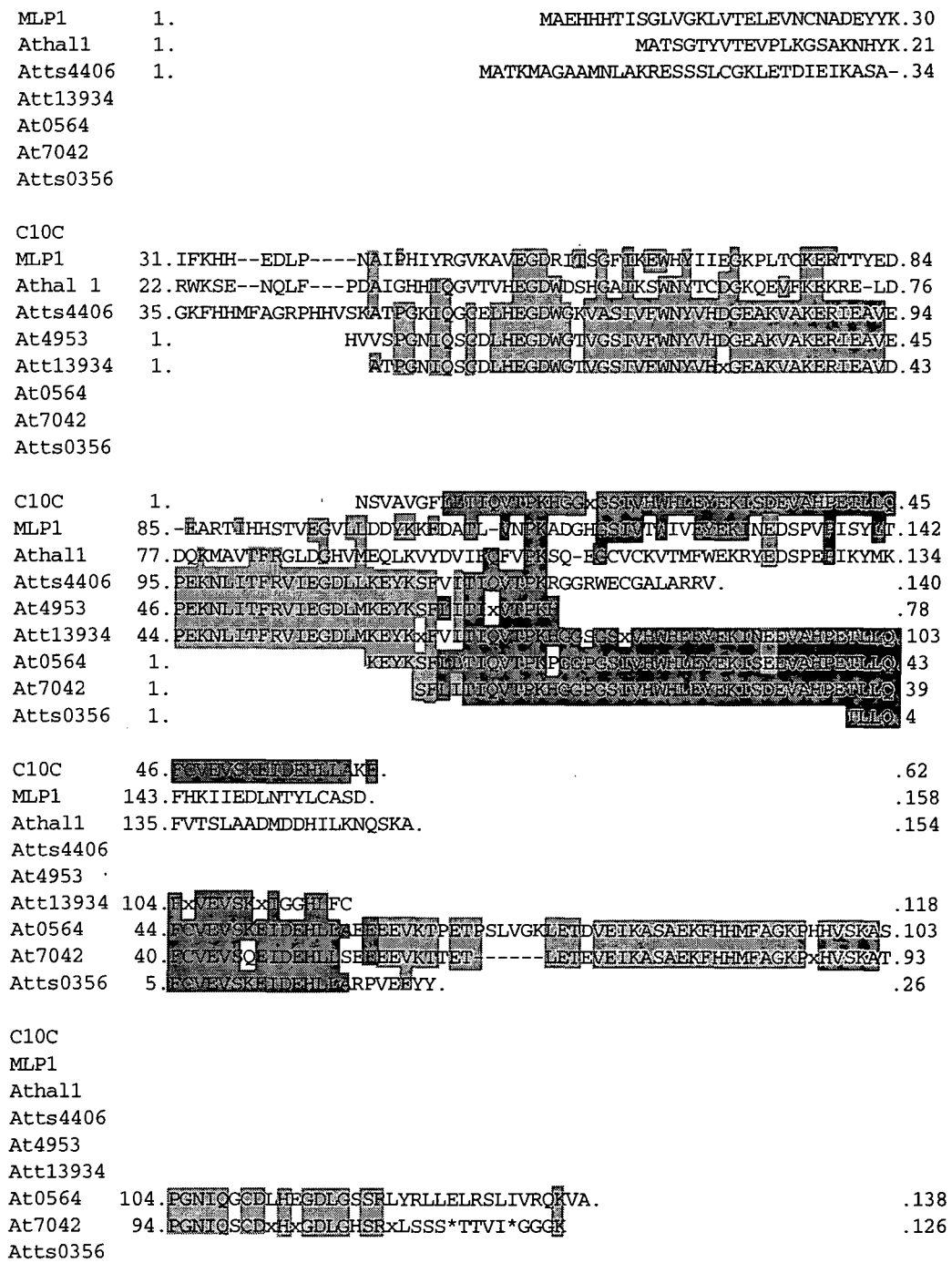


Fig.45. Alignment of the MLP protein sequences with the putative proteins of C10C and selected class 1, 2 and 3 EST sequences. Alignment was performed using the Clustal W 1.5 program. Dark grey shading indicates identical amino acid residues between the C10C protein and the other sequences. Light grey shading indicates consensus sequences between the class 1 EST's and matching amino acids in the MLP sequences. * indicates stop codons.

4.3.2.5 C10C transcript size and expression in the flower.

The size of C10C-homologous transcripts was estimated from a blot of the RNA from different organs of mature *B.oleracea*. This approach was also used to assess the levels of expression. Initially the blot was washed in 1xSSC, 0.1%(w/v)SDS at 60°C which gave an estimated stringency of 78%, sufficiently low to allow hybridisation to the C10C-like regions of the class I EST's. Under these conditions the C10C clone hybridises to transcripts in all the main floral organs while no cross-hybridising transcripts could be detected in the vegetative tissue (fig.46ai). The detected transcripts were all ~920nt. in length, similar to the 860nt. length of the *P.somniferum* MLP mRNA (Nessler and Vonder Haar, 1990) and confirms that the C10C insert represents only a partial clone. This uniformity of transcript size suggests that all the observed transcripts correspond to a single class of C10C-like sequences as opposed to the family of extended sequences as represented by the EST's At0564 and At7042. Such a class of sequences could occur on similarly sized transcripts if it lacks 5' sequence which occurs in the C10C class of transcripts although there is no evidence available to substantiate this possibility.

The level of expression in the floral tissues is not uniform: The highest level of transcript was observed in the petal, while expression in the stigma is much weaker that the other tissues. The reason for this lower level of expression in the stigma is unknown but suggests that the mRNA represented by C10C does not encode a protein involved in a major stigmatic function. In fact the lower level of C10C-like transcript in the stigma, compared to the rest of the flower, suggests a stigmatic repression of expression

A similar pattern of expression was obtained by washing same the blot in 0.2xSSC, 0.1%(w/v)SDS at 65°C, equivalent to a stringency of 91%, and re-exposing to X-ray film (fig.46aii).

The C10C probe was also used against a carpel development blot. Hybridisation was detected at equivalent levels at all stages of development under low (78%) stringency wash conditions (fig.46). A similar pattern of expression was also obtained by rewashing the blot at high (91%) stringency as described above. As with the cross carpel blot all the transcripts were of uniform size suggesting that they don't correspond to class I EST sequences

4.3.2.6 Genomic organisation of C10C.

In order to obtain some estimate of the number of C10C-homologous sequences in the *B.oleracea* genome a genomic southern blot was probed with the C10C cDNA. After EcoRI, XhoI and TaqI digestion only 2-3 bands are visible in each lane (fig.47). For the former two digestions the majority of bands are >8126bp so they may represent single or tandomly arranged multiple copies of the homologous sequences. Whichever these are it suggests that only a limited number of copies of the C10C-like sequences occur in the genome. At the stringency used, these sequences could include class I EST sequences as well as C10C-like genes

The bands in the TaqI lane have a weaker signal than those in the other two lanes. The reason for this is unknown as equal quantities of DNA were digested and loaded into each lane. The low (~600bp and ~300bp) size of these bands could have caused them to become diffuse during blotting though other fragments of this size have blotted effectively under the conditions used (section 4.1.2.2).

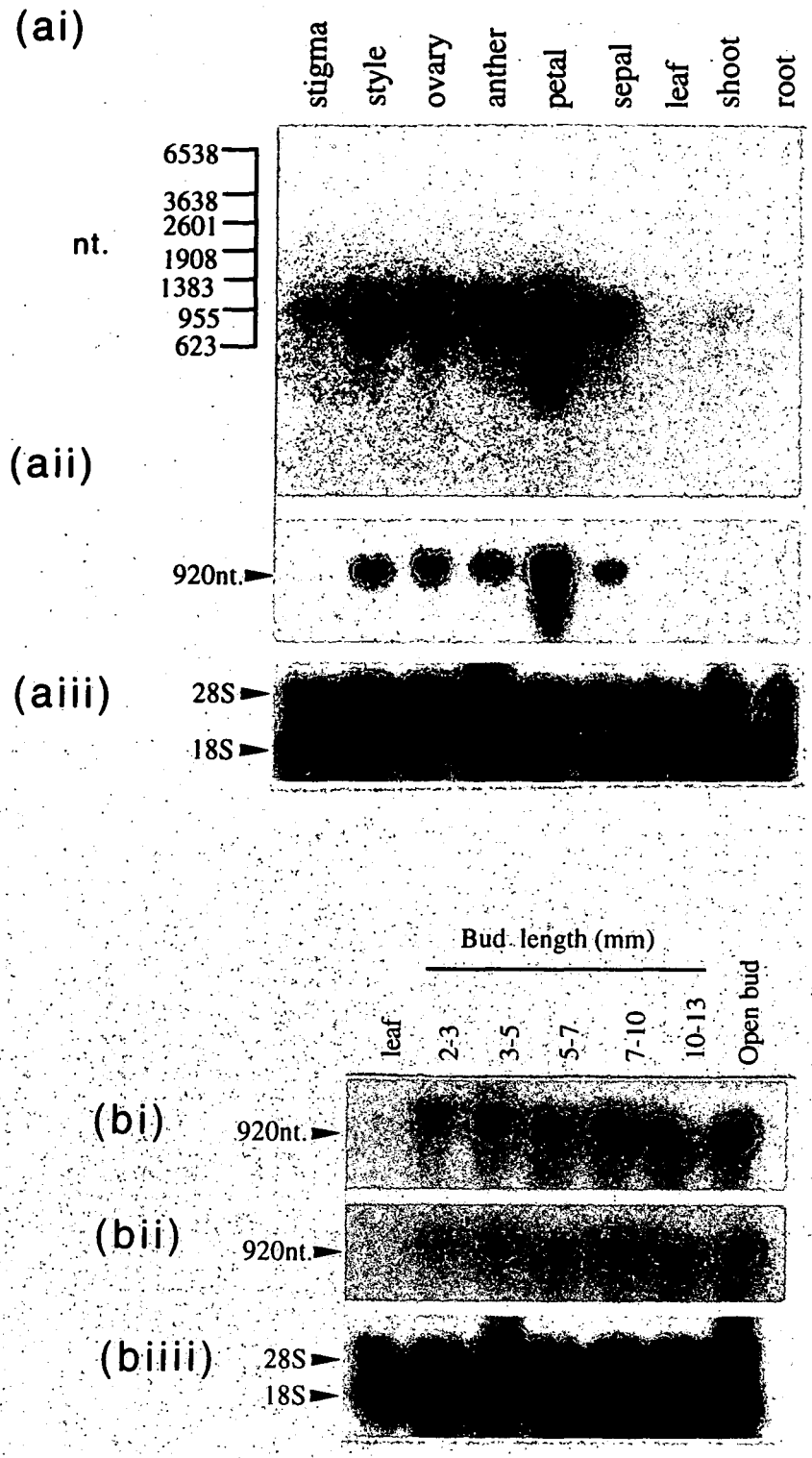
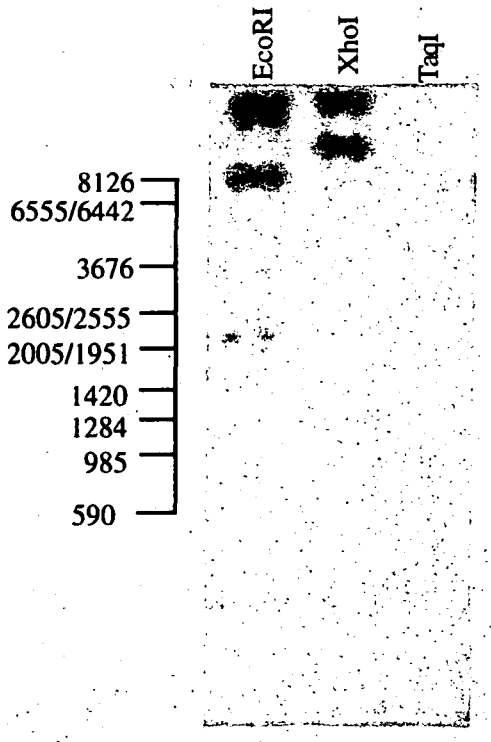


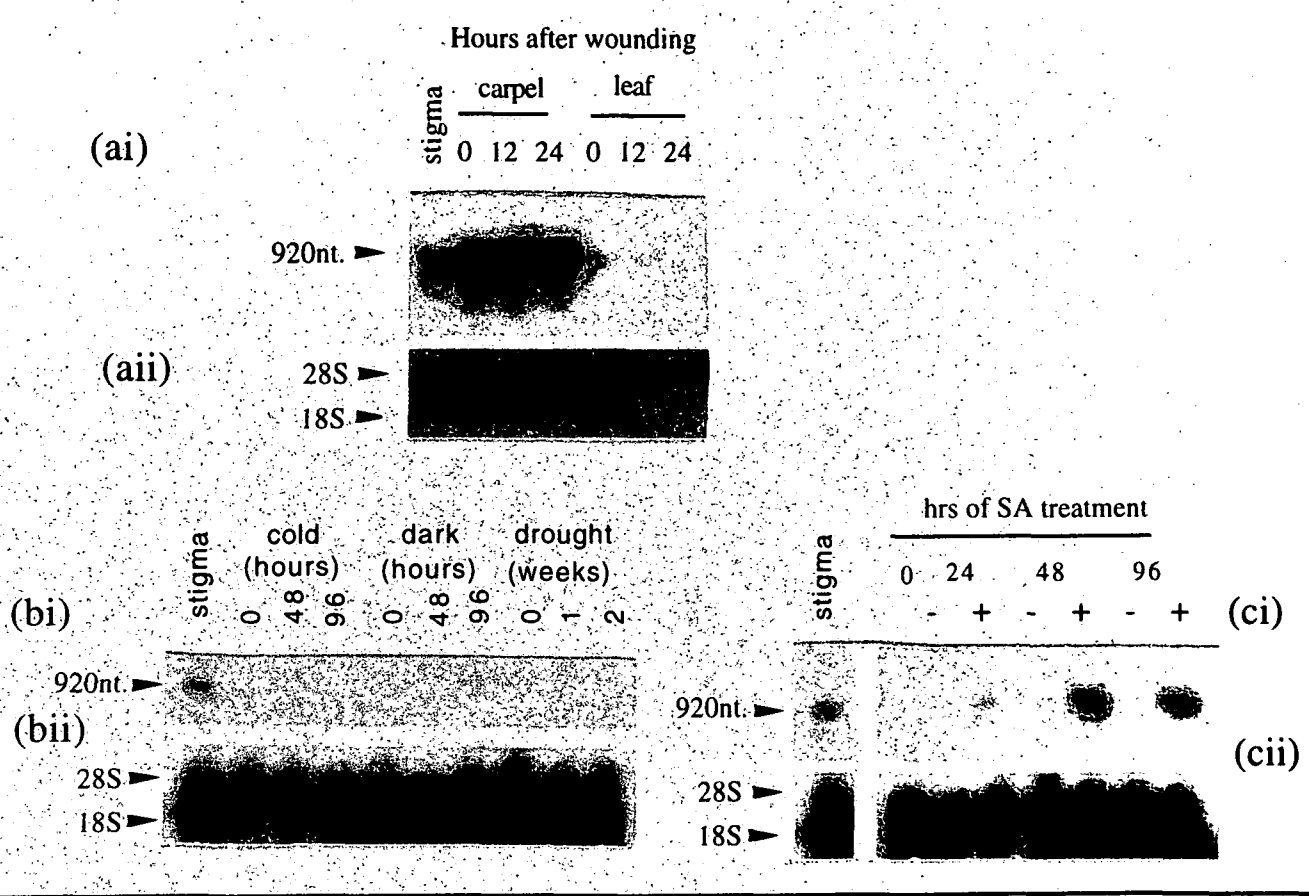
Fig.46. Expression of C10C-like transcripts in the flower of *B.oleracea*. (a) Northern blot analysis of the expression in the organs of mature *B.oleracea*.. (i) Autoradiograph of blot washed in 1xSSC, 0.1%SDS, 65°C (78% stringency). (ii) Autoradiograph of the same blot washed to 0.1xSSC, 0.1%SDS, 65°C (91% stringency. aiii) Autoradiograph of the blot hybridised to the clone pHA1 to confirm equivalent loading of RNA samples (20ug per lane). (b) Northern blot analysis of the expression during carpel development. (i) Autoradiograph of blot (10µg per lane) washed in 1xSSC, 0.1%SDS, 60°C (78% stringency). (ii) Autoradiograph of the same blot washed in 0.1xSSC, 0.1%SDS, 65°C (91% stringency). (iii) Autoradiograph of the same blot hybridised to the clone pHA1.

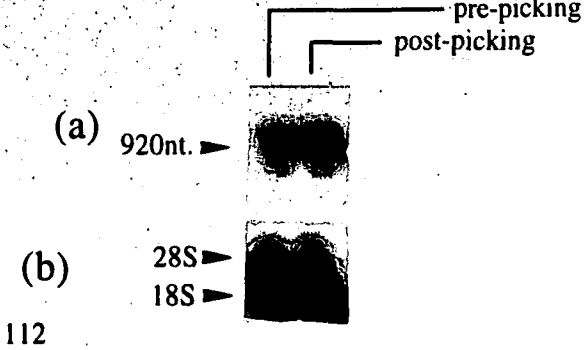
Fig.47. Autoradiograph of a *B.oleracea* Southern blot probed with the C10C probe. 5µg of genomic DNA was loaded into each lane. The blot was washed to 1xSSC, 0.1%(w/v)SDS, 65°C which gives a predicted stringency of 84%. The size markers were derived from AvaII digested lambda DNA.

Fig.48. Northern blot analysis of the effect of stress on the expression of transcripts similar to C10C in *B.oleracea.* 10µg of RNA were loaded into each lane. (a) Autoradiograph illustrating the effect of wounding on expression in the carpel and leaf. (c) Autoradiograph illustrating the effect of heat, light and water deprivation on seedlings. (e) Autoradiograph illustrating the effect of salicylic acid on expression in the leaf.(ai), (bi) and (ci) were washed in 1xSSC, 0.1%(w/v)SDS at 60°c. (aii), (bii) and (cii) are the autoradiographs of the blots in (ai), (bi) and (ci) after hybridisation to pHA1 to determine equivalent loading of RNA samples.

Fig.49. Northern analysis of the effect of bud harvesting on the expression of C10C-like transcripts in the buds of *B.oleracea.* (a) Autoradiograph of blot probed with C10C. (b) Autoradiograph of the same blot as in (a) probed with pHA1to confirm equal (10µg) loading in each lane. (a) was washed in 1xSSC, 0.1%(w/v)SDS at 60°C.


EcoRI
XhoI
TaqI
8126
6555/6442
3676
2605/2555
2005/1951
1420
1284
985
590


Hours after wounding
(ai)
stigma
carpel
leaf
0 12 24 0 12 24
920nt.
(aii)
28S
18S
(bi)
stigma
cold (hours)
0 48 96
dark (hours)
0 48 96
drought (weeks)
0 1 2
920nt.
(bii)
28S
18S
hrs of SA treatment
stigma
0 24 48 96
- + - + - +
(ci)
920nt.
(cii)
28S
18S


pre-picking
post-picking
(a)
920nt.
(b)
28S
18S

4.3.2.8 Expression of C10C-homologous genes under stress conditions.

As the putative C10C protein is homologous to a protein with a defensive role the effect of stress on expression of C10C-like transcripts in the vegetative organs of *B.oleracea.* was assayed. No effect on expression could be detected, at low stringency, by wounding of leaves or by cold, dark or drought treatments on seedlings as indicated in fig.48a and b. Expression could be induced by treatment with salicylic acid (fig.48c), a physiologically naturally-occurring inducer of systemic acquired resistance (SAR) (White, 1979) and the relevant PR proteins (Ward et al., 1991). Therefore C10C, like 13G6(i) represents a gene family that is both stress-inducible and present in floral tissue and is analogous to the solanaceous stress genes in that it is similar (or identical) to a PR sequence. Unlike these latter sequences, however, C10C is expressed in all the floral organs rather than being restricted as with the chitinases and ß (1,3)glucanases (Lotan et al., 1989; Leung, 1992; Wemmer et al., 1994 and Harikrishna, 1996).

As with the floral transcripts, all the salicylic acid-induced transcripts are the same size as the stigma transcript. Therefore these are also unlikely to be similar to the extended class of transcripts represented by At0564 and At7042.

As salicylic triggers the expression of stress genes in healthy tissue in response to attack on another part of the plant then, as with 13G6(i), floral expression of C10C could be an artefact induced by picking buds. Therefore the RNA isolated from pre-picked and post-picked buds was probed with the C10C cDNA to determine if floral expression is induced by the way the plants were handled. No apparent difference between the two populations of buds could be detected (fig.49) indicating that the level of C10C in whole flowers was not affected by the way the plants were handled. As only whole buds were used changes in the distribution of the C10C expression can not be ruled out although as expression was very similar in all the organs assayed, except the stigma, co-ordinated changes in expression that maintain the overall expression level in the whole flower are unlikely.

4.3.2.9 Isolation of a longer C10C-homologous cDNA clone.

As the C10C clone appeared to be weakly expressed in the stigma, it was predicted that a large number of clones might have to screened. Therefore, in order to obtain a longer length clone, with the minimal number of rounds of plaque purification required two plaque densities (~8000 and 50000pfu per plate) were plated out on 82mm diameter petri dishes so that if sufficient clones from the lower density were available these could be used. Initially hybridisation was performed at 65°C in the standard hybridisation buffer (section 3.6.4) then washing done to 1xSSC, 01%(w/v)SDS. This caused apparent hybridisation to all the plaques so that the true positive clones could not be discerned. Washing at higher stringency failed to alleviate this high background (data not shown). therefore the plating was repeated on bigger plates (137mm diameter) to reduce the density of non-target λ DNA and hybridisation performed at 68°C. Background was still present on these plates but positive clones could still be identified. Twelve positives were identified on the 8000 pfu plate while ~40 were identified on the 50000 pfu plate where background was higher and so positives more difficult to discern. Therefore the twelve 8000 pfu-plate

positives were picked as 0.5mm diameter plugs. Division of the plating area by the area of a single picked plug indicates that ~10 different clones should occur in each plug so that only one round of plaque purification would be needed to obtain pure clones

In order to identify the largest C10C-similar clone a PCR-base approach was adopted. The clones were to be amplified using primers based on the SP6 and T7 primer sites either side of the cloning site in λ ExCell and then identified by blotting and hybridisation to a C10C insert probe. To test the feasibility of PCR amplification direct from eluted λ ExCell phage a trial experiment, using three purified phage plug eluates, C10C, C10A and D5A2, as templates for amplification, was performed with an extension time of 1.5 min. Discrete bands corresponding to the predicted sizes of these clones were obtained (data not shown) indicating that amplification could be carried out on unpurified phage.

The twelve selected positives were amplified by this method and run on a gel. With the exception of clone12, from which no products were obtained, all extracts gave a smeared product that contained some just-visible, diffuse bands (fig.50a). This contradicts the expected pattern of approximately 10 bands, predicted from the number of clones per plug. Therefore some component in the phage eluant appeared to be interfering with the PCR or degrading the DNA that was not present in the pure clones. The trial clones had been preserved in SM/chloroform at 4°C for five months whereas the C10C plugs had been stored for only three days. To determine if the age of the phage eluant affects its amplification the two phage from the primary differential screen (~1 year old) were amplified together with overnight eluants from plugs from a one day old plate using a 1 min. extension time. Discrete bands were obtained from the former but not the latter, confirming that the age of the template phage affects its suitability for PCR by this method.

As some banding was still visible in the products from the C10C positives, this gel was blotted and probed to determine the size of any C10C-similar amplification products. All the amplification products gave positive signals, with positives 9, 10 and 11 having producing the largest hybridising bands with sizes of ~850bp (fig.50b). From this screen positive 5 gave a band that hybridised much more strongly than those in the other clones, although this had a size of only 450bp.

Therefore positives 5, 9 and 10 were replated at a density of approximately 100pfu per plate at which individual clones could be identified and picked. Hybridisation of this blot and washing as above produced a blot in which >90% of the clones registered positive (data not shown), even after washing at 0.2xSSC, 0.1%(w/v)SDS (93% stringency). Selection of one positive from each plate, DNA purification and digestion indicated that all the clone had inserts of <500bp (data not shown). Therefore these clones were unlikely to represent homologues of C10C.

The conclusions from this result is that the C10C probe is hybridising non-specifically to these clones and PCR products. The reason for this is unclear given that apparently specific hybridisation occurred with the Southern and Northern blots. Under the conditions used hybridisation stringency was 84% while washing stringency was 93% so non-specific binding of the insert should have been prevented. Two factors shared by all the inserts are poly(dA) tails and the EcoRI adapter sequences at the 5' end of the clone. The former should not produce non-specific hybridisation as poly(dT) was included as a blocking reagent in the mixture. The latter should not hybridise in the hybridisation solution used as their T_M can

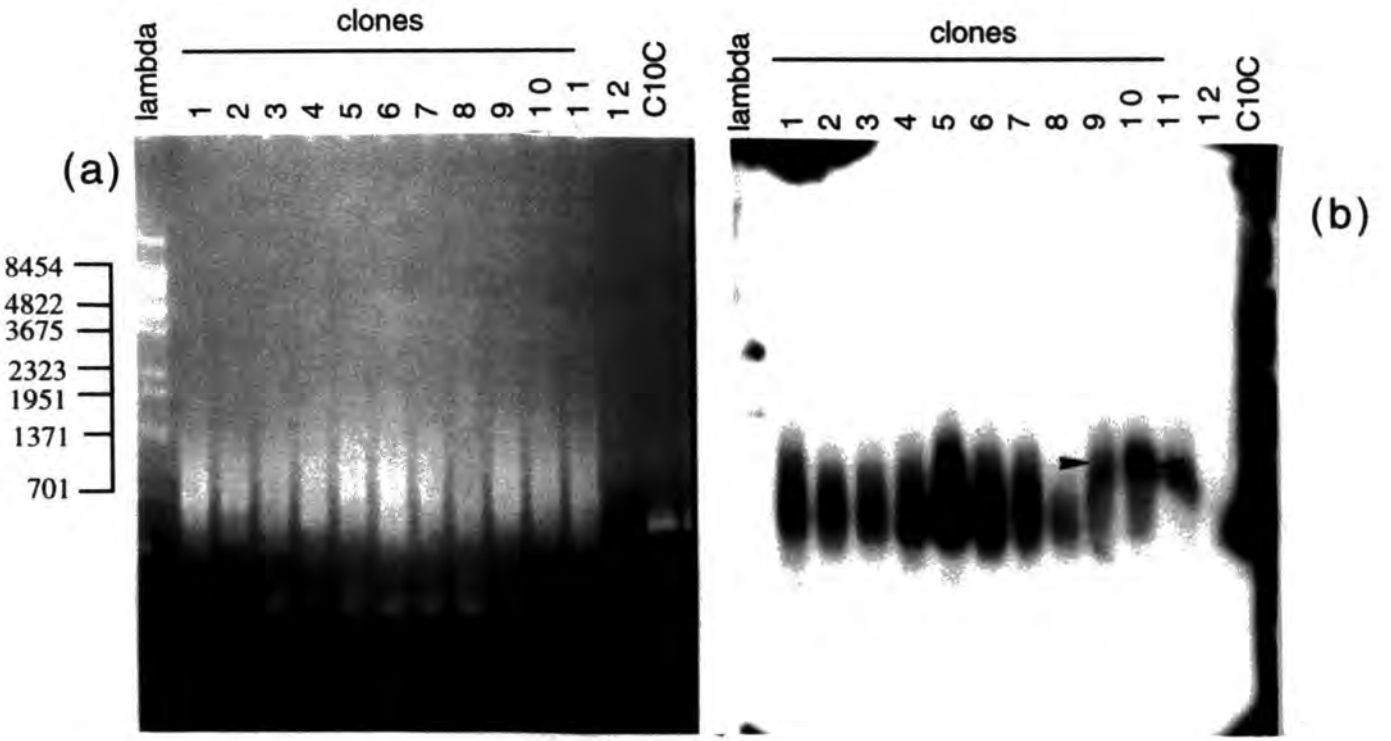


Fig.50. PCR of C10C-cross-hybridising clones. (a) DNA gel stained with ethidium bromide after electrophoresis of PCR products obtained by amplification from phage indicated by lane numbers. (b) Autoradiograph of the blot taken from (a) after hybridisation to a C10C probe. Arrows indicate PCR products derived from phage 5,10 and 11 which were lead to further purification of these clones. Size markers are lambda Bst EII

be calculated as 81.5+16.6log(5x0.195)+0.41(9/16x100)-(500/16)=73°C. Therefore hybridisation at 65°C is well above TM-20°C. Similarly washing at 0.2xSSC, 0.1%(w/v)SDS would require the temperature to be 49°C to maintain any hybridisation. Therefore the adaptor sequences are unlikely to contribute to the observed hybridisation. A final cause of the cross-hybridisation is contamination of the cDNA probe with vector sequence. Although the insert from which the probe was derived was cloned in pKS+Bluescript while the library was in λ ExCell, these two vectors share regions of sequence identity (Amp^r and Ori^c sequences occur in both the plasmid and the phagemid region of the lambda vector) which would allow cross hybridisation.

Due to time limitations, a full study of why non-specific hybridisation was occurring and what could be done to prevent it could not be carried out. Therefore further work on determining the full length sequence of the transcript from which C10C was derived was not carried out. Therefore the relationship between C10C and the poppy latex protein genes is still unclear.

As the above strategy was unsuccessful different approaches could be taken. PCR could be used with primers that bind specifically to the C10C insert could be used on both the library or stigma/floral mRNA. Direct PCR using a C10C primer and the SP6 primer on the library could be used although given the aberrant hybridisation of a C10C insert probe this may produce false positives again. A C10C primer could also be used in a 5' RACE reaction, which given the apparently specific binding of the C10C probe on Northern blots would be predicted to have more chance of success than amplification from a library. If these fail the isolation of genomic clones could be used as a last resort as C10C appears to represent a member of a small gene family. However if this approach was used reprobing of Northern blots with the 5' sequence of any clones would have to be carried out to confirm that they correspond to the same transcript as C10C.

4.3.2.10 Conclusions for C10C.

The data available for C10C is insufficient for many conclusions to be made concerning it. Despite its isolation from a stigma library, the pattern of the clone's expression suggests that it does not play a major stigma-specific role and so at first glance is not directly of interest, given the aims of this project. However the induction of similar transcripts after salicylic treatment suggest that C10C can be classified with 13G6(i) as a member of a gene family that is expressed in healthy flowers but is induced by stress in vegetative organs. The full implications of this are discussed in section 5.2.

The similarity of C10C to the MLP's is of interest not only in *Brassica* flower biology, where other MLP's such as those described in section 4.3.2.2. occur but also in its implication for the biology of defence. If C10C does encode a latex protein then this is the first report of such a sequence being expressed as part of the SAR response. Although both lactation and PR gene activation are pre-emptive mechanisms, in that expression of the defence proteins occurs before the organ expressing them is attacked, the former is constitutive while the latter requires attack on another part of the plant for activation. Similarly lactation has been principally implicated in wounding defence while PR genes are principally involved in defence against microbial attack (Ward et al., 1991). These different roles are mirrored by the initial cytoplasmic localisation of latex while most PR proteins predominate in the

extracellular matrix. Nevertheless similarities between the two are becoming apparent. Some PR enzyme activities such as chitinases (van Parijs et al., 1991) and ß-1,3-glucanases (Churngchow et al., 1995) have also been detected in the latex of *Heavea brasiliensis* although the prior treatment of the plants in which activity was detected may have been responsible for the observed expression (Churngchow et al., 1995). A reverse situation, where latex protein genes are expressed with PR proteins in SAR is therefore not unreasonable although there are no reports of the formation of latex as part of the PR process.

Even if C10C does not encode a latex gene it represents a novel SAR gene as no homology with previously characterised PR sequences was indicated by the database searches. Sequencing of the PR proteins of *A.thaliana* (Uknes et al., 1992) indicates that they show high homology (>50% identity) to the respective tobacco PR families. Therefore C10C is unlikely to encode a homologue of a previously characterised tobacco PR protein whose sequence has diverged sufficiently to prevent identification by alignment.

From this data, C10C also appears to be part of a structurally divergent gene family. The high degree of identity between C10C and the class I EST's indicates that they must be related. However very different protein sequences are produced by the EST's At0564 and At7042 which have >60 amino acid extensions at their C-terminal ends compared to the putative C10C protein. It is notable that all the EST's which were sequenced around the point of divergence appear to have the extended domain suggesting that all the *A.thaliana* sequences are of this extended class of transcript, rather than like C10C. This bias could be interpreted in two ways:

(i) the class I EST's represent a different class of transcript to that represented by C10C. This class of transcript is more highly represented in the library from which the EST's were isolated.

(ii) the class I EST's represent the *A.thaliana* equivalent of the *B.oleracea* transcript encoded by C10C.

Theoretically, a difference of 60 amino acids between equivalent proteins of two such closely related species as *A.thaliana* and *B.oleracea* is unlikely, especially as the regions of protein present in both classes of sequence are so conserved (fig.45). However floral tissue was a component of the library used by Newman et al. (1994) so C10C-like transcripts would be predicted from their sequencing.

Discrimination between these two possibilities could be achieved simply by the probing of a Northern blot containing *Arabidopsis* and *Brassica* floral tissue with either C10C or one of the class I EST's as a probe. If transcripts in equivalent tissues in the two species have different sizes with the *Arabidopsis* form being suitably larger then this would suggest that the extended transcript is the *A.thaliana* form of C10C. However if transcript in equivalent tissues are the same size then this would suggest that C10C and the relevant class I transcripts do represent different genes. A PCR approach could also be used as described in section 5.2.3.

4.3.3 Stigma enriched clones

4.3.3.1 K8A2 and F4B are predominantly expressed in the stigma

Screening of the cDNA library identified the clones K8A2 and F4B which both cross-hybridised with the stigma RNA probe but showed no detectable hybridisation with the leaf probe. The cDNA inserts were prepared from the purified clones, sequenced and used to probe Northern blots. These blots were washed in 1xSSC, 0.1%(w/v)SDS at 60°C (fig.51). The results from these experiments confirmed that transcripts similar to both clones were enriched in the stigma, not only when compared to vegetative tissue but also to the other organs of the mature flower. Expression is not exclusive to the stigma for either class of transcript as relatively weak bands were also detected in the style, anther, sepal and petal of F4B and in all the organs assayed for K8A2. This low level of expression could be the result of lower expression than seen in the stigma or by hybridisation to transcripts only weakly similar to the two clones. For both clones, the weak transcripts have the same length as those in the stigma so suggesting that the former suggestion is more likely.

Expression of K8A2-like transcripts in the leaf would appear to contradict the expression indicated by the differential screen. However as expression in the leaf was much lower than that in the stigma it may have been too low to produce a detectable signal when using the leaf probe in the differential screen.

The stigmatic transcripts, homologous to K8A2 and F4B were estimated to be ~2800nt. and ~1450nt. long respectively. As the inserts in the clones were only ~800bp and ~500bp long they represent only partial sequences. Despite this, the clones were both sequenced to determine whether they contained any sequence information that could indicate biological activity or function for their encoded proteins.

4.3.3.2 Sequence of K8A2.

The sequence K8A2 was initially sequenced from each end which gave the full sequence although not completely on both strands. To complete the sequencing on both strands, a Sac1 site, identified by the preliminary sequencing, was used to generate two subclones from which the full sequence was obtained (fig.52a).

The K8A2 insert is 798bp in length and represents the 3' end of its respective transcript as it has a 21bp polyA tail which allows orientation of the clone. No exact matches with the polyadenylation consensus signal are present but two close matches are present at nucleotide positions 717 and 761, the latter of which is sufficiently near the poly(A) site to act as a possible signal motif. A match with the partial far upstream element sequence is also present at position 702 (fig.52b).

ORF scanning reveals no obvious reading frame (fig.52c) as in each frame stop codons occur near the 5' end of the clone. Therefore as the clone represents only ~30% of the full size transcript K8A2 may consist predominantly of 3'UTR. In one reading frame however, only one of the first 155 codons is a stop codon. This would appear to indicate that this stop codon may be a mutation caused by misreading during transcription or the cDNA synthesis/cloning procedure. To assess whether this might be so the nucleotide sequence was analysed for codon bias.

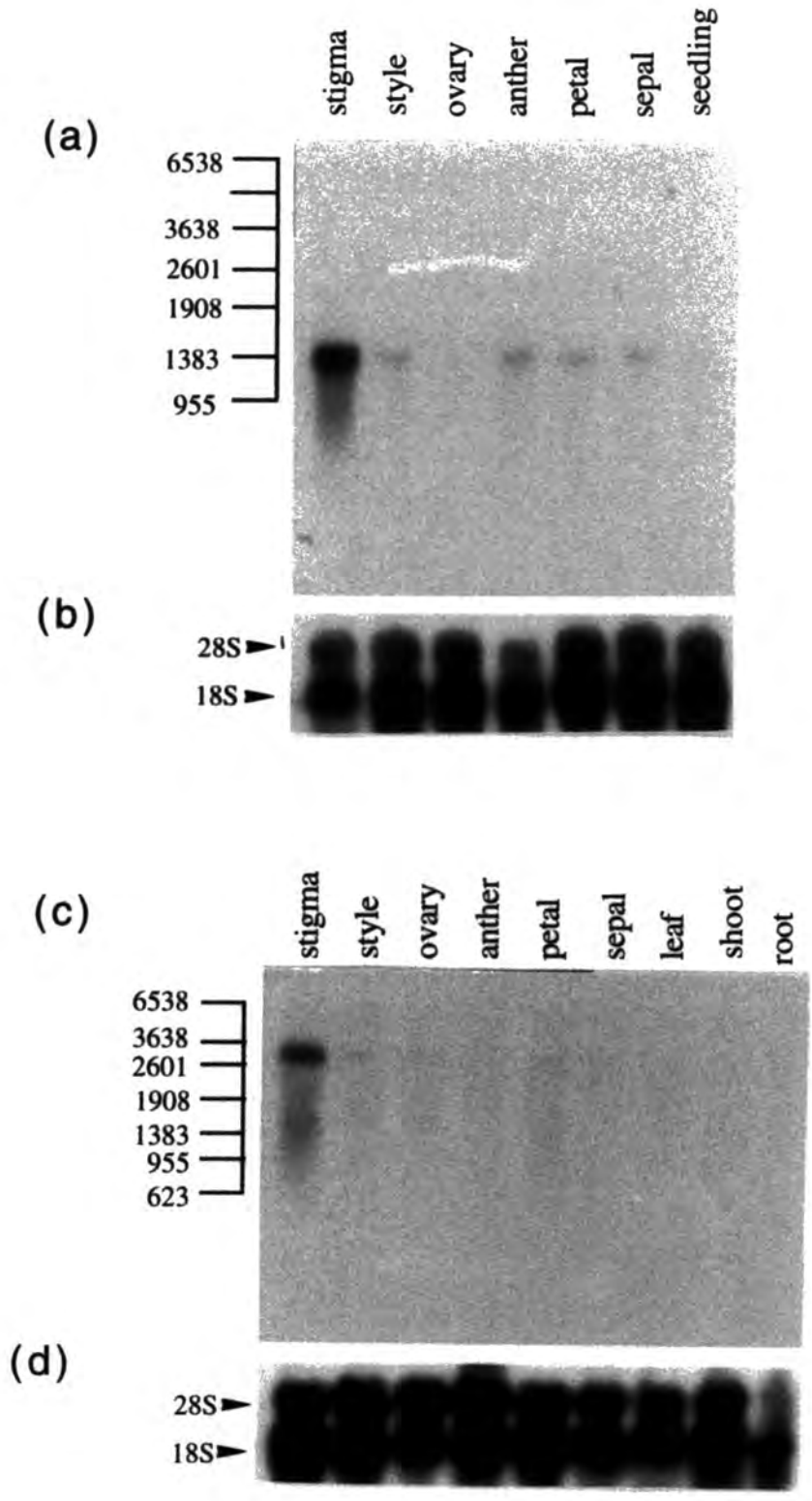
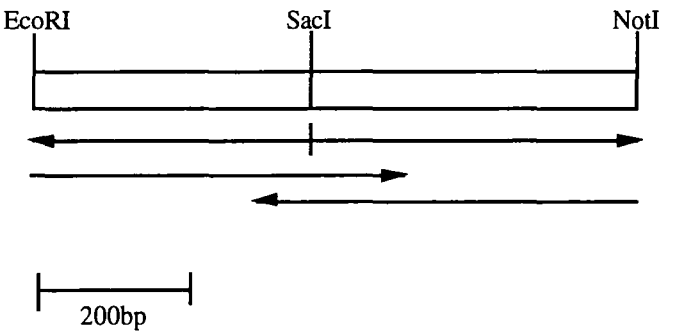


Fig.51. Expression of transcripts similar to the clones K8A2 and F4B in the major organs of mature *Brassica oleracea*. (a) expression of F4B-like transcripts. (c) expression of K8A2-like transcripts. (b) and (d) are the blots used in (a) and (c) probed with the pHA1 clone to confirm equivalent loading of RNA. (a) and (c) were washed to 1xSSC, 0.1%(w/v)SDS, 60°C which gives predicted stringencies of 76% and 78% respectivly. All lanes contain 10µg total RNA

(a)



(b)

```
  1.GCGGCCCCAGGAATTTTTTTTGTGGTGAATGCAGAGTTCTTGTTCAGTTTTGGTGTAGTTATGGCACTGTGCCTCTCAAT.80
 81.GTTATTGTCACTCAGGTTCTTCTTAACCTTTCATATATCTTCAACTTCCAAAATTTTAAGTCTATGAAGGTCTATGACTA.160
161.ATCATTGTTCATCTATTGGATTCACCTGAATGTGAACAGATGGGGTCTCGGTGCAAGAAGCTGTGATAGCAGAGAATGTA.240
241.AGAGACTCACTTCACAGCTGGTGCAAGAGAGTGAAAGAGAGGTCAAAGCACGCGAGATCAGTGTGTTCACTAGACACAGC.320
321.AACAATAGACGAGAGAGACGAGATGACAGTGGGAACATTGTCAAGGAGCTCATCCATGACCTCGCTGAATCAGATCACCA.400
401.TAAACTCCACAGACCAAGCAGAGTCCATATTCGGAGCAGCAGCTTCATCGAGCAGTCCTCAAGATGAATACACTTCGAGA.480
481.GTCGAAGAATATCTGTCTGAAACATTCAATAACATCGGTTCGATGCAACCTTTAAACGATGAGATTGAGATTGAGGTTGA.560
561.AAGTGAAGAAGATAATGAAAGGCGAGGGAGTGAGAGTGAAGAGACTAATGGTGAAGGTGGAGAAACACTTTTTGATTTGT.640
641.TTAGGAGGACTTGATCTTGTTTCTTCTTCATCTTCTACTTCCTTATTGAAGCTTGGTAAAGTTGTAATTACATGACAATA.720
721.ACATCCATAAAAGCAAAAACTACTGTAACAATGTAACAGAGTAAATCTTTTGTTGGTAAAAAAAAAAAAAAAAAAAAA .798
```

(c)

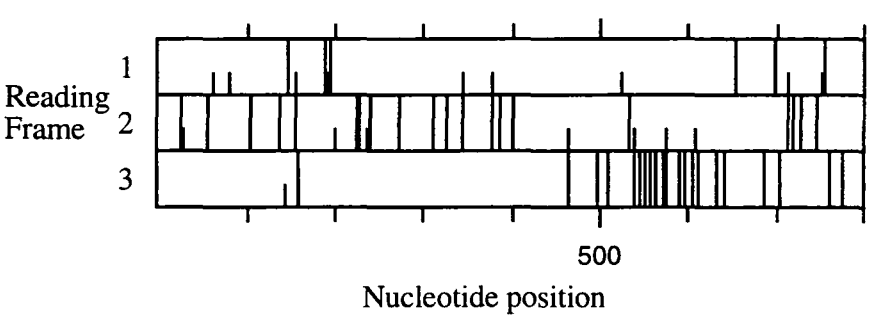


Fig.52. Sequence of the clone K8A2. (a) sequencing strategy used to obtain the full sequence of K8A2 on both DNA strands. Arrows indicate the direction of sequencing while the length of each bar indicates the extent of each sequencing reaction. (b) Full nucleotide sequence of the K8A2 cDNA. The 5/6 matches to the poly(A) signal sequences is singly underlined while the partial FUE consensus is doubly underlined. (c) ORF map of the clone K8A2. Half height bars indicate ATG codons while full height bars indicate possible stop codons.

	RF1 1-781	**RF2 2-781**	**RF3 3-161**	**RF3 162-487**	**RF3 488-781**
%NNG/C	40.0	37.0	44.3	52.0	38.4
%NTA	6.1	5.0	0	0	4.8
%NCG	2.3	2.7	0	0	1.9

Table 8. Codon bias table for the cDNA sequence K8A2.

Bias in the sequence of dicot codons has been reported by Campbell and Gowri (1990). By analysing the choice of codons in characterised dicot genes they calculated that C or G occurs as the third nucleotide in the codon at a frequency of 45%. Preferentially avoided codons, in these plants, have the sequence NTA (3.2%) and NCG (1.8%). While these figures are approximate for all dicots some differences occur between different families (Campbell and Gowri, 1990). For the Brassicaceae the NNC/G frequency is 51.9% while the NTA and NCG have frequencies of 3.0% and 1.9% respectively. Therefore these latter figures were taken as guidelines for *Brassica* coding regions.

To assess the possible coding capacity of the reading frames the frequency of the three indicator codon classes were determined in the sequence upto the first stop codon (nt. 2-161), from the first stop codon to the second stop codon (nt. 162-467) and from the second stop codon to the 3' end of the clone, not counting the poly(A) tail which would act as a biasing factor. As controls the complete first and second frames were also analysed so as to rule out any bias in composition of the sequence affecting the codon frequency. The results are presented in table 8. These suggest that the sequence both 5' and 3' to the first stop codon in RF3 has a similar bias to that seen for coding regions in the Brassicaceae. The predicted 3' UTR in RF3 and the other two reading frames do deviate from the Campbell and Gowri standards in that their NNG/C frequencies are low while the NTA frequencies are all high. Therefore the predicted coding region of K8A2 runs from nt. 3-467 with the internal stop codon being due to a putative erroneous base in that codon. This gives the protein sequence as indicated in fig.53.

While this sequence has a high chance of being an accurate representation of the protein encoded by the K8A2-like transcripts further confirmation is required. One approach would be to isolate other cDNA sequences representing the same class of transcript and determine if they too have a stop codon at the equivalent position to nucleotides 159-161 of K8A2. Isolation of cDNA's longer than K8A2 would also allow more sequence information to be obtained on this class of clone.

The nucleic acid sequence of K8A2 were used to screen the EMBL sequence library. No similar sequences were identified indicating that K8A2 is novel. The SWISSPROT database was screened with both the predicted RF3 protein and the protein encoded by the first 144nt. of RF1 as although the codon composition suggests otherwise, the latter could represent the K8A2 protein. The protein encoded by RF2 was not used as the coding region from the start of the clone in this reading frame is only 8 codons long. No significant matches were found suggesting that, if either reading frame is correct, K8A2 encodes a novel protein

```
  1.GPRNFFCGECRVLVQFWCSYGTVPLNVIVTQVLLNLSYIFNFQNFKSMKV.50
 51.YD*SLFIYWIHLNVNRWGLGARSCDSRECKRLTSQLVQESEREVKAREIS.100
101.VFTRHSNNRRERRDDSGNIVKELIHDLAESDHHKLHRPSRVHIRSSSFIE.150
151.QSSR                                                .154
```

Fig.53. Predicted protein sequence from the clone K8A2. The stop codon inside the sequence, which is predicted to be a mutation, is indicated by *.

4.3.3.3 Sequence of F4B.

The F4B insert was only 500bp in length so was sequenced via vector primers from each end to give the full sequence on both DNA strands. The clone insert is 467bp long and represents the 3' end of its transcript as there is a 23bp polyA tract at the 3' end. As with K8A2 no match to the consensus polyadenylation signal site is present although three copies of an element proposed to act as a far upstream element in the 3'UTR processing of pea *rbcS*-E9 gene (Mogen et al., 1992) are present (fig.54a).

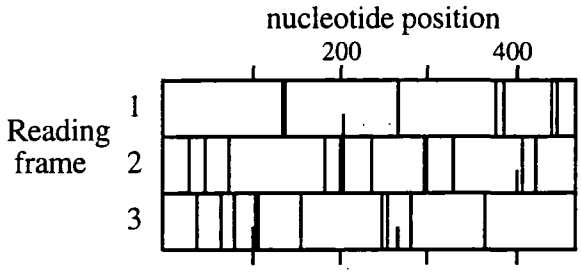
ORF screening indicated that the clone did not contain an open reading frame greater than 70bp. (fig.54b). This is not surprising as the clone represented the 3' end of the transcript from which it was derived, a proportion, if not all of which would consist of untranslated message. The three possible reading frames were therefore screened from the 5' end of the clone to the first in frame stop codon, for their potential to be a *Brassica* coding sequence using codon bias tables (Table 9). None of the three reading frames appeared to be coding suggesting that the entire clone may be 3'UTR. However, as all the reading frames are small (10-45 codons) this could bias their codon frequencies so no firm conclusions can be drawn from this data.

Screening the EMBL nucleotide database with the F4B sequence revealed one sequence, Atts4051, which showed some similarity to it. This sequence is an *Arabidopsis* EST (Phillips and Gigot, 1994) derived from sliced *A.thaliana* leaves maintained in liquid media for 2-4 days and has 41% identity to the relevant region of F4B (fig.55a). The level of identity is low for equivalent sequences from such closely related species so Atts4051 may represent a different gene. The level of identity could be affected by the possibility that if the two clones predominantly represent 3'UTR then significant divergence could occur between the *Arabidopsis* and *Brassica* forms of a gene. Although most EST sequencing is performed from the 5' end of the cDNA inserts, the Atts4051 sequence is quoted as being 3' sequence (Phillips and Gigot, 1994). Conformation that this is the case comes from ORF screening which indicates that stop codons occur throughout all 3 reading frames in the clone (fig.55b). Therefore it is probable that Atts4051 represents 3'UTR as does F4B

4.3.3.4 Effect of wounding on F4B.

The identification of a putative homologue of F4B expressed in *Arabidopsis* leaves is in contrast to the mode of expression seen in *Brassica* where no expression in seedlings, most of which is leaf tissue, is observed. Therefore if F4B and Atts4051 do represent the same class of gene then the action of wounding is probably responsible for inducing expression, so the expression of F4B in response to leaf wounding (as described in 3.2.2, a method similar to that reported by Phillips and Gigot (1994) as it involved cutting the leaf tissue into strips) was investigated. However, no increase in the level of transcripts similar to F4B could be detected at 78% stringency, one or two days after wounding (fig.56). This implies that F4B and Atts4051 do not represent the same class of transcript although they may still represent closely related genes as transcripts less than ~80% similar to F4B would not be detected by this analysis. However, different patterns of equivalent gene expression in different species, as has been suggested for 13G6(i) cannot be ruled out.

(a)



(b)

```
  1.CCAGACCTTCACCGGCGCACAGATACGCTAGGAAGGGGTGATTTGGTAAAAATTCTCATATCATTTGACCAAGTGAAAAC.80
 81.TAGTGGTTCGTGTGGTGTTGTATGTAGTAGAAAGATTGGAGAGGGGGCATATTAGTAATTTGGTCCGTTCTTACCTAAAA.160
161.AGAAGTGGTGCTTCGAGTGCTTGATTTTCCACAAAGCCTTGATGACGTTTCCTTTTCGGACCTTTCTATTATTGGTAAGG.240
241.TTTTTTTTTAATACTAAGTTTTATTGATGCAGAGGTGGAATTAAGGACCACCTGCTAGTAGTACATTTATCGTGTTTGTA.320
321.TCCCCTCTTGAAGCGAGAGGGAGAGTTTCACCTTTCAAGAATTGGTAGATTGTATTAACGTCGTTAAGTCATCATCATCT.400
401.ATGTTGTAAATTTTCAATTTTTAAAAGTACGGAATATTTGAGATTAAAAAAAAAAAAAAAAAAA          .465
```

Fig.54. Sequencing on the cDNA clone F4B. (a) ORF map of the full nucleotide sequence. Full length bars indicate stop codons, half size bars indicate potential initiator ATG's. (b) Nucleotide sequence of F4B. The two 4/6 poly(A) signal consensus sequences are singly underlined. The two matches to the partial FUE are doubly underlined while the full FUE consensus is indicated by bold text.

4.3.3.5 Effect of stress treatment on the expression of K8A2.

The effect of stress treatments on K8A2 in vegetative organs was also assayed to determine if it could belong to a floral/stress gene family, such as described for many of the Solanaceae stigma/style sequences. The same blot as was used to assay stress effects on 15H11 was used (section 4.1.2.3). No increase in expression under salicylic acid treatment or light, heat or water deprivation was observed (fig.57). Similarly wounding (as described in 3.2.2) failed to increase expression in leaves.

4.3.3.6 Genomic organisation of K8A2 and F4B.

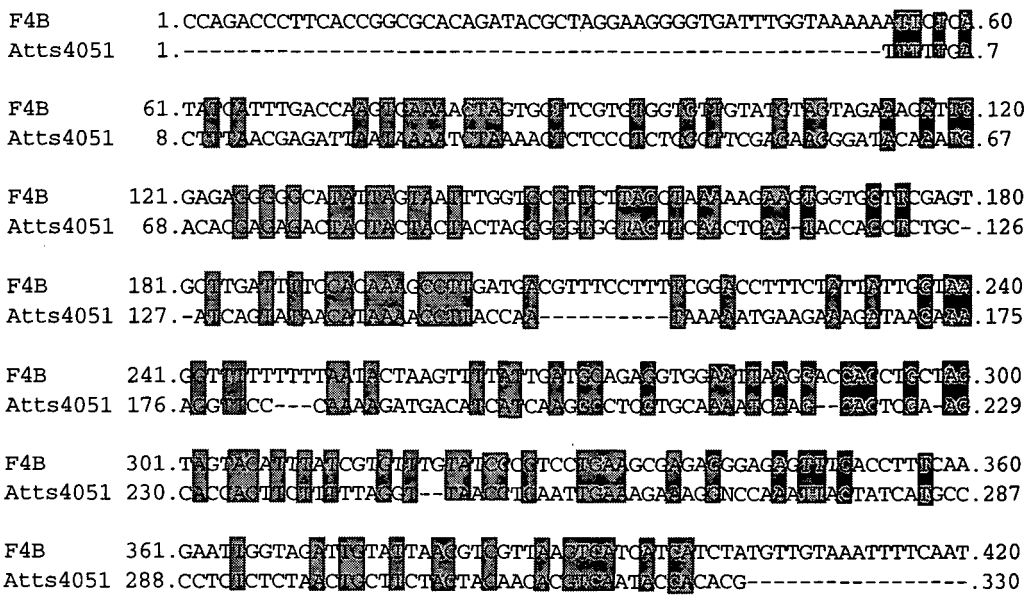
Southern blots were probed with both the K8A2 and F4B inserts in order to assess the complexity of their gene families. Both clones produced only a small number of bands with each restriction enzyme digestion tested suggesting only a small number of copies of each gene (fig.58). For the F4B sequence the genomic fragments all appear to be high molecular weight (>8000bp) except for one 4300bp fragment in the XhoI digestion). These are sufficiently large that they could contain tandem arrays of the gene.

The K8A2 probe hybridised to between 2-5 bands per restriction enzyme digestion. Two of the enzymes used do not cut in the K8A2 cDNA and give 2 (EcoRI) and 5 (BamHI) bands while the third (HindIII) cuts 108bp from the 3' end of the cDNA clone and gives 4 bands on the blot. The two HindIII

	RF1 coding region	RF1 whole clone	RF2 coding region	RF2 whole clone	RF3 coding region	RF3 whole clone
%NNG/C	0.0	5.4	8.9	10.1	10.0	7.7
%NTA	7.1	4.1	4.4	4.1	0.0	0.6
%NCG	50.0	35.8	33.3	37.8	80.0	37.8

Table 9 Codon composition of the three reading frames in the clone F4B. The coding region is the region from the 5' end of the cDNA to the first in frame stop codon in that reading frame. The codon bias for the whole reading frame, not including the poly(A) tail is also included as a reference.

(a)



(b)

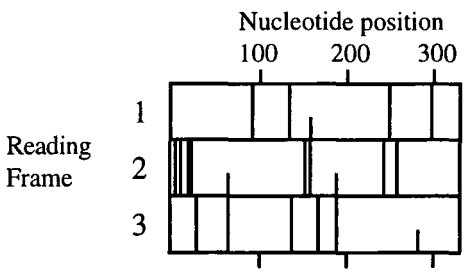


Fig.55. Relationship between F4B and Atts4051. (a)Alignment of the nucleotide sequences of F4B and Atts4051 (Philips and Gigot, 1994) by Clustal W 1.5. Shading indicates exact matches. (b) ORF screen of the Atts4051 sequence. Half height bars indicate potential start codons while full height bars indicate potential stop codons.

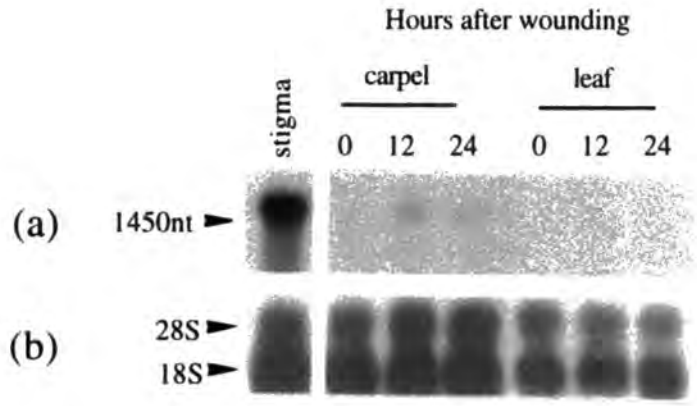


Fig.56. Effect of wounding on the expression of F4B-like transcripts in the carpel and leaf of *B.oleracea*. (a) blot probed with the F4B cDNA. The blot was washed in 1xSSC, 0.1%(w/v)SDS at 60°C which gave a stringency of 76%. (b) the blot in (a) reprobed with pHA1 to confirm equal loading of RNA (10μg per lane).

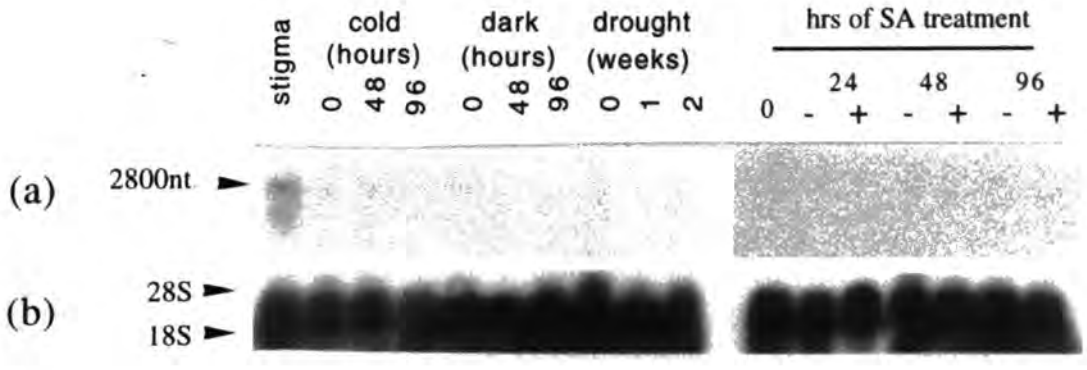


Fig.57. Effect of various stresses on the expression of K8A2-like transcripts in seedlings and leaf of *B.oleracea*. (a) blot probed with the K8A2 cDNA. The blot was washed in 1xSSC, 0.1%(w/v)SDS at 60°C which gave a stringency of 78%. (b) the blot in (a) reprobed with pHA1 to confirm equal loading of RNA (10μg per lane).

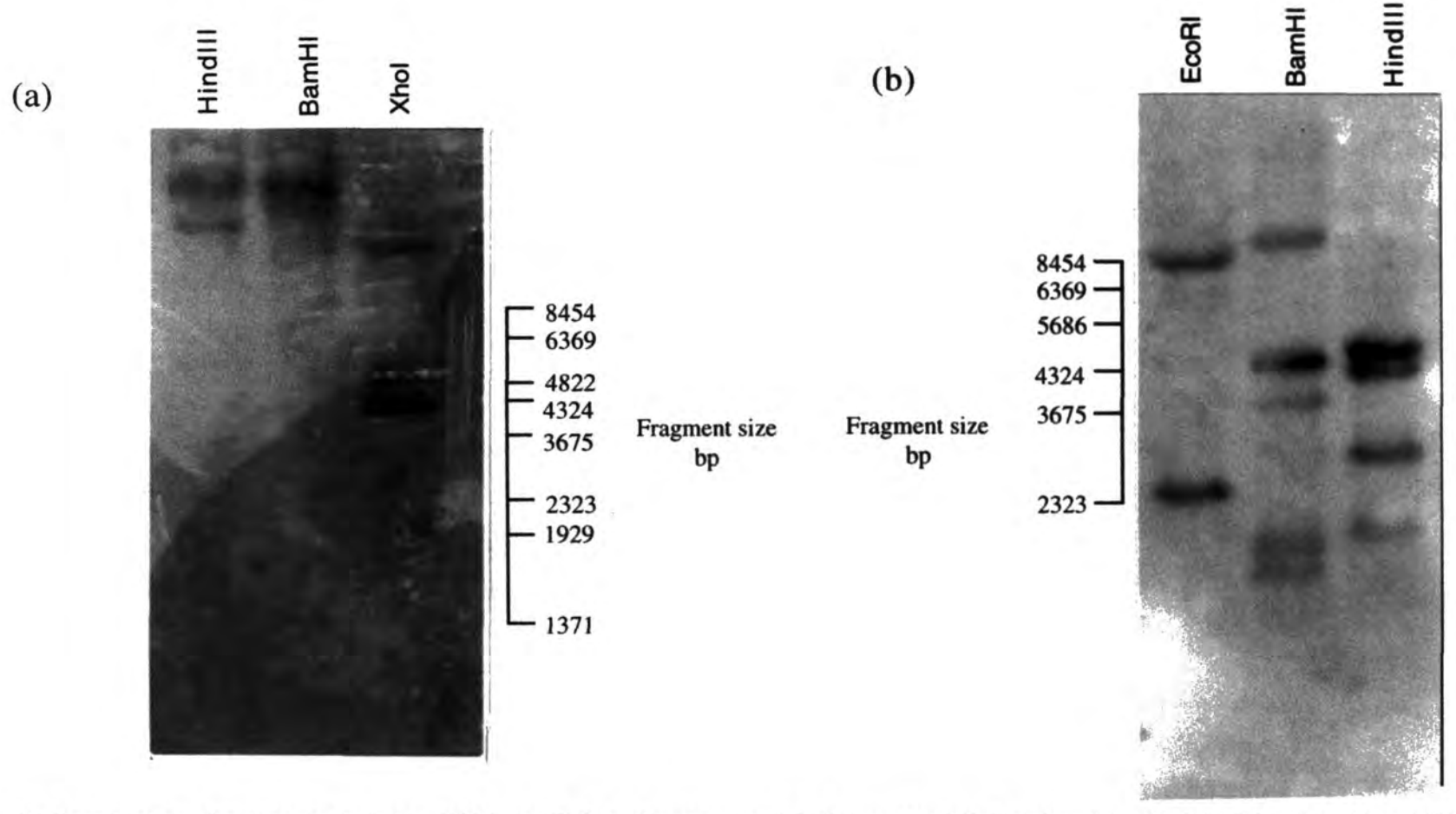


Fig.58.Organisation of sequences similar to K8E2 and F4B in the genome of *B.oleracea*.(a) F4B Southern blot.(b) K8A2 Southern blot. All lanes contain 10µg of DNA. Both blots were washed in 1xSSC, 0.1%(w/v)SDS at 60°C which gave stringencies of 84% and 85% respectivly. Sizes are derived from lambda BstEII markers co-run on the same gels as the genomic DNA.

fragments produced from the cDNA have sizes of 690bp and 108bp, both of which would hybridise to the cDNA under the conditions used (T_M=84°C and 89°C in the wash buffer, respectively).

The simplest interpretation from the EcoRI and HindIII digestions is that only two copies of K8A2 occur in the haploid genome of *B.oleracea*. This, however is unlikely due to the result with BamHI. The greater than expected number of bands after digestion with BamHI suggests that the bands seen after EcoRI and HindIII digestion represent more than one copy, either by different but identically sized restriction fragments or by two gene copies on the same restriction fragment. Some of the bands in the BamHI digestion could have been produced by relevant restriction sites lying within the genomic region equivalent to the cDNA clone: such sites being produced by mutation or being present in introns. However because bands produced by BamHI digestion are all >1500bp, only two of the extra bands could come from point mutations or introns as neither of these could generate more than two bands from a region covered by a clone the size of K8A2. Therefore, K8A2-like sequences are predicted to occur at a low copy number greater than two in the genome of *B.oleracea*.

4.3.4 Other clones.

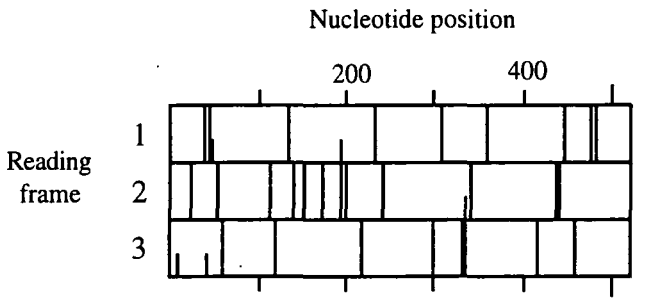
The clones C10A and D5A2 displayed different patterns during differential screening. D5A2 displayed hybridisation to the stigma probe but not to the leaf probe, while C10A hybridised to both probes. The extent of hybridisation of C10A to the stigma probe appeared greater than to the leaf probe but as differential screening is not entirely quantitative, the relationship of this result to the expression pattern of transcripts homologous to the clone could not be directly inferred. Both clones were purified and shown to have estimated insert sizes of ~500bp (C10A) and ~1400bp (D5A2).

4.3.4.1 Sequence of C10A

As C10A was only ~500bp long, the entire sequence was obtained on both strands by sequencing from either end of the clone using pKS+Bluescript Reverse primer and oligo-dT primer. This gave the sequence described in fig.59b which is 519bp long excluding the poly(A) tail whose size could not be accurately determined from the 5'→3' sequence and which was not fully represented on the 3'→5' sequencing run. From the 5'→3' sequence the poly(A) tail was estimated to be 15-25bp long which does not significantly affect the clone size so that the sequence length corresponds with the insert size estimated by gel electrophoresis. It was not deemed necessary to accurately obtain the full length of the poly(A) tail by further sequencing.

As a poly(A) tail is present this indicates that C10A represents the 3' end of its relevant transcript. One AATAAA motif is also present in the clone (fig.59b) although as this is 165bp from the poly(A) site it is unlikely to act as a poly(A) signal in the transcript from which C10A was derived. One 4/6 match, within 40bp of the poly(A) tail, is also present (fig.59b), although due to the current lack of knowledge about what constitutes a plant poly(A) signal (Hunt, 1994) what, if any, role this motif played in 3' end formation is not known.

(a)



(b)

```
  1.CAATTCCAATGGAATTTATCTTTGAACGGGAATCTATAGATATGAATGGAACTGAAAAGTAACAACACTATCAGCATCAC.80
 81.TCGGGTTTGTCTGGACTTTTGGTCCATCGTCGTAAGTCGTAGTATTGCGTCGTAAGTCGTAGTATTGCACATAAAAGAGA.160
161.CACTTCTCAAAGTAGGAATCATCTCAGGGTTTATGAGTTTGAGACAAGAAAGACTCCGTAAGTACTGCATTTGAAGGCTT.240
241.CTAGCGGTTCACCGTACCTACACCCCCTTTTGGCTTGTGCTGTGTTACAGACAAGTGGGTGAGAAGTGAGAACTTACTCC.320
321.CGAGTCCAAGTATAATGAAATAAAGTCGAGAAGTATATAGGGAGCAGGAAACACTGTATCGGGTTTTCTTTGGGTTTCTT.400
401.TGGAGGTCTCGGAACGTGATTTTCAGGATTTCACCATAATAGTTTGATTACTGCGAAGTAATTTTTTTTACTGCTTAATAT.480
481.TAATCTCGTTGGCAAGTTTGGTTATTTC(A)15-25                                          .519
```

Fig.59. Sequence analysis of C10A. (a) ORF map of the C10A insert. (b) full nucleotide sequence of the C10A insert. An accurate assessment of the size of the poly(A) tail was not possible as the forward sequence was of too poor quality, while backward sequencing was performed using a poly(dT)G primer so that only the sequence 5' to the poly(A) tail was obtained. The putative poly(A) consensus signal sequence is underlined.

Searching for ORF's in the clone revealed that no reading frame greater than 58 codons could be identified, while the largest ORF was only 32 codons in length (fig.59b). As the clone represents the 3' end of the transcript this suggests that the majority of the clone represents the 3' untranslated region of the relevant transcript. The 5' region of the clone could contain coding information but as only 7-20 codons could be present no further analysis was carried out as such a small dataset is insufficient for codon bias analysis to determine which, if any of these sequence fragments code for the C10A protein.

4.3.4.2 Sequence of D5A2.

The size of the D5A2 insert was such that subcloning was necessary to obtain the full sequence. Initially the sequencing was carried out from either end of the clone which revealed the presence of an SpnI site ~756bp from the 5' end of the clone from which further sequence was derived as indicated in fig.60. This gave the full sequence of the clone but not on both strands as complete sequencing required subcloning from a site in the 5' half of the clone. Restriction mapping of the sequence indicated that the only suitable site was NsiI (fig.61). The strategy chosen to use this site was digestion with MpH1103/EcoRI and MpH1103/NotI (MpH1103 is an isoschizamer of NsiI), polishing with T_4 DNA polymerase. Recircularisation would then give two clones to allow sequencing forward and back from the NsiI/Mph1103 site (fig.62). This approach, however failed to generate any clones so that such sequencing

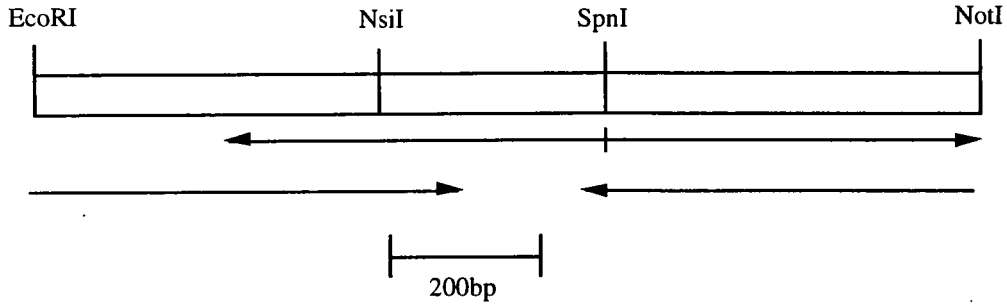


Fig.60. Subcloning strategy to obtain the full sequence of D5A2 on both strands of the cDNA. The sequence obtained is indicated by bars with the arrows indicating the direction of sequencing and the bar length indicating the amount of sequence obtained.

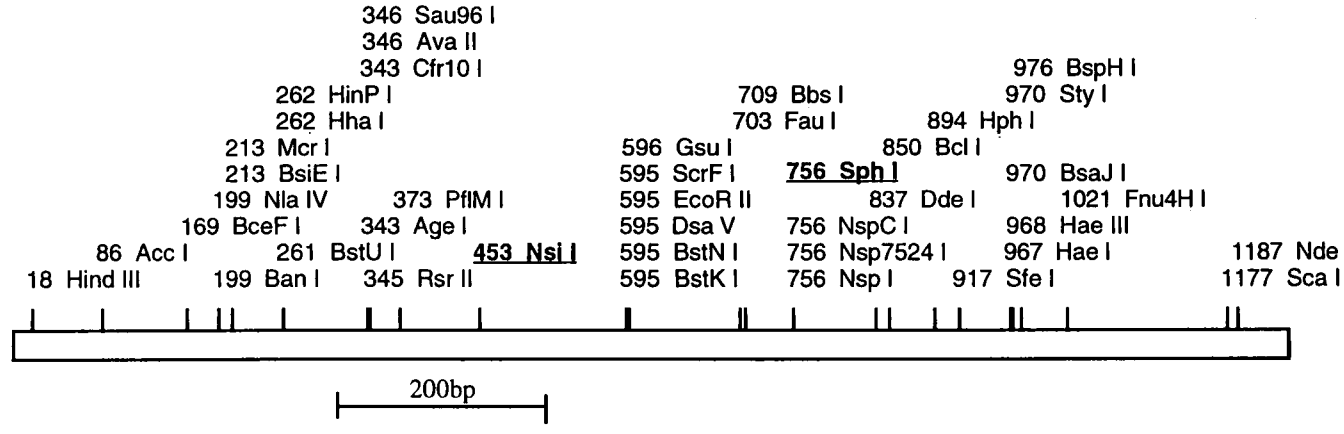


Fig.61. Restriction map of the full sequence of the cDNA clone D5A2. Only unique sites are shown. The two sites chosen for subcloning are indicated by bold, underlined type.

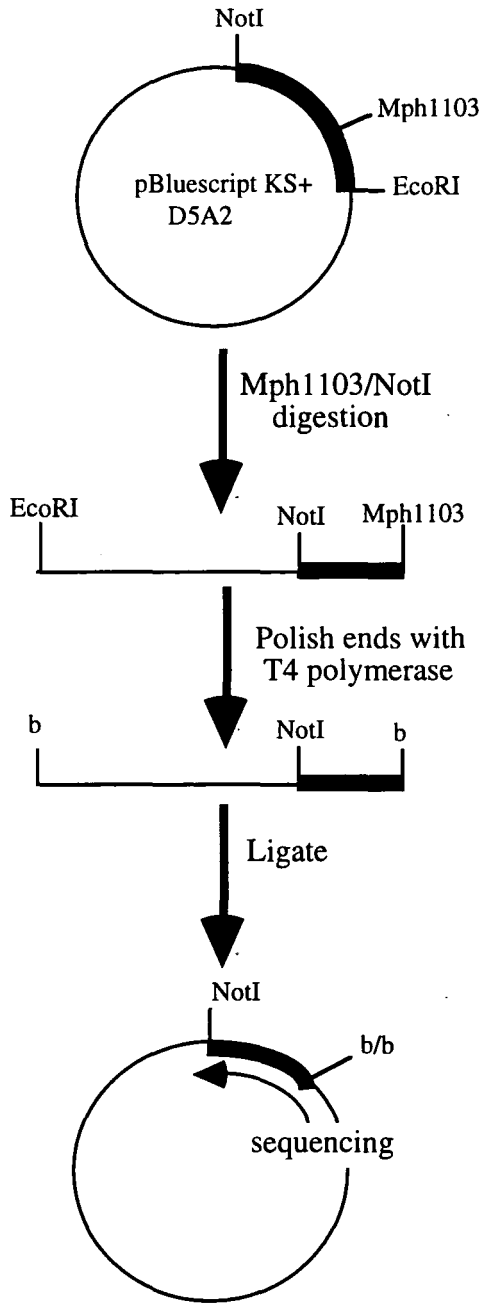
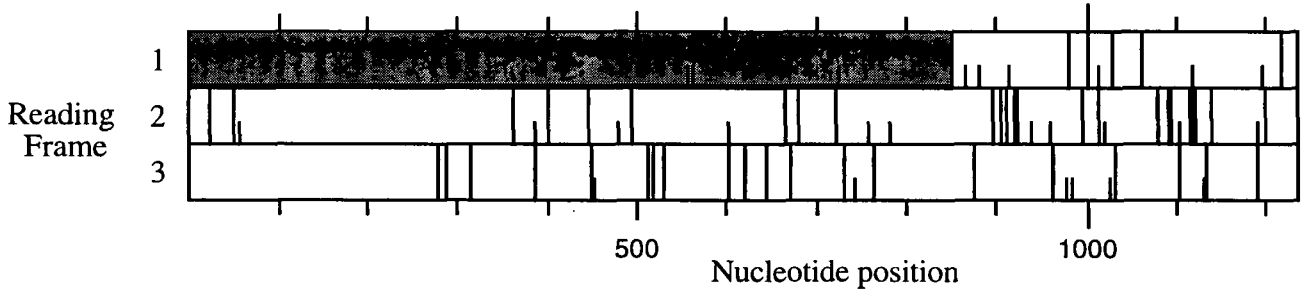


Fig.62. Sequencing strategy (using Mph1103) to obtain the 5' end sequence of the clone D5A2. The pKS+Bluescipt D5A2 construct was digested with Mph1103 and EcoRI, the ends polished to allow religation ('b' indicates a blunt end), then the construct recircularised. Sequencing in the other direction was to be carried out by an identical procedure with the exception that NotI digestion was substituted for EcoRI.

was not possible. Therefore as the sequence on at least one strand was available for the whole clone, no further work to obtain the full sequence on both strands was carried out. If the full length sequence is required in the future it could be obtained by sequencing directly from the full length clone using primers that can bind in the region encompassed by nucleotides 400-500 of the clone

The full length of the D5A2 insert, including poly(A) tail is 1234bp Searching for coding regions in this sequence revealed one long 283 codon reading frame which contains a 98 codon ORF (fig.63a).

(a)



(b)

```
   1.AGAGGGAGAGAGAAGAGAAGCTTGAAGAAGGTTCTCTATCTCTCGTTCATAGATTATGGAACTGCTCTATCTCCTCTGCT.80
     R  G  R  E  K  R  S  L  K  K  V  L  Y  L  S  F  I  D  Y  G  T  A  L  S  P  L  L
  81.CAATCGTCTACACTTCAACCACAACGCTCTTCCTCTCTCTCCTGCTCCCCTTCCGCTCCTTTTCCACCGTCTCCTACCTT.160
       N  R  L  H  F  N  H  N  A  L  P  L  S  P  A  P  L  P  L  L  F  H  R  L  L  P
 161.CACGCTCCGCCGTCGATTCCAACGTCTCTTACTACGAAGGCACCGTCTGGCACGACCGTCTCCGTCCCGTCCGTCACTCG.240
     S  R  S  A  V  D  S  N  V  S  Y  Y  E  G  T  V  W  H  D  R  L  R  P  V  R  H  S
 241.TTCCGTTACACCGTCCGTTACGCGCTCTTCGACCTCGATAAATCCCTTGAAACTCCGCCGGATCATCTCTCCGCTGACGA.320
     F  R  Y  T  V  R  Y  A  L  F  D  L  D  K  S  L  E  T  P  P  D  H  L  S  A  D  E
 321.AGCTCGCCTTCTCGCTCGTACCACCGGTCCGATTTTCTTATTGACAATACCTCCAAGCGTTGGATATGAGCAGAACCCGT.400
       A  R  L  L  A  R  T  T  G  P  I  F  L  L  T  I  P  P  S  V  G  Y  E  Q  N  P
 401.TGAGTTTGTATTATTGCTACAACTTGGAAGGATCAAGCAAACGCTTAAGTAAATGCATTGCACAGGTTACAAATACGCAT.480
     L  S  L  Y  Y  C  Y  N  L  E  G  S  S  K  R  L  S  K  C  I  A  Q  V  T  N  T  H
 481.GGGGGGGAACGAGTGACATTTGTTTTCGACCCTGAATCTGACTTGGTTGCTAAATCATTACAAGTCAGTCCTTTCATGGA.560
       G  G  E  R  V  T  F  V  F  D  P  E  S  D  L  V  A  K  S  L  Q  V  S  P  F  M  D
 561.TATGCTTGGGAATTGGAAGATCAGAGCAAACGAACCTGGAGATGAGTTATCTGTATCAATTGAATCACAGCATCCTCATC.640
         M  L  G  N  W  K  I  R  A  N  E  P  G  D  E  L  S  V  S  I  E  S  Q  H  P  H
 641.ACGGTAACTACTTCTCTGCCAAGTTGAAGGCTAAAAGAATAGACCAAACACGGGTTTCAGATCCCGCTGTCTTCTTCTGG.720
     H  G  N  Y  F  S  A  K  L  K  A  K  R  I  D  Q  T  R  V  S  D  P  A  V  F  F  W
 721.TTGATGCCTCATAAGGTTGCTATATGGATCTATTGGCATGCACTTAAGCTCTGGTGGAGGAATGTACCCTTTCATCCAAC.800
       L  M  P  H  K  V  A  I  W  I  Y  W  H  A  L  K  L  W  W  R  N  V  P  F  H  P  T
 801.ACCCCAGATACTACAAACCCATTCATACAGAGAGGACTCAGCGAAAACGTGATCAAAAACTTCGATGTGTTGGTTTAAGA.880
         P  Q  I  L  Q  T  H  S  Y  R  E  D  S  A  K  T
 881.TGGATCAAACTCTGGTGAAACCATTAGCTTTGATGGCTGTAGTAGTGGTTTTGGAGGATGGTCGCTTTGCGTGGCGAGAT.960
 961.GCTAATTGGCCTTGGTCATGAATATGCACACACTTGAGATAGCAGATCGGCATGACTCATGCCGCATGAGTTAAGAATCT.1040
1041.ACAATCATTCATCGAGAGGTAGGAGTGGTTTACTTGTTTAGATTTGGCTTAATAAGCAAAGATGAACGTTGTTTGATGA.1120
1121.TAAGCCGACGATGTAACTTGACTTCTTGCAAGTTGGAGAACAATCCTGCATTTTCAAGTACTATACCATATGAAATGTTA.1200
1201.GCGTACCATTTGCTCTAAAAAAAAAAAAAAAAAA                                              .1234
```

Fig.63. Sequence of the cDNA clone D5A2. (a) ORF map of the clone. Half height bars indicate ATG sequences. Full height bars indicate potential stop codons. The shaded region indicates the predicted reading frame. (b) Nucleotide sequence of D5A2 together with its predicted protein. The T rich regions are indicated by bold text.

The sequence around the ATG in the ORF matches only one nucleotide, except the ATG, of the Lutcke (1987) consensus (TTTCATGGA vs. AACAATGGC) indicating that this is probably not the translation initiation site, which suggests that D5A2 is only a partial sequence. As with the previously described sequences, the presence of a poly(A) tail indicates that the 5' end of the clone is missing. From this analysis the full sequence of the D5A2 cDNA and its predicted protein is as described in fig.63b.

4.3.4.3 Sequence analysis of the C10A and D5A2 clones.

In an attempt to identify any possible function for both clones or their encoded proteins the C10A and D5A2 sequences were used to scan the EMBL sequence database. No significant matches for the C10A sequence were found indicating that it is a completely novel sequence.

The SWISSPROT database was screened with the predicted D5A2 protein sequence (the lack of any identifiable C10A sequence meant that no justifiable data could be gained for similar analysis with that clone). No protein matches to the putative D5A2 protein sequence were found so that no function prediction from this data was possible. To complete sequence analysis of the D5A2, the predicted protein sequence was screened for conserved motifs in the PROSITE database. No positive matches were identified, while searching for putative glycosylation sites similarly also gave a negative result. As a reading frame could not be predicted for C10A no attempt to screen for homologies to the possible proteins encoded by this clone was made

From this analysis no prediction on the function of C10A or D5A2 can be proposed. Neither full nucleic acid sequence has been previously reported although the three D5A2-like EST's probably represent the equivalent *Arabidopsis* genes. The amount of available protein sequence for D5A2 means that any similar proteins would be picked up by database screening so its predicted protein is novel.

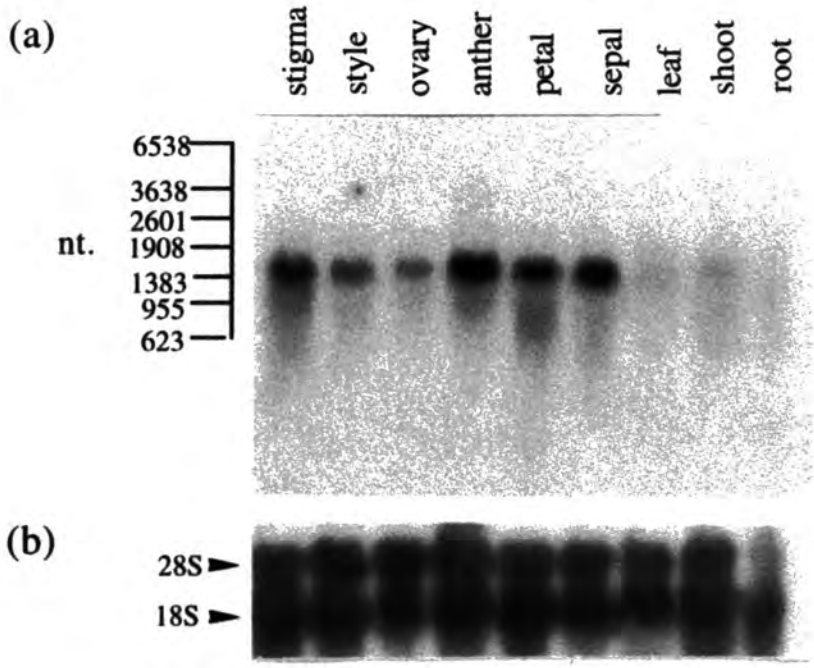
4.3.4.4 C10A and D5A2 are not stigma specific.

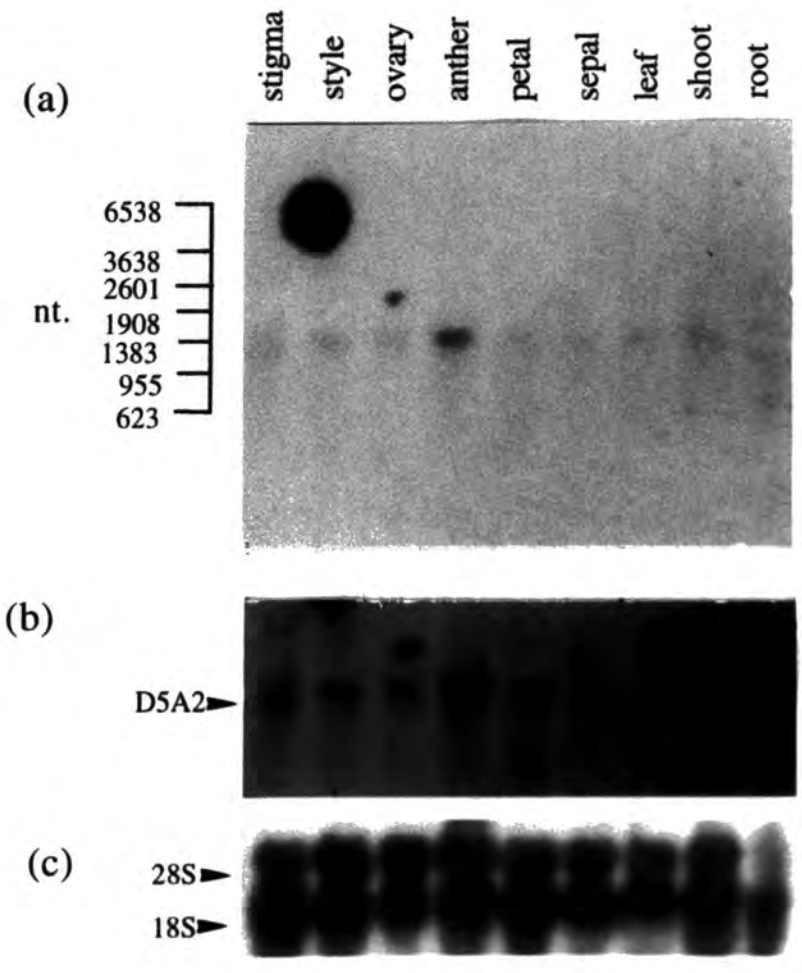
To determine the pattern of expression of transcripts, that cross-hybridise to the C10A and D5A2 cDNA's, Northern blots were probed with the inserts. C10A-like transcripts were detected in all the tissues of the mature plant, as expected from the clone's behaviour in differential screening (fig.64). Expression was weaker in the vegetative tissues compared to the flower indicating that the pattern of expression predicted by the differential screen was correct. Some variation in the flower was also evident, with a lower level of expression in the style and ovary compared to the other organs. The size of the transcripts in all the tissues was ~1400nt. indicating that C10A represents approximately one third of the transcript from which it was derived.

Probing Northern blots with the D5A2 cDNA gave a very weak signal, even after 2 weeks exposure, which allowed only faint bands to be determined from which a description of the expression of the clone could be not be clearly deduced (fig.65a): positive expression was too weak to determine whether RNA samples showing no hybridisation were true negatives or were due to a failure to detect D5A2-like transcripts.

Fig.64. Expression of the clone C10A in the mature organs of *B.oleracea*. (a) Autoradiograph of a Northern blot probed with the C10A cDNA. (b) the same blot probed with pHA1 to confirm equivalent loading of RNA (10μg per lane). The blot in (a) was washed to 1xSSC, 0.1%(w/v)SDS at 60°C giving a stingency of 78%.

Fig.65. Expression of D5A2. (a) Autoradiograph of a Northern blot probed with the D5A2 cDNA. (b)The autoradiograph in (a) after enhancement by scanning and contrast adjustment. The blot was washed in 1xSSC, 0.1%(w/v)SDS at 60°C giving a stringency of 73%. (c) the same blot probed with pHA1 as described in 3.6.4.3 to confirm equivalent loading (10μg per lane).


(a)
stigma
style
ovary
anther
petal
sepal
leaf
shoot
root
6538
3638
2601
1908
1383
955
623
nt.
(b)
28S
18S


(a)
stigma
style
ovary
anther
petal
sepal
leaf
shoot
root
6538
3638
2601
1908
1383
955
623
nt.
(b)
D5A2
(c)
28S
18S

In an attempt to overcome this problem the D5A2 autoradiograph was artificially enhanced by input, via a flatbed scanner, into the Hewlett-Packard DeskScan utility (PC version) which was used to increase the level of contrast to increase the relative intensity of the bands. This revealed that D5A2-like transcripts, of length 1500nt. were present in all the organs of the flower (fig.65b). Expression levels appear roughly equivalent in all tissues except the anther where there is a stronger signal than in the rest of the flower. Contrary to the differential screening results expression was also detected in RNA from leaf and stem tissue. However, as this image was artificially obtained these results may be somewhat artifactual

The contrasting results seen between differential screening and Northern probing suggests that the enhanced blot is not giving an accurate representation of the expression of transcripts homologous to this clone. Similarly, the low level of detection of D5A2 on Northern blots is in contrast to its detection by differential screening. Northern analysis of total RNA should be able to detect transcripts that comprise 0.001% or more of the total mRNA population while differential screening should only detect mRNA species that constitute 0.01% or more of the RNA (Sambrook et al., 1989). Therefore all of the isolated clones should be easily detectable on Northern blots. As the Northerns used for D5A2 probing had previously been through two rounds of probing, this may have affected their sensitivity although such a decrease would have to be severe to prevent effective detection. A fall in sensitivity would also not explain the differences in leaf expression between the differential screening and the blot probing.

One explanation for this apparent pattern of expression could be that the observed bands are due to weak binding of the D5A2 probe to a transcript with which it shares only weak identity. If this is so then an explanation for the failure of the probe to bind to D5A2-identical transcripts, which from the differential screen must be present in the stigma, is required. One possible explanation is that the 'real' D5A2 transcript may be similar in size to either the 18S or 28S rRNA's. These nucleic acids have been reported to interfere with the transfer of co-migrating transcripts to the membrane during blotting, possibly by blocking binding sites on the relevant region of membrane as, in total RNA preparations, they occur at levels that can saturate the binding capacity of the membrane (Sambrook et al., 1989).

Resolution of the problems associated with the expression of D5A2-like transcripts could be achieved with the use of poly(A) rather than total RNA. As well as eliminating any interference from ribosomal species, this would enable a larger amount of target RNA to be loaded into each lane so enhancing the signal achieved by binding. However, the quantity of tissue required to obtain suitable yields of poly(A) RNA coupled with the small size of the *Brassica* floral organs meant that harvesting of tissue was not possible in the time available. Therefore confirmation of the pattern of expression of D5A2-like transcripts was not completed during this project.

From the available data, the pattern of expression of these two clones suggests, as for 13G6(i), that they are unlikely to encode proteins which play a defining role in the stigma or carpel. C10A is not floral-specific and is unlikely to play a determining role in the flower. Expression in the flower is higher than in the vegetative organs but the difference in signal is such that it is unlikely to have a major role physiological impact.

Due to the problems in assessing its expression firm conclusions cannot be made about D5A2 but if the observed pattern of expression is correct it is also unlikely to play a defining role in the flower.

4.3.4.5 Genomic organisation of C10A and D5A2.

To complete the analysis of these clones their genomic organisation in *B.oleracea* was studied. After digestion with HindIII, BamHI and XhoI genomic DNA gives only simple patterns of hybridisation (fig.66). D5A2 hybridises with two bands in BamHI and XhoI digestions and one band with HindIII. Only the latter enzyme cuts within the cDNA clone and that is only 18bp from the 5' end so that equivalent digestion of genomic DNA would give a fragment too small to hybridise to the probe under the conditions used (calculated TM=70°C). This suggests that D5A2-like sequences occur as one copy with introns containing XhoI and BamHI sites or as two copies which exist on identically sized HindIII fragments.

Only interpretable data on the hybridisation pattern with HindIII and XhoI digested genomic DNA is available for C10A. In both digestions one high molecular weight band cross-hybridises with the cDNA. This could be a single copy gene, or a tandem array of genes on one fragment. Due to the size of the smallest fragment (~7500bp) in the BamHI digestion any tandem array would be restricted to a maximum of <4 copies given the size of the C10A transcript.

Therefore both C10A and D5A2 are probably representatives of small gene families, although confirmation of this would require gene reconstruction. However as the Northern analysis suggested that both these clones are of little interest, this work was not carried out.

4.3.5 Comparison of *Brassica* and *Nicotiana*.

4.3.5.1 Clone detection in *Brassica* and *Nicotiana*.

The sequence of eleven cDNA's have been reported in this thesis, none of which resemble cDNA and protein sequences that have been reported in the Solanaceae. Therefore Northern blots containing RNA from the reproductive organs of *N.tabacum* (as an example of a plant with a wet stigma) together with *B.oleracea* were prepared. The use of the former as a representative of the Solanaceae was determined by availability. Due to the differences in biology between the Cruciferae and Solanaceae, a combined stigma/style RNA extract was used for comparison.

Expression in *Nicotiana* was assayed for 13G6(i), 15H11, C10C and K8A2. No signal of any kind was obtained for the latter two sequences, even at low (1xSSC, 0.1%(w/v)SDS, 60°C) stringency (data not shown). This is probably due to the restriction of expression to the stigma for K8A2, which is diluted, so becoming undetectable when using stigma/style RNA.

Clear positive signals were obtained in *Brassica* tissue for 13G6(i), 15H11 and C10C at the stringency used, but no homologous transcripts could be detected in the *Nicotiana* samples (fig.67). This result suggests that no transcripts which share of 76%, 73% and 78% identity to these clones are expressed in *Nicotiana* stigma and styles.

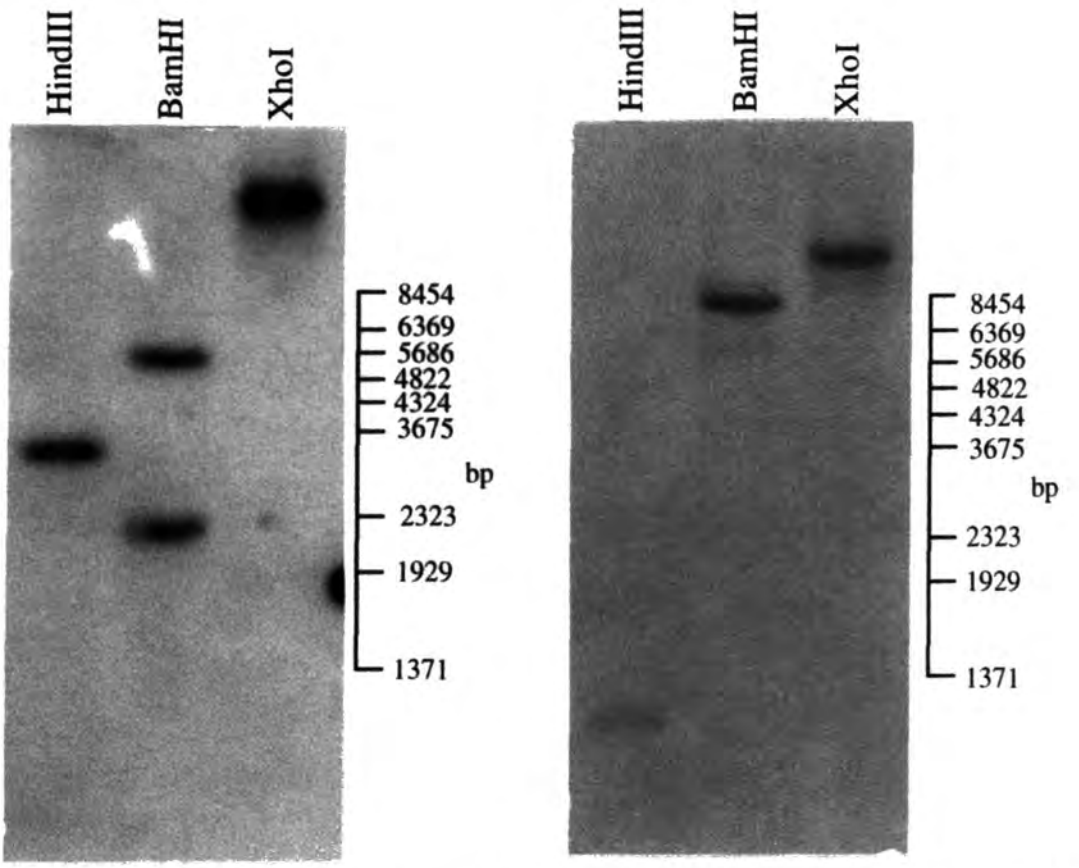


Fig.66. Southern analysis of *Brassica oleracea* genomic DNA using the clones C10A and D5A2 as probes. (a) hybridisation of a D5A2 probe with *B.oleracea* genomic DNA (b) hybridisation of a C10A probe with *B.oleracea* genomic DNA. Both blots were washed in 1xSSC, 0.1%(w/v)SDS at 60°C giving stringencies of 80% and 82% respectivly. Indicated sizes are derived from lambda BstEII size markers co-run on the gels from which the blots were taken.

This result has different implications for the different clones. For C10C it does not necessarily rule out expression as a large part of this clone is 3'UTR which can show high variability between transcripts encoding homologous proteins.

For 13G6(i) the result implies that no 13G6(i)-like transcripts are expressed in tobacco reproductive organs. Although no tobacco homologues of this clone are available their existence can be predicted, given the apparently ubiquitous presence of genomic copies of 13G6(i) in other plants (section 4.1.1.4). From the percent similarities between 13G6(i) and its EST homologues and the hybridisation pattern on the multi-plant genomic blot, the predicted *Nicotiana* 13G6(i)-like gene should be 70-75% identical to the *Brassica* genes. Therefore, if similar transcripts are present in *Nicotiana*, they should have been detected under the hybridisation and washing conditions used in this experiment.

A similar conclusion can be drawn for 15H11. The high GC content of this clone means that under the conditions used it should hybridise to clones with only 73% identity. At this level of identity it is unlikely that a clone would encode a GRP of similar structure to that predicted for 15H11. Expression of other GRP's with different primary structures cannot be ruled out, however.

4.3.5.2 Polypeptide patterns in the reproductive tissue of Brassica and Nicotiana.

Differences between the floral organs of *B.oleracea*. The hybridisation results suggest that different genes are expressed in the female reproductive organs of *Brassica* and *Nicotiana*. In order to investigate this further the pattern of proteins from these organs were compared by SDS-PAGE to determine if there are any significant qualitative differences. If stigma/style proteins could be identified it was hoped that they could be partially sequenced

Initially the polypeptide patterns in different *Brassica* organs were compared to each other (fig.68a). A number of bands are common to all organs tested, the most prominent of these have Mrs of 70,66,63,51,46,41,31,16 and 14kDa. Two of these can be identified as the large (51kDa) and small (16kDa) subunits of Rubisco. .

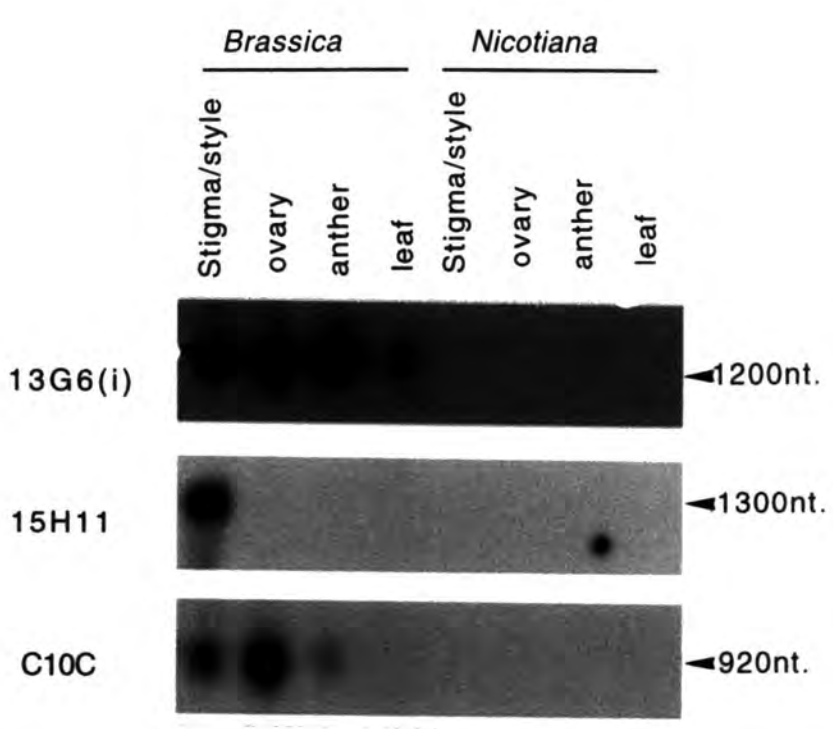
Three clearly defined floral-specific band at 40, 36 and 19kDa can also be identified. Unfortunately no stigma/style specific bands could be determined from this gel although two ovary specific bands could be identified with Mr's of 56 and 49kDa.

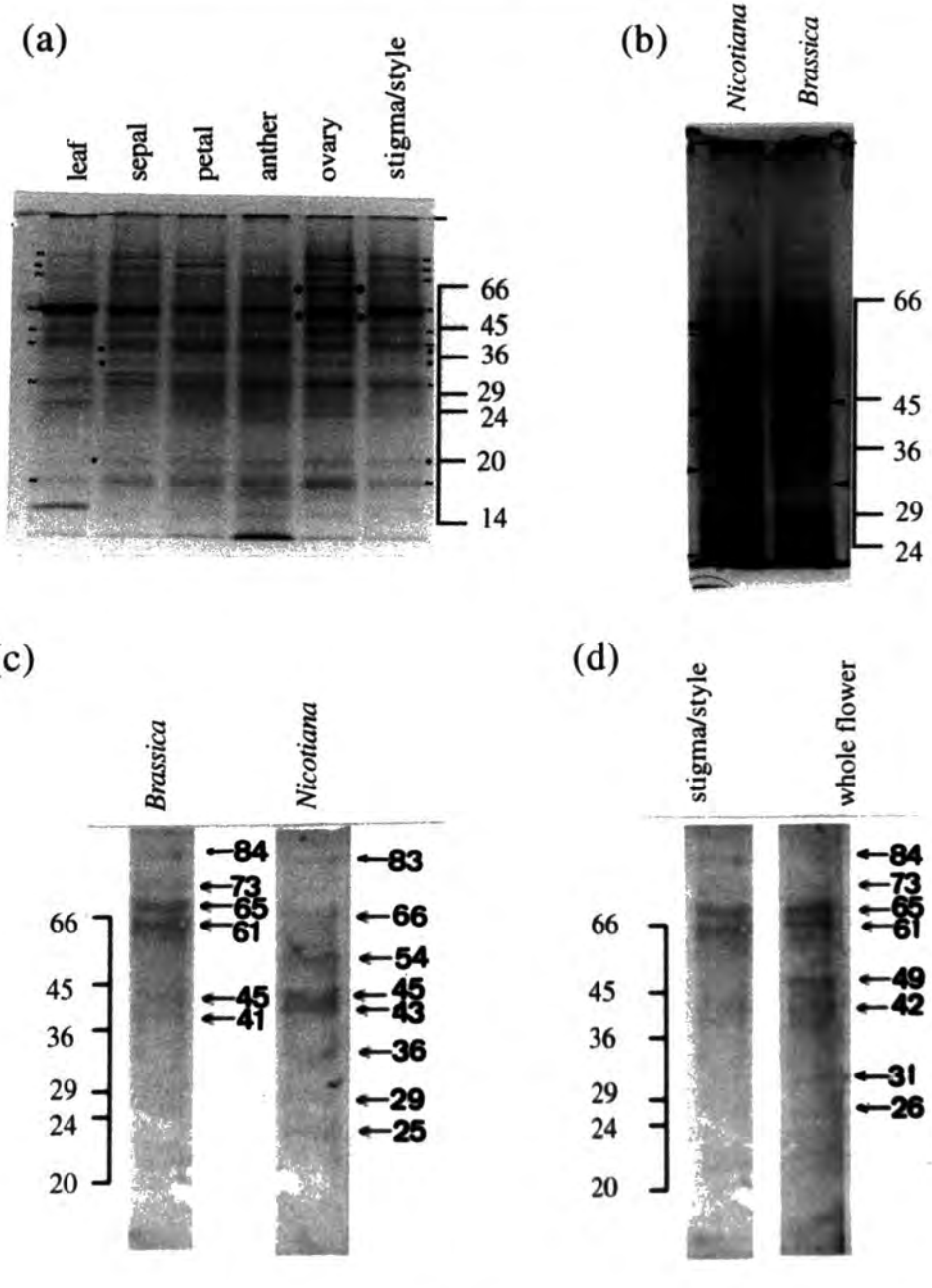
These results suggest that the molecular differences between the floral organs of *B.oleracea* are not dictated by the major proteins which appear to be consistent in their expression across the flower. Likewise the difference between the floral organs and leaves must be primarily determined by proteins not visible in total extracts.

Differences between *Brassica* and *Nicotiana*. The electrophoretic patterns of *Brassica* stigma/style extracts were then compared to those of *Nicotiana*. Few differences could be detected (fig.68b): *Nicotiana*-specific bands are visible at 60, 42 and 32kDa while *Brassica*-specific bands occur at 41 and 31kDa. As neither of the *Brassica* bands are stigma, or even floral, specific, it is unlikely that either is of any interest in the study of stigma/style biology.

Fig.67. Expression of transcripts that cross hybridise with 13G6(i), 15H11 and C10C in *B.oleracea* and *N.tabacum*. 10μg of total RNA were loaded into each lane.

Fig.68. Polypeptide pattterns of *B.oleracea* and *N.tabacum* tissue. (a) Coomassie blue stained gel showing the protein patterns for different tissues from *Brassica.* ~ indicates major bands present in all tissue, ° indicates bands present in all floral but not leaf tissues, * indicates ovary specific bands. (b) Coomassie blue stained gel showing the protein patterns for *Brassica* and *Nicotiana* stigma/style tissue. Arrows indicate bands present in one tissue but not the other. (c) ConA glycoprotein stained blot showing the different patterns of glycoproteins between*Brassica* and *Nicotiana* stigma style extracts. All bands are labelled with their size in kDa. (d) ConA binding to *Brassica* stigma/style and whole flower extracts. The floral bands are labelled with their estimated size.


Brassica
Nicotiana
Stigma/style
ovary
anther
leaf
Stigma/style
ovary
anther
leaf
13G6(i)
1200nt.
15H11
1300nt.
C10C
920nt.


(a)
leaf
sepal
petal
anther
ovary
stigma/style
66
45
36
29
24
20
14
(b)
Nicotiana
Brassica
66
45
36
29
24
(c)
Brassica
Nicotiana
84
73
65
61
45
41
83
66
54
45
43
36
29
25
66
45
36
29
24
20
(d)
stigma/style
whole flower
84
73
65
61
49
42
31
26
66
45
36
29
24
20

Both these results suggest that despite the large number of floral organ-specific genes predicted in the carpel (Kamaly and Goldberg, 1980), housekeeping genes still predominate when comparing different organs by SDS-PAGE. However it is unclear whether this is due to the housekeeping genes predominating throughout the organs or whether it is a consequence of expression in the non-specialised tissues of the stigma and style masking that in the specialised tissue.

Concanavalin A staining of *Brassica* and *Nicotiana* stigma/style extracts. In a final experiment proteins from the stigma and style of the two species were separated by SDS-PAGE, electroblotted and probed for Concanavalin A (ConA) binding. This lectin, which binds to terminal glucose/mannose residues in glycoproteins, has previously been shown to bind strongly to the stigmas of *Gladiolus* and inhibit the interaction with pollen suggesting that it interacts with key components of the stigma (Clarke et al., 1979).

The binding pattern of this lectin does indicate differences between *Brassica* and *Nicotiana* stigma/style extracts (fig.68c). The major bands in the *Brassica* extract occur at 83, 73, 65, 61, 45 and 40kDa. Those of *Nicotiana* are at 84, 72, 66, 54, 45, 43, 36, 29 and 25kD. Therefore only one band (45kDa) occurs in both the stigma/style preparations from the two species. The differing bands could be due to either the presence of different proteins or differential glycosylation. Given the larger number of bands in the *Nicotiana* sample the former is more likely, although the multiple bands could be due to multiple glycoforms of the same protein or proteins.

To determine if the *Brassica* ConA-bonding proteins are unique to the stigma and style the binding pattern was compared to that of whole flowers (fig.68d). Many of the bands seen in the stigma/style extract are also present in the whole flower extract suggesting that these are not stigma/style specific. Only the 45 and 40kDa stigma/style bands are unique to the stigma and style. As the 45kDa is also seen in the *Nicotiana* stigma/style extract this may represent a cross-species stigma/style specific protein.

This leaves the 40kDa protein as the only *Brassica* stigma/style specific protein identified in this study. The failure to detect this protein using coomassie stain may have been due to lack of sensitivity or masking by housekeeping proteins.

5
CONCLUSIONS

5.1 Assessment of results.

The principal aim of this project was to isolate and identify sequences expressed in the stigma of *B.oleracea* in order to partially redress the strong bias in this area towards work with the Solanaceae. *Brassica* species are the models for studying sporophytic SI and a complete description of the molecular basis of this phenomenon will require a description of the action of recognition of compatible pollination. The morphological and physiological differences between the Solanaceae and Cruciferae precludes direct extrapolation of results from the former to the latter. The success of this project is now considered:

5.1.1 Differential Screening as an approach to study.

The bulk of this project has been involved with the isolation and characterisation of stigma cDNA clones by a differential screening approach. Such an approach, while having been partially superseded by other methods such as subtraction libraries (Zimmerman at al., 1980) and differential display (Liang et al., 1993), is a well established method for identifying genes with tissue-specific expression. The principal limitations of this method is that it is restricted to identifying highly expressed sequences. This has not, however, proved to be a handicap when applying this method to the stigma and style of members of the Solanaceae from which a large number of clones have been isolated. The majority of these have had their products identified, although the relationship between the activity of the encoded proteins and carpel biology has, for the most part, been unclear.

Due to its success in the Solanaceae, differential screening has been adopted as the basis for identifying stigma-specific clones from *B.oleracea* in both this and previous work (Scutt, 1990). Identification of differentially expressed clones was technically simple as almost a hundred clones could be easily identified from only a 15000pfu primary screen. The results of the screen are summarised in table 10. Of the nine clones, isolated in this project, six were shown to be stigma-specific (B6E, C10B, F4B, K2B, K8A2 and K8E1) so that two thirds of the clones, chosen for final characterisation, fell into the target category. This success rate, however, is tempered by the fact that two thirds of these stigma-specific clones were derived from previously identified transcripts. The three non-stigma specific clones could have been eliminated by differentially screening with stigma and whole flower mRNA probes. Although the floral probe would contain stigma-specific mRNA these would be sufficiently diluted by the mRNA's from the other organs to be rendered ineffective in detecting stigma-specific clones. This approach, however, was not chosen as it may have stopped the identification of clones of interest such as 15H11, which could be detected by the floral probe due to the presence of similar sequences in the petals.

Unlike, the work on the Solanaceae, this approach has not generated significant information on the molecular nature of the reproductive tissues. The predominant factor in this failure has been an inability to identify the function of the putative proteins encoded by the isolated cDNA's. This is a risk taken by anyone using this approach as the usefulness of any expressed gene sequences is determined by the identification of their function. An inability to identify randomly selected sequences is also seen in the large scale EST sequencing projects such as those reported by Newman et al. (1994) and Cooke et al. (1996) for *A.thaliana* where only about one third of the sequenced clones could be assigned identities. Of

Clone	Insert Size (bp)	Transcript Size (nt.)	Stigma-enriched Expression	Floral-specific Expression	Stress-Induced Expression	Identification
C10B	368	1465*	yes	no	N/D	SLG29
K2B	354	1465*	yes	no	N/D	SLG29
K8E1	~1000	1465*	yes	no	N/D	SLG29
B6E	584	1416*	yes	no	N/D	SLR1
15H11	776	~1200	yes	no	no	GRP
13G6	1001	~1200	no	yes	yes	?
C10C	373	~920	no	yes	yes	MLP?
F4B	798	~2800	yes	no	no	?
K8A2	519	~1400	yes	no	no	?
D5A2	1234	?	no	no	N/D	?
C10A	467	~1450	?	?	N/D	?

Table 10. Summary of the clones isolated and characterised during this PhD project. Insert and transcript sizes are exact if they were derived from a sequence and approximate if estimated from a gel. *=transcript size derived from Trick and Flavell (1989). N/D=not determined.

the investigated clones, only 15H11 and the S-sequences have been identified, with the former identification only possible because of the extreme amino acid bias of the predicted protein. Isolation of the latter clones would be predicted as they constitute a high proportion of the RNA species in the stigma (Nasrallah et al., 1985). These sequences have already been extensively investigated (Nasrallah and Nasrallah, 1993), and so their isolation here provides no new information with respect to the biology of the female tissue.

Therefore in this study the differential screening approach has not proven entirely successful. Alternative approaches that could have been used are discussed in 5.3.

5.1.2 Quality of the isolated clones.

A major factor in the identification problem is the relatively poor quality of the clones. Only three of the reported clones (13G6(i), 15H11, and K8E2) represent more than 50% of their respective transcript (although D5A2 may also represent greater than 50% of its respective transcript, for which an accurate size could not be determined) and one of these represents the already characterised S29 SLG.

Isolation of partial clones could have been due to a bias towards small inserts in the library, even though preliminary analysis of insert size (section 4.2.2) suggested that the clone size of the library was relatively high. However, this study used a limited number of clones and it is possible that the large 2.2kb clones may have been partial sequences of even longer transcripts. Therefore the inserts in the library may have been of too poor quality to obtain suitable amounts of sequence information. However the isolation of small inserts may also have been a matter of luck. To determine which of these two factors were responsible sizing of a more extensive number of clones would need to be carried out, although due to time limitations this was not done.

Some of the smaller clones (C10C and K2B) could have been eliminated by more rigorous size-selection during library production by using Sephacryl S-400 rather than S-300 spin columns. This modification would have only eliminated the smallest clones, however, while leaving clones such as C10A from which no coding information could be derived, still in the library.

A more suitable approach would have been the elimination of small clones early in the screening procedure. Such a step could have been achieved by PCR of clone mixtures, after the primary screen, then differential screening of the PCR products to identify the longest clones. This approach is not perfect as the longer cDNA clones may still only represent partial sequences while small full-length sequences may be lost. However selection for longer length clones would have provided more coding sequence information and therefore enhanced the chances of identifying any homologies to known sequences.

5.1.3 Molecular differences between the Solanaceae and Cruciferae.

Although identification of function has not been possible in most cases these clones still provide information on the Brassica stigma and its relationship to the stigmas and styles of the Solanaceae. As previously stated the different physiological characteristics of the stigma surfaces of these two families would strongly suggest different patterns of gene expression around these tissues. Similarly the low levels of protein in the transmitting tissue of the Brassicaceae, compared to the Solanaceae, suggests that the

molecular environment in this specialised tissue may also be different. Despite this, underlying molecular similarities, representing a common ancestry for the two families have not been ruled out

Most of the Solanaceous genes/proteins are expressed in the transmitting tissue whereas the stigma-expressed genes isolated here have not been localised. Therefore while the two sets of sequences can be compared the fact that they may come from different tissues must be kept in mind. Not withstanding this, and the limitations of the sequence information, the set of clones isolated here appears to represent different classes of sequence from those previously identified in the Solanaceae. While this could be used as evidence that there are significant molecular differences between the carpels of the respective families, the limited number of clones presented here is insufficient to make such judgements on alone.

This data does however correlate with data from other workers who have shown differences between the two families. S-RNases have not been detected in *Brassica*, while the level of RNase activity in the stigmas and styles of the Cruciferae is lower than the self-incompatible Solanaceae. Similarly there have been no reports of S-like sequences in the solanaceous reproductive organs, although absence has not been confirmed as no positive searches for such sequences have been published. Similarly attempts to identify solanaceous stigma/style PR-sequences and cell wall proteins in other families have been unsuccessful (Harikrishna et al., 1996; Sommer-Knudsen et al., 1996).

5.2 Stress genes in Brassica and Nicotiana.

5.2.1 The Solanaceous stress-induced sequences.

Although their biochemical function is not known 13G6(i) and C10C can be classified with many of the solanaceous stigma/style genes as they are members of sequence families whose other members are expressed in the vegetative organs under stress. The function of the stress-induced family members in the Solanaceae is not currently known, although from the available evidence, two possible physiological roles for these sequences can be proposed:

(a) a constitutive defensive role in which the proteins encoded by the respective genes have similar biochemical functions to those expressed under stress in the vegetative tissue (Wemmer et al., 1994; Atkinson et al., 1993).

(b) a specific role in the floral tissue where the gene products are directly involved in the physiology of the flower. This does not necessarily preclude similar biochemical activities to those seen in defence as several analogies have been drawn between pathogen invasion of plant tissue and pollination (Hodgkin et al., 1988).

For the solanaceous sequences much of the argument around these two models comes from the proposed biochemical activities of the relevant proteins. Evidence for a role analogous to that seen in defence comes principally from the lack of potential substrates or deposition, rather than degradation, of

substrate in the floral tissue. Such evidence, however, does nothing to promote the defensive-role hypothesis other than in a negative way by offering objections to the rival model.

Positive evidence comes from Jung (1956) who reported that fungal hyphae growing on the carpel were arrested in the style in a region equivalent to the expression of the ß(1,3)-glucanase and chitinase transcripts. The pattern of expression between families, however would appear to contradict this as high levels of the two transcripts have not been reported together in the same plant (Harikrishna et al., 1996), whereas in all stress response situations the two enzyme activities are induced together (Bowles, 1988). Moreover, from cross-hybridisation studies the presence of one transcript appears to preclude the other: mRNA's that cross-hybridise with stylar (1,3)ß-glucanase and chitinase have only been found in the Cestroidae and Solanoidae subfamilies respectively (Harikrishna et al., 1996). However, this inability to detect the mRNA/proteins may be due to sequence divergence.

The tissue-specific role hypothesis overcomes the substrate problem by proposing that the proteins may have previously unreported activities that predominate in the flower. This is analogous to the model proposed for the embryonic chitinases (De Jong et al., 1992) which are essential for the development of the carrot embryo where no chitinous material has been detected. Such an activity could allow the relevant stress-induced genes to be implicated as key components in plants where rapid changes in the cell wall and ECM occur as this is the principal shared characteristic of pollination, embryogenesis and wound/pathogen response. Indeed it may prove that the currently defining anti-pathogen activities of many of the stress-induced enzymes may be secondary to their activities towards the plant itself.

Experimental evidence to support this model is lacking in pollination, The restricted expression in subzones of the carpel could be considered evidence for a specific role in the biology of those tissues as a more widespread expression would be expected if the relevant proteins were to act against pathogens. However, once again this is negative evidence in that it utilises experimental results against the opposing hypothesis rather than for the proposed model. The restriction of expression of different transcripts to different sub-families is also a problem for this model as it implies that the proteins are not key components of pollination otherwise they would be expressed throughout the Solanaceae family.

To support the physiological-role-in-pollination hypothesis, second activities for these floral sequences need to be demonstrated, together with evidence linking this activity with the biochemical constitution of the pistil and/or developing pollen. However, no such evidence has been reported.

A second approach to resolving which of these two models are correct would be to down regulate expression of the relevant genes by the expression of antisense constructs. Transgenic plants could then be assayed for correct carpel structure and the ability to support pollination. If either of these two factors are affected by the suppression of a gene then this would provide conclusive evidence for the floral-role hypothesis. A negative result, however, would not provide any conclusive proof for the defensive-role hypothesis as it is probable that multiple factors act in the stigma to promote pollination and so elimination of one may not be sufficient to derail the ability of pollen to grow through the tissues of the stigma. Similarly the presence of multiple factors means that testing for increased pathogen sensitivity in the carpels of transgenic plants may also lead to negative results even though the gene in question may

play a defensive role. However, given the development of routine methods for transformation of the *Nicotiana* and *Solanum* (Klee et al., 1987) such experiments are still justifiable.

5.2.2 Similarities and differences between the Brassica and Solanaceous stress-induced sequences.

The two hypotheses for stress genes in the Solanaceous carpel can also be applied to the two expressed sequences from *B.oleracea,* that have been identified in this work. The expression of such sequences could act as an initial reference point for defining similarities in carpel biology between the two plant families. However, differences do occur between the sequences from the two families which may preclude direct comparisons to be drawn.

(i) 13G6(i) and C10C represent different transcripts and putative proteins from those seen in the Solanaceae. 13G6(i) definitely represents a novel stress-induced sequence while C10C may represent a transcript encoding a protein similar to the MLP's. As such, both cDNA's could be considered to represent different classes of gene, compared to those in the Solanaceae which encode proteins predicted to act against attack by other organisms and so may play a different role in the flower.

(ii) The relationship between the floral and stress-induced gene family members differs between the two families. The *S.tuberosum* chitinase and *N.tabacum* proteinase inhibitor show 66% and 50% identity to their nearest stress-induced relatives (Wemmer et al., 1994; Atkinson et al., 1993) while the putative tomato γ-thionin protein shares only 35% identity to its most homologous non-floral relative (Milligan and Gasser, 1995). This is contrast to the floral and wound induced 13G6(i) transcripts which, from the hybridisation experiments, must share >90% identity while the light-starved 13G6(i) and salicylic acid-induced C10C transcripts share >79% and >81% identity to the isolated floral sequences. Therefore these two cruciferous floral-expressed sequences show significantly greater levels of sequence identity to their stress-induced homologues than do most of the solanaceous sequences. The one exception to this trend is Pr10a, for which the stress-induced and stigma sequences are identical, being expressed from the same gene.

The implication of these sequence differences is that, in these more closely related nucleic acid sequences, the encoded proteins are more likely to have greater homology and so there is less scope for the differently expressed proteins having differing activities and functions. Indeed if the reported 13G6(i) cDNA and genomic sequences do represent the two different gene copies then the expressed proteins, of this gene family, are predicted to be <2% different from each other. This suggests that the model whereby the

stress-induced gene sequences play a floral specific role, via a different biochemical activity to that evident in defence, is less likely in the Cruciferae than in the Solanaceae.

(iii) the expression patterns of the clones are different. Both 13G6(i) and C10C are expressed throughout the flower with all assayed organs having roughly equivalent levels of expression, with the exception of C10C in the stigma. This is in contrast to the solanaceous clones which tend to hybridise with transcripts in carpel RNA only, the one exception being the γ-thionin which is also expressed in mature petals and immature stamens.

This restricted expression is one of the principal arguments for a specific role in the carpel for these sequences in the Solanaceae. Such a specific role is, therefore, less likely for the genes represented by C10C and 13G6(i) whose expression across the flower suggests that their encoded proteins are not specifically involved in defining the physiology of the specialised carpel tissues. Moreover, given the different morphologies and physiologies of the floral organs a specific physiologically-defining function across all the whorls is not immediately apparent .

This lack of organ-specificity in the expression of C10C and 13G6(i) coupled with the high level of identity between their respective floral and stress-induced proteins, suggests that a floral-specific activity for the encoded proteins is unlikely. This in turn would suggest that the proteins encoded by these clones play a defensive role in the flower. This is supported by the fact that 13G6(i) transcript levels increase after bud picking so suggesting that at least these additionally expressed transcripts are acting defensively

However, given the unknown biochemical nature of the predicted proteins, encoded by C10C and 13G6(i), a third option is available. This is that the biochemical and physiological activity is the same in both the flower and under stress but the proteins play a role important for the biology in both systems. Campillo and Lewis (1992) have previously reported the presence of many cell-wall metabolising enzymes, including β(1,3)glucanase, that are present in both flowers and abscission zones. In the former these enzymes have been suggested to be involved in major cell wall disruption, a process which in floral tissue is analogous but not identical to that occurring in abscission. Therefore the proteins encoded by 13G6(i) and C10C may also play a role in such processes, although what such a role might be is open to speculation. Whatever the role was, it would have to be continuous throughout the development of the flower, given the expression of transcripts similar to the two clones. Such a role could also be predicted for some Solanaceous stress genes where the specific patterns of spatial and temporal expression could prove more easy to relate to a specific physiological process than is the case with the Cruciferae.

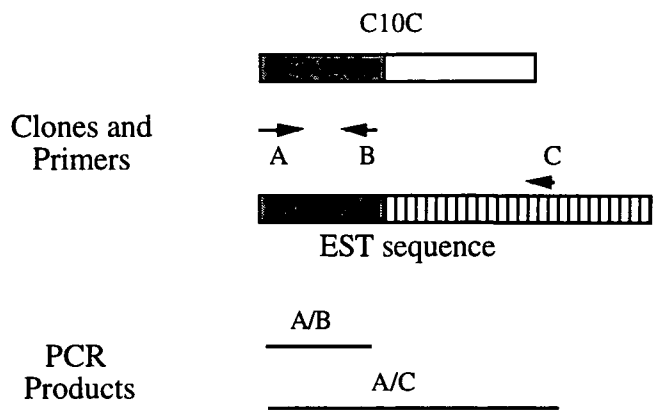
5.2.3 Determination of the role of the Brassica sequences in the flower.

The expression of members of stress-induced gene families in the reproductive tissues of the Solanaceae and Cruciferae is the one common link, between the two families, identified in this work.

Therefore elucidation of the role of these genes in *Brassica* would provide a starting point for the comparison of the specialised carpel tissues in these two families at the molecular level. Further analysis of these sequences may also prove useful if other stress induced sequences are isolated from cruciferous carpel tissue as data on the role of 13G6(i) and C10C may prove to be applicable to similar sequences More data to help resolve the question of the role of the 13G6(i) and C10C proteins could come from the determination of their functions. However, while the relationship of the former to latexes may enable determination of its role, determination of the function of the latter may require an element of luck, as discussed in section 4.1.1.11, and so cannot be relied upon.

Direct determination of the role of these genes by knocking out their expression may be the best approach to ascertaining their function. Given the powerful genetics available for *A.thaliana*, antisense gene transformation may be more productive, in that organism, if similar patterns of expression to *B.oleracea*, can be demonstrated for these two genes.

(a)



(b)

Outcome	A/B product	A/C product	Transcript expression in flower
1	No	No	No C10C-like transcripts in organs
2	Yes	No	C10C but not class1 extended transcripts present
3	No	Yes	Not Possible
4	Yes	Yes	C10C and/or class 1 transcripts present

Fig.69. PCR experiment to determine the presence and nature of C10C-like transcripts in the floral organs of *A.thaliana*. (a) relationship of the primers and predicted PCR products to the C10C and extended class 1 EST sequences. Shading indicates the C10C coding region common to the two clones while the vertically lined block indicates the extended coding region in the EST. (b) possible outcomes of the PCR experiment. Outcome 2 is the desired outcome. Outcome 4 is ambiguous as either extended or C10C-like and extended transcripts could be present. The PCR product could be labelled and used to probe a blot containing the relevant RNA's from *Brassica* and *Arabidopsis*. If the probe hybridises to only an *Arabidopsis* transcript larger than in *Brassica* then this would indicate only extended transcripts while if two or more transcripts were to be present, one of which was roughly similar in length to that in *Brassica* then C10C-like transcripts could be predicted to be present

Such a pattern cannot be directly inferred given the differences between the two Crucifers, in gene copy number for 13G6(i) and in the sequence for C10C. Therefore similar patterns of expression in the related crucifers would need to be confirmed. Given the problems associated with harvesting suitable quantities of *Arabidopsis* floral organs for Northern analysis, a PCR approach may be more effective than Northern analysis. Using primers that bind to both the *Brassica* sequences and the available *Arabidopsis* EST's, semi-quantitative PCR could be used to assess 13G6(i) expression in the latter organism. Analysis of C10C would be more complex, given the possible differences between the *B.oleracea* and *A.thaliana* sequences and so a three primer system could be used (fig.69). One common primer (primer A), that binds to the common 5' sequence could be used with a primer for the 3' end of the common coding sequence (primer B) and with a primer to the coding region of the *Arabidopsis* EST's (primer C). The first combination of primers would determine if and where C10C-like transcripts are expressed. The second combination of primers would determine if these transcripts were of the C10C or extended EST class.

If similar expression patterns were to be confirmed then antisense knockout of the expressed sequence in *Arabidopsis* could be performed. If this affects the morphology of the flowers then this would imply a specific role in the flower and from the generated phenotype some clues to the physiological/biochemical role of the relevant proteins may be obtained. However, no determinable effect on the phenotype may be obtained due to conditions discussed for similar work in the Solanaceae. If this is so then determination of the role of this class of genes may rely on determination of their biochemical role and linking of that role with physiological processes in the flower.

5.3 Other molecular approaches to investigating the carpel in the Cruciferae.

The approach taken to investigate the molecular biology of the female reproductive organ of *Brassica* has provided little insight into the biology of this organ. Much of this can be attributed to the lack of coding information obtained, although there is no reason to believe that isolation of full length clones would provide enough data to identify the encoded proteins, or their function. So, while the approach used in this project is guaranteed to generate data, the possibility that the genes expressed in the stigmas of the Cruciferae may code for proteins of unknown function would leave most of this data as useless in the investigation of the biology of this tissue. Therefore continuation of this work must consider alternative approaches to studying the molecular biology of these tissues. Three broad approaches are available using already established molecular methods

5.3.1 Isolation and characterisation of mutants.

Although, due to its size, it is not the obvious model system for studying stigma/style and pollination biology, the powerful genetics available for *Arabidopsis* means that this is the ideal system for isolating genes by mutation. One approach would be to mutate plants and then screen them for the inability of stigmas and styles to support wild-type pollen growth and subsequent fertilisation. Such an

approach has the advantage that it would allow the isolation of stigma/style genes which by definition are essential for the pollination process. However this approach also has several drawbacks which may prevent limit its use:

(i) Mutants may not be generated. This could be due to gene redundancy or the previously discussed possibility that the components of the stigma and transmitting tract may provide an environment where the loss of one component may produce a cryptic or undetectable phenotype. Similarly once mutants had been isolated each gene would have to be mapped and isolated.

(ii) The strategy would be very labour intensive as each plant in the screen would have to hand pollinated then the relevant carpels removed and stained with aniline blue to assess pollen development.

(iii) It is possible that key genes involved in pollination may play a role in other aspects of plant development and so viable mutants may not be generated in the screen. The common presence of chitinases and glucanases in reproductive tissues and embryo development suggests that this may be a real possibility.

(iv) maintenance of the mutant stocks would be difficult as if the carpels do not support fertilisation then the mutants would be essentially sterile, although mutants that allow some fertilisation could still be isolated. Similarly if the carpel does not support pollination then crossing mutations to determine which mutant plants are mutated in the same gene would not be possible. The maintenance problem could be overcome via the ability to generate embryos and then new plants from developing pollen (Zaki and Dickinson, 1990), assuming the stamen is unaffected by the mutation. The crossing problem, however, has no obvious solution so that each mutant plant would have to be genetically mapped to identify its mutation.

Therefore, although the mutation approach may help the isolation of key genes from the carpel, involved in pollination, there are sufficient theoretical problems to suggest that it would not guarantee success. Even if it were successful, it may only allow the identification of a subset of genes, while others, not being essential for pollination, would not be isolated by this approach. It also carries the doubt that even if any genes are isolated, other than what can be deduced from the phenotypic observations, their biochemical/physiological function may still be cryptic

5.3.2 Promoter trapping

Many of the problems associated with the isolation of mutants could be overcome by using a promoter trapping approach whereby mutation is mediated by insertion of transposons containing reporter gene constructs that are activated by integration next to promoter sequences (Topping et al., 1994). Mutants are then isolated by identification of reporter gene expression in the tissues of interest. This would allow isolation of genes expressed in the specialised tissues of the stigma and style which might not show up on a phenotypic screen. The presence of transposon sequences would also act as a tag to allow isolation of the locus into which the transposon had integrated.

However, as with more classical mutation studies, integration into the genes of interest would inactivate them so that viable plants might not be produced if the genes are essential for other aspects of plant development.

5.3.3 Analysis of the proteins of the carpel.

Rather than the nucleic acid approach a protein approach could be taken. In the Solanaceae this has been successfully used to determine the role for TTS1 and the S-RNase in the pollination process. Isolation of proteins that may be involved in the transmitting tract and specialised stigma tissues, would allow sequence to be obtained from them which may aid identification or allow isolation of their respective cDNA's. However the work in the Solanaceae was performed on *N.alata* which has a relatively large carpel allowing easy harvesting of large quantities of tissue. None of the Cruciferae have carpels similar in size (*Brassica* and *Raphanus* are ~1cm long at maturity) so that harvesting of sufficient fresh tissue may prove an insurmountable obstacle for this such an approach.

Therefore work on the Cruciferae would have to involve approaches which require minimal quantities of tissue or which use tissue that can be harvested then stored frozen. The simplest is the sequencing of protein directly isolated by SDS-PAGE. However, this approach in its simplest form is not entirely suitable for *Brassica* stigma/styles given the lack of discrete protein bands after electrophoresis of their extracts (section 4.3.5.2). The use of 2D-separation methods may enable specific polypeptides to be identified, that could not be determined from 1D gels but as previously discussed many of these proteins may come from the non-specialised epidermal, parenchymal or vascular tissue of the stigma and style and so may not play a role in pollination.

As with the creation of the cDNA library, stigma material could be used instead of stigma/style as it contains higher concentrations of the tissues of interest, although once again, identified proteins may not be from the tissues involved in pollination. Therefore isolation of specific tissue may be necessary. Isolation of transmitting tissue cells from *N.alata* by gentle grinding and filtering has been reported (Gane at al., 1994), although as this requires fresh tissue it may not be possible to harvest sufficient Cruciferae material.

A complementary approach would be to wash the stigma surface to remove and isolate extracellular stigma components for analysis as performed in *Gladiolus* by Clarke et al. (1979). Once again, although the wash could be frozen and then different washes pooled it may not be possible to obtain sufficient quantities of tissue to generate enough protein for sequencing or identification by the use of specific stains.

Therefore a protein-based approach is unlikely to be practical given the size of the reproductive organs in the Cruciferae. Even if it was successful any sequence information may still not provide clues to the function of the protein. If the isolated protein could be purified it could be assayed for any effect on pollen growth *in vitro*, as has been performed with TTS-1 and the S-RNase, but such experiments are only of value if they produce positive results.

5.3.4 Extrapolation of other cellular systems.

The pollination process is not the only plant cell biological system being studied and so one approach to its analysis in the Brassicaceae might be to use results from other systems as a starting point for experiments. This approach could be particularly productive when considering intracellular signalling in pollination as a large body of evidence suggests that the core messenger systems discovered in animals also predominate in the plants (Ranjeva and Boudet, 1987; Soomarin and Sandelius, 1988; Roberts and Harmon,1992). The use of inhibitors for these components could help elucidate what, if any role they play in the Crucifer carpel. Some work of this kind has already been carried out with okadaic acid, an inhibitor of phosphatases (Rundle et al.,1993) but other signal cascade components such as G-proteins, via cholera and pertussis toxin, (Stryer and Bourne, 1986) and inositol, via lithium(Berridge, 1988) have not been investigated.

Such an approach is a relatively non-labour intensive method as only a few carpels would be needed for each initial experiment. PCR, antibodies or screening with clones for relevant genes from other systems could then be used to isolate the genes/proteins involved with the positively identified aspect of pollination

The usefulness of these experiments, however, is dictated by whether they produce positive results. Negative effects of inhibitors need not be due to the absence of that system: it could be argued that the enzymes in the carpel may differ from those in previously characterised systems or that the inhibitor may not be taken up by the carpel so that it can interact with its target (as has been reported for tunicamycin (Roberts et al., 1984b; Sarker et al., 1988)). Even if positive results are obtained, in pollination experiments it may not be possible to determine whether it is pollen, carpel or pollen and carpel factors which are being inhibited. A final problem with this method is that only cell biology processes that have been previously characterised can be studied. Therefore if aspects of pollination are unique to it and the tissues involved, which would be predicted to be so as it involves the only known case of directed cell movement in plants, then extrapolation of the molecular mechanisms from other biological processes would prove an inappropriate approach.

5.4 Summation.

In summary, although the experiments performed in this project have provided little information on the biology of the female reproductive tissue in *B.oleracea*, alternative approaches may prove just as poor in elucidating the molecular basis for the biology of this organ. A combination of differential screening coupled with the above methods could be used, though such an approach would be highly dependant on a large supply of carpel material to support all possible strands. Despite this the pollination process still provides a potentially lucrative model system for studying aspects of plant cell biology.

6
Appendices

Appendix A Electrophoresis size markers.

All sizes on agarose/acrylamide gels and associated blots were obtained from the co-running of the following size markers.

DNA size markers (bp).

kb ladder	λBstEII	λHindIII	λAvaII
12000	8454	23130	8126
11000	7242	9416	6555
10000	6369	6557	6442
9000	5686	4361	3676
8000	4822	2322	2605
7000	4324	2027	2555
6000	3675	564	2134
5000	2323	125	2005
4000	1929		1951
3000	1371		1612
2000	1264		1420
1500	702		1284
1000	224		985
500	117		974
400			894
300			597
200			590
			513
			511
			433
			398
			345
			310
			308
			272
			242
			215
			151
			8 <100bp fragments

RNA size markers (nt)

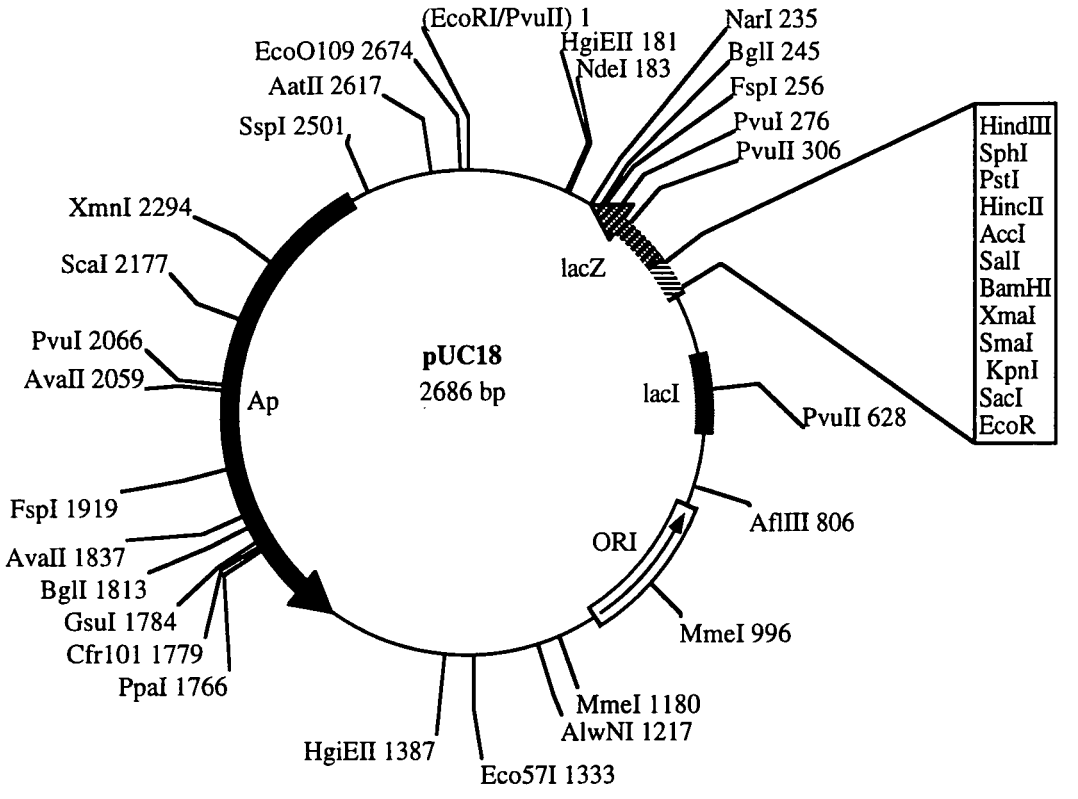
Promega RNA marker	Ribosomal RNA
6583	~3300
4981	~1900
3638	
2604	
1908	
1383	
955	
623	
281	

Protein size markers (Da).

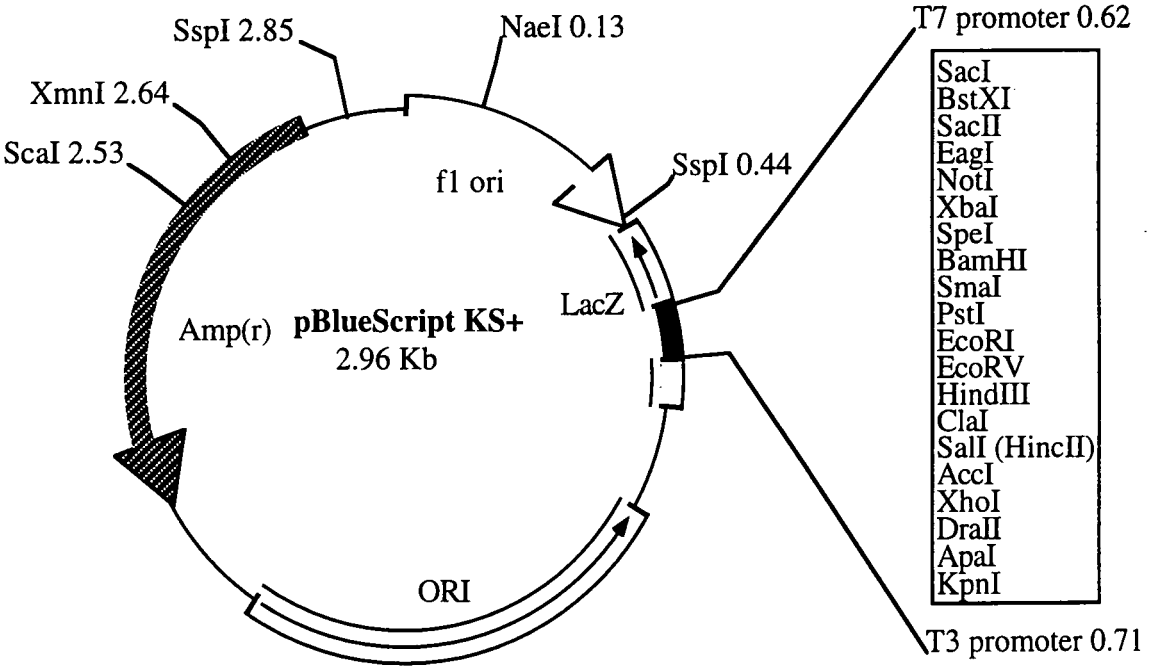
66000	Albumin (bovine)
45000	Albumin (egg)
36000	Glyceraldehyde 3-phosphate dehydrogenase
29000	Carbonic anhydrase
24000	Trypsinogen (PMSF treated)
20000	Soybean trypsin inhibitor
14000	Lactalbumin (bovine milk)

Appendix B Restriction maps of vectors.

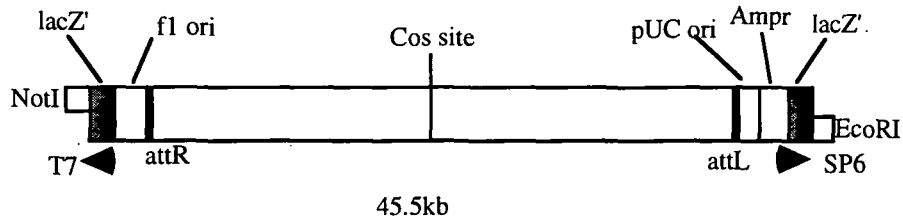
(i) pUC18 as from Sambrook et al., 1989



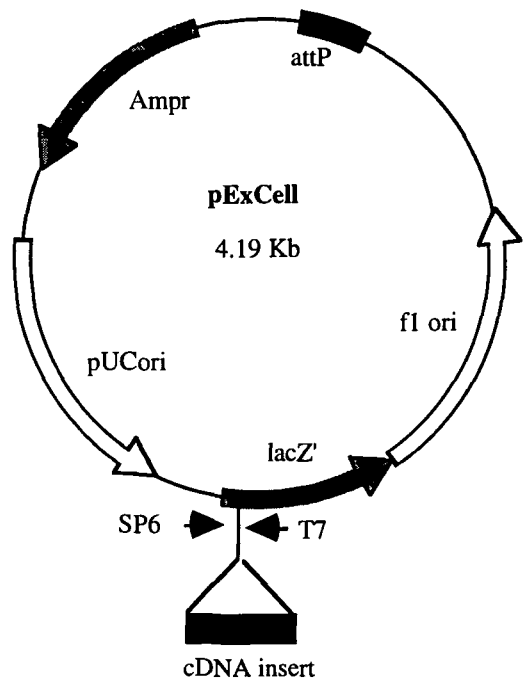
(ii) pKS+Bluescript as from Stratagene catalogue (1996)



(c) lambda ExCell phage vector as from Promega catalogue (1995)



(d) pExCell, phagemid obtained by in vivo rescue from lambda ExCell from Promega catalogue (1995)



Appendix C EMBL Database entries for clones sequenced in this project.

All the sequences reported in this PhD project, with the exception of the SLG/SLR sequences which have been reported previously, were submitted to the EMBL database using the AUTHORIN program for the Apple Macintosh. The sequences were given the following accession codes.

Clone	Entry name	Accession name
13G6(i) cDNA	BO13G6IC	Z74950
13G6(i) genomic sequence	BO13G6IG	Z74962
15H11	BO15H11	Z74892
C10C	BOC10C	Z74875
F4B	BOZ74821	Z74821
K8A2	BOK8A2	Z74921
C10A	BOC10A3UT	Z74963
D5A2	BOD5A2	Z74841

The full entries are as follows:

```
ID   BO13G6IC   standard; RNA; PLN; 985 BP.
XX
AC   Z74950;
XX
NI   e1001208
XX
DT   29-JUL-1996 (Rel. 48, Created)
DT   29-JUL-1996 (Rel. 48, Last updated, Version 1)
XX
DE   B.oleracea mRNA (unknown)
XX
KW   .
XX
OS   Brassica oleracea
OC   Eukaryota; Plantae; Embryobionta; Magnoliophyta; Magnoliopsida;
OC   Dilleniidae; Capparales; Brassicaceae.
XX
RN   [1]
RP   1-985
RA   Booker J.P.;
RT   ;
RL   Thesis (1996), Biological Science, University of Durham
XX
RN   [2]
RP   1-300
RA   Scutt C.P.;
RT   ;
RL   Thesis (1990), Biological Science, University of Durham
XX
RN   [3]
RP   1-985
RA   Croy R.D.;
RT   ;
RL   Submitted (26-JUL-1996) to the EMBL/GenBank/DDBJ databases.
RL   Ronald R.D. Croy, Biological Sciences, University of Durham, South
RL   Road, Durham, County Durham, DH1 3LE, United Kingdom
XX
FH   Key             Location/Qualifiers
FH
FT   source          1..985
FT                   /organism="Brassica oleracea"
FT                   /strain="gemifera"
FT                   /clone="13G6(i)"
FT                   /dev_stage="mature"
FT                   /tissue_type="stigma"
FT                   /cell_line="Plant line S29"
FT                   /clone_lib="lambda ExCell"
FT   polyA_site      985
FT                   /evidence=EXPERIMENTAL
FT   mRNA            &lt;1..985
FT                   /note="wound and dark induced"
FT   CDS             &lt;1..781
FT                   /codon_start=2
FT                   /gene="unknown"
FT                   /note="protein similar to bacterial YRN1 and HEAHIO
FT                   proteins"
FT                   /db_xref="PID:e256044"
FT                   /translation="LDRTIQFYTECFGLKVLRKRDVPEEKYSNAFLGFGPETSNFVVEL
FT                   TYNYGVSSYDIGTGFGHFAISTQDVSKMVEAVRAKGGNVTREPGPVKGGGSVIAFVKDP
FT                   DGYTFELIQRGPTPEPLCQVMLRVGDLDRAIKFYEKALGMRLLRRIERPEYKYTIGMMG
FT                   YAEEYESIVLELTYNYGVTEYTKGNAYAQIAIGTDDVYKSAEVVKIANQELGGKITREA
FT                   GPLPGLGTKIVSFLDPDGWKTVLVDNEDFLKELNEEE"
FT   3'UTR           782..985
FT   polyA_signal    816..821
FT   polyA_signal    944..949
XX
SQ   Sequence 985 BP; 273 A; 184 C; 249 G; 279 T; 0 other;
     ccttgatcgc actatccagt tttacacaga gtgctttggc ttgaaggtgt tgaggaaaag        60
     agatgtacct gaggagaagt actctaatgc tttccttggt tttggtcctg aaacctccaa       120
     cttcgttgtc gagctcactt acaattatgg tgttagctca tatgacattg gaactggatt       180
     tgggcatttc gctatttcaa ctcaagatgt ttctaagatg gttgaggcgg ttcgtgccaa       240
     gggtggaaat gtcaccagag agcctggtcc ggtcaaaggt ggaggcagtg tcattgcctt       300
     tgtcaaggac cctgatggtt acacttttga gcttatccag agaggtccaa ctcctgaacc       360
     tctctgtcaa gtcatgcttc gtgttggtga tcttgaccgc gccatcaagt tctatgaaaa       420
     ggcccttggg atgagactct tgagaaggat tgagaggcct gaatacaagt acaccatagg       480
     catgatggga tatgctgagg aatatgagtc aatagttttg gagctgacct ataactacgg       540
     tgtgactgag tacacaaagg gcaacgcata tgcacagatt gcaataggca cggatgatgt       600
     gtacaaaagc gctgaagtag tgaagatagc caaccaagag ctaggaggaa agatcactag       660
     agaagccgga cctcttcctg gactcggcac caagattgtc tcattcctcg atccagatgg       720
     ctggaaaaca gttctggtag acaacgaaga ttttctgaag gaactgaatg aggaagagta       780
     atgatggtga tgatcacaga caaaaccata ataataataa aatgtgatct tatgaagtag       840
     agtgaagact tagtatttaa tctaaataag atgttggttc agtgtgcgtc gtgtttctat       900
     gtgtttgttc cttgtctttg agctaaactc tgttacgtgc ctgaataaat atatttccac       960
     cacctggttt aggtgctttc cggtt                                            985
//
```

```
ID   BO13G6IG   standard; DNA; PLN; 2843 BP.
XX
AC   Z74962;
XX
NI   e1001209
XX
DT   29-JUL-1996 (Rel. 48, Created)
DT   29-JUL-1996 (Rel. 48, Last updated, Version 1)
XX
DE   B.oleracea mRNA (unknown)
XX
KW   .
XX
OS   Brassica oleracea
OC   Eukaryota; Plantae; Embryobionta; Magnoliophyta; Magnoliopsida;
OC   Dilleniidae; Capparales; Brassicaceae.
XX
RN   [1]
RP   1-2841
RA   Booker J.P.;
RT   ;
RL   Thesis (1996), Biological Science, University of Durham
XX
RN   [2]
RP   1-2843
RA   Croy R.D.;
RT   ;
RL   Submitted (26-JUL-1996) to the EMBL/GenBank/DDBJ databases.
RL   Ronald R.D. Croy, Biological Sciences, University of Durham, South
RL   Road, Durham, County Durham, DH1 3LE, United Kingdom
XX
FH   Key             Location/Qualifiers
FH
FT   source          1..2843
FT                   /organism="Brassica oleracea"
FT                   /strain="gemifera"
FT                   /clone="13G6(i).lambda 10"
FT                   /cell_line="plant line S29"
FT                   /clone_lib="lambda fix"
FT   mRNA            join(&lt;797..894,981..1105,1180..1244,1323..1533,1623..1670,
FT                   1763..1867,1966..2117,2229..2490)
FT                   /evidence=EXPERIMENTAL
FT                   /note="mRNA structure deduced from homology to 13G6(i)
FT                   cDNA"
FT   CDS             join(797..894,981..1105,1180..1244,1323..1533,1623..1670,
FT                   1763..1867,1966..2117,2229..2273)
FT                   /gene="unknown"
FT                   /note="protein similar to bacterial YRN1 and HEAHIO
FT                   proteins"
FT                   /db_xref="PID:e256075"
FT                   /translation="MAENADLVEWPKKDKRRFLHVVYRVGDLDRTIQFYTECFGMKVLR
FT                   KRDVPEEKYSNAFLGFGPETSNFVVELTYNYGVSSYDIGTGFGHFAISTQDVSKMVEAV
FT                   RAKGGNVTREPGPVKGGGSVIAFVKDPDGYTFELIQRGPTPEPLCQVMLRVGDLDRAVK
FT                   FMEKALGMRLLRRIERPEYNTIGMMGYAEEYESIVLELTYNYGVTEYTKGNAYAQIAIG
FT                   TDDVYKSAEVVKIVNQELGGKITREAGPLPGLGTKIVSFLDPDGWKQVLVDNEDFLKEL
FT                   E"
FT   3'UTR           2271..2490
FT   polyA_signal    2322..2327
FT   polyA_signal    2449..2454
FT   exon            &lt;797..894
FT                   /evidence=EXPERIMENTAL
FT                   /number=1
FT   intron          895..980
FT                   /evidence=EXPERIMENTAL
FT                   /number=1
FT   exon            981..1105
FT                   /evidence=EXPERIMENTAL
FT                   /number=2
FT   intron          1106..1179
FT                   /evidence=EXPERIMENTAL
FT                   /number=2
FT   exon            1180..1244
FT                   /evidence=EXPERIMENTAL
FT                   /number=3
FT   intron          1245..1322
FT                   /evidence=EXPERIMENTAL
FT                   /number=3
FT   exon            1323..1533
FT                   /evidence=EXPERIMENTAL
FT                   /number=4
FT   intron          1534..1622
FT                   /evidence=EXPERIMENTAL
FT                   /number=4
FT   exon            1623..1670
FT                   /evidence=EXPERIMENTAL
FT                   /number=5
FT   intron          1671..1762
FT                   /evidence=EXPERIMENTAL
FT                   /number=5
FT   exon            1763..1867
FT                   /evidence=EXPERIMENTAL
FT                   /number=6
FT   intron          1868..1965
FT                   /evidence=EXPERIMENTAL
FT                   /number=6
FT   exon            1966..2117
FT                   /evidence=EXPERIMENTAL
FT                   /number=7
```

```
FT   intron          2118..2228
FT                   /evidence=EXPERIMENTAL
FT                   /number=7
FT   exon            2229..2490
FT                   /evidence=EXPERIMENTAL
FT                   /number=8
FT   polyA_site      2490
FT                   /evidence=EXPERIMENTAL
XX
SQ   Sequence 2843 BP; 789 A; 488 C; 616 G; 950 T; 0 other;
     tctagagaag ttaaattcaa aacatttgtt atgacatgtc ttttgttcat atggcatatg        60
     catccaaaca ttaatatggt ttgagattag aaaaatagat tgagatctaa atattttagt       120
     tatttttatt acatgcatca tgcatgtaga tgaaaaatgt aaaattcaaa acaaactaat       180
     caaacaaaaa aaattcatgg agtgggcttt tctacttcat ctatcaattc atatgtatgg       240
     gctttacacc catgtgcctt gggctctttt ctcttatcat taataataca cttcattatt       300
     gtactttctg ccgacaaaaa aaaatcgttt atgtacgcca cgtgtatccc catgggcgtt       360
     taggtcgcat cccatcgctc cccaaacaat taatctaacg gctataccgc gtctgtaatg       420
     gttacaccac atcacaggca gtatccgaga agttatcagt cacagttttt ttttttata        480
     agtacagtga cggacttagc ttttttatca aagcccattg ttggtggtgt atctgctccg       540
     tttcgttaat cttctcccgg gtgagtttct gttgcgcacg atctcgtaat gatcgcattc       600
     tcgttaatcg cgtcatgaat catgtgattg ttgtttattt ttgatccatt tgatgatatt       660
     ttccggatcg gcgttaattg ttgtcattat tcgattgatc ttatgggttt tatgggattt       720
     aagtgttacg atgatgatca ttgagaaagg tgtttaaagg tttggatctt tatctctctt       780
     ctttttttgt agaaaaatgg cagaaaatgc tgatttggtg gagtggccaa agaaggataa       840
     gcgtcgtttt ctccatgttg tctaccgcgt tggtgatctt gatcgcacta tccagtatct       900
     ctctttcctt tactctgttt cattcataca ctattgtgtt attcattcta tatactaaca       960
     ggagttgttt tcttgaatag gttttacaca gagtgctttg gcatgaaggt gttgaggaaa      1020
     agagatgtcc ctgaggagaa gtactctaat gctttccttg gttttggtcc tgaaacctcc      1080
     aacttcgttg tcgagctcac ttacagtata tgcctcttat ctctccttgt ctacacaaaa      1140
     gtgatattac ttgactctta tgagaaatct ggtttacaga ctatggtgtt agctcatatg      1200
     acattggaac tggatttggg catttcgcta tttcaactca agatgtaagt tgtcctcatg      1260
     ctttgattat ttagatttct aggatcagtc ggtccttgaa tgtggagaga gttttgtttc      1320
     aggtttccaa aatggttgag gccgttcgtg ccaagggtgg aaatgtcacc agagagcctg      1380
     gtccggtcaa aggtggaggc agtgtcattg cctttgtcaa ggaccctgat ggttacactt      1440
     ttgagcttat ccagagaggt ccaactcctg aacctctctg tcaagtcatg cttcgtgttg      1500
     gtgatcttga ccgcgccgtg aagtttatgg aaaaggtaat gcaacgccct tgtataaggc      1560
     aatgaacatt gctgttttct agaaagcttc ttgcttttat atctcacttt tagaatccat      1620
     gcaggccctt gggatgagac tcttgagaag gatcgagaga cctgaataca aggtgaacta      1680
     atcaatgata cttcttagac aagtagttta ccctttgag atacgttgcg actcttaacg       1740
     agtttttgtt taatgaaaca gtacaccata ggcatgatgg gatatgctga ggaatatgag      1800
     tcgatagttt tggagctgac ctataactac ggcgtgactg agtacacaaa gggcaacgca      1860
     tatgcacagg taccggctta agactattgc tactgtttat gttgtcttta aacggtttgt      1920
     gctctttctt ttgatcgtta acatttggtc ttaaatgtaa ttctcagatt gcaataggca      1980
     cggatgatgt gtacaaaagc gctgaagtag tgaagatagt caaccaagag ctaggaggaa      2040
     agatcactag agaagccgga cctcttcctg gactcggcac caagattgtc tccttcctcg      2100
     atccagatgg ctggaaaaca gtaagtaact gtgcatttga gagatggatg aatctggtga      2160
     ttttattaaa gattctactt agcttcattt tggtttttgc taagaagttt gtttggtggt      2220
     taaaatgaca ggttctggta gacaacgaag attttctgaa ggaacttgaa tgaggaagag      2280
     tgatgatgat gatggtgaaa gatcacatac aaaaccataa taataaaatg tgatcttatg      2340
     aagtagaggg aagactttgt aattaatcta aataagatgt tgcagtgtgt gtcgtgtttc      2400
     tatatgtttg ttctttttat ctttgtgagc tgaactctgt tacgtgcctg aataaatata      2460
     tttccaccac ctggtttagg tgttttccgg ttatggatgg tttagtagaa ttcaaaattt      2520
     ctgtaccaaa ttggtttact cgattgaaga cggattacaa tggataataa agatgcgaaa      2580
     gtaatgaatg tcttatgtct acaagactac aactcacaag aaagttgtat cacaaaagag      2640
     tacaaagata tggtgcaaag gtctcgtgga ggtcgttgtt taaaatcatc tcagaaacta      2700
     aaggatatgg agaaaggaca tcaagtggtt catcaccatc gctaaaaaaa agacaaggcg      2760
     tggcagaggt ggtgggcact gctcttggtc gtggatacca ctagaatttg aagatcatgg      2820
     agttcttcat attcccctgc tat                                             2843
//
```

```
ID   BO15H11    standard; RNA; PLN; 762 BP.
XX
AC   Z74892;
XX
NI   e1001206
XX
DT   29-JUL-1996 (Rel. 48, Created)
DT   29-JUL-1996 (Rel. 48, Last updated, Version 1)
XX
DE   B.oleracea mRNA for glycine-rich protein
XX
KW   glycine-rich protein.
XX
OS   Brassica oleracea
OC   Eukaryota; Plantae; Embryobionta; Magnoliophyta; Magnoliopsida;
OC   Dilleniidae; Capparales; Brassicaceae.
XX
RN   [1]
RP   1-762
RA   Booker J.P.;
RT   ;
RL   Thesis (1996), Biological Science, University of Durham
XX
RN   [2]
RP   1-762
RA   Scutt C.P.;
RT   ;
RL   Thesis (1990), Biological Science, University of Durham
XX
RN   [3]
RP   1-762
RA   Croy R.D.;
RT   ;
RL   Submitted (26-JUL-1996) to the EMBL/GenBank/DDBJ databases.
RL   Ronald R.D. Croy, Biological Sciences, University of Durham, South
RL   Road, Durham, County Durham, DH1 3LE, United Kingdom
XX
FH   Key             Location/Qualifiers
FH
FT   source          1..762
FT                   /organism="Brassica oleracea"
FT                   /strain="alboglabra"
FT                   /clone="C10C"
FT                   /dev_stage="mature"
FT                   /tissue_type="stigma"
FT                   /cell_line="plant line S29"
FT                   /clone_lib="lambda ExCell"
FT   polyA_site      762
FT                   /evidence=EXPERIMENTAL
FT   mRNA            &lt;1..762
FT                   /note="stigma/style specific expression"
FT   CDS             &lt;1..663
FT                   /product="glycine-rich protein"
FT                   /function="cell wall protein"
FT                   /codon_start=1
FT                   /db_xref="PID:e256074"
FT                   /translation="GGAGGGFGGGAGGGHAGGGAGGGHGGGAGGGFGGGAGGGHGGGAG
FT                   GGFGGGAGGGKGGGLGGGGGVGGGHGGGVGGGFGGGAGGGKGGGVGGGFGGGAGGGKGG
FT                   GLGGGHGGGVGGGFGGGAGGGKGGGVGGGHGGGAGGGFGGGAGGGHGGGAGGGFGGGAG
FT                   GGHGGGAGGGFGGGAGGGHGGGAGGGFGGGAGGGHAGGGAGGGFGGGAGGGGGGGF"
XX
SQ   Sequence 762 BP; 114 A; 78 C; 379 G; 191 T; 0 other;
     ggaggagctg gtggaggttt tggcggtggt gcaggaggag gccatgctgg cggtggagct        60
     ggtggaggtc acggtggtgg tgctggagga ggatttggcg gtggagctgg tggaggtcac       120
     ggtggtggtg ctggaggagg attcggcggt ggtgctggtg gaggtaaagg tggtggtctt       180
     ggtggtggtg gtggcgttgg aggaggccac ggtggtggtg ttggaggagg gttcggtggt       240
     ggtgctggtg gaggcaaagg tggtggtgtt ggaggagggt tcggtggtgg tgctggtgga       300
     ggcaaaggtg gtggtttagg tggaggtcac ggtggtggtg ttggaggagg atttggtggt       360
     ggtgctggtg gaggcaaagg tggtggcgtc ggaggaggcc atggtggtgg tgctggtgga       420
     gggttcggtg gtggtgctgg tggaggacat ggtggtggtg ctggaggagg atttggcggt       480
     ggagctggtg gaggtcacgg tggtggtgct ggaggaggat tcggcggtgg tgctggtgga       540
     ggccacggtg gaggagctgg tggaggtttt ggcggtggtg caggaggagg ccatgctggc       600
     ggtggtggtg ctggtggagg gtttggtggt ggtgccggag gaggcggtgg tggtggattt       660
     tgaatctagc ttttagggc tacttacttg tgtaccacgc caaaaaagaa tgagattata       720
     aatgctggga ttttattatg aatgaaataa atgatactta ac                         762
//
```

```
ID   BOC10C     standard; RNA; PLN; 349 BP.
XX
AC   Z74875;
XX
NI   e1001205
XX
DT   29-JUL-1996 (Rel. 48, Created)
DT   29-JUL-1996 (Rel. 48, Last updated, Version 1)
XX
DE   B.oleracea mRNA for putative major latex protein
XX
KW   major latex protein.
XX
OS   Brassica oleracea
OC   Eukaryota; Plantae; Embryobionta; Magnoliophyta; Magnoliopsida;
OC   Dilleniidae; Capparales; Brassicaceae.
XX
RN   [1]
RP   1-349
RA   Booker J.P.;
RT   ;
RL   Thesis (1996), Biological Science, University of Durham
XX
RN   [2]
RP   1-349
RA   Croy R.D.;
RT   ;
RL   Submitted (26-JUL-1996) to the EMBL/GenBank/DDBJ databases.
RL   Ronald R.D. Croy, Biological Sciences, University of Durham, South
RL   Road, Durham, County Durham, DH1 3LE, United Kingdom
XX
FH   Key             Location/Qualifiers
FH
FT   source          1..349
FT                   /organism="Brassica oleracea"
FT                   /strain="alboglabra"
FT                   /clone="C10C"
FT                   /dev_stage="mature"
FT                   /tissue_type="stigma"
FT                   /cell_line="plant line S29"
FT                   /clone_lib="lambda ExCell"
FT   mRNA            &lt;1..349
FT                   /note="expressed in flowers"
FT                   /note="induced by salicylic acid in leaf tissue"
FT   CDS             &lt;1..198
FT                   /codon_start=1
FT                   /product="major latex protein"
FT                   /note=" clone product similar to P.somniferum major latex
FT                   protein"
FT                   /db_xref="PID:e256043"
FT                   /translation="FLLTIQVTPKHGGEGVLFTGTLSMRRLATRLLIPRLSSSSVWKSP
FT                   KRSTNISWPRNKGMCVCVCF"
FT   polyA_site      349
FT                   /evidence=EXPERIMENTAL
XX
SQ   Sequence 349 BP; 94 A; 58 C; 90 G; 107 T; 0 other;
     ttcttgctca cgatccaggt gacccctaag cacggtgggg agggagtatt gttcactggc        60
     accttgagta tgagaagatt agcgacgagg ttgctcatcc cgagactctc ctccagttct       120
     gtgtggaagt ctccaaagag atcgacgaac atctcttggc caaggaataa aggaatgtgt       180
     gtgtgtgtct gtttttaata aagaaatgtg tgatgtgcct ttctttctgt attttaatta       240
     aacgcaataa gggattaagg agccaagtcc gcgtttgtga gagatataaa tgatgagtgt       300
     gccatgatta tgtgtgtata ataagatcaa ttacgtttga agttgtatc                   349
//
```

```
ID   BOZ74821   standard; RNA; PLN; 446 BP.
XX
AC   Z74821;
XX
NI   e1001203
XX
DT   29-JUL-1996 (Rel. 48, Created)
DT   29-JUL-1996 (Rel. 48, Last updated, Version 1)
XX
DE   B.oleracea mRNA (unknown)
XX
KW   .
XX
OS   Brassica oleracea
OC   Eukaryota; Plantae; Embryobionta; Magnoliophyta; Magnoliopsida;
OC   Dilleniidae; Capparales; Brassicaceae.
XX
RN   [1]
RP   1-446
RA   Booker J.P.;
RT   ;
RL   Thesis (1996), Biological Science, University of Durham
XX
RN   [2]
RP   1-446
RA   Booker J.P.;
RT   ;
RL   Submitted (26-JUL-1996) to the EMBL/GenBank/DDBJ databases.
RL   jonathan J.P. Booker, Biological Sciences, University of Durham,
RL   South Road, Durham, County Durham, DH1 3LE, United Kingdom
XX
FH   Key             Location/Qualifiers
FH
FT   source          1..446
FT                   /organism="Brassica oleracea"
FT                   /strain="alboglabra"
FT                   /clone="F4B"
FT                   /dev_stage="mature"
FT                   /tissue_type="stigma"
FT                   /cell_line="plant line S29"
FT                   /clone_lib="lambda ExCell"
FT   mRNA            &lt;1..446
FT                   /gene="unknown"
FT   polyA_site      446
FT                   /evidence=EXPERIMENTAL
XX
SQ   Sequence 446 BP; 120 A; 67 C; 103 G; 156 T; 0 other;
     gccagacctt caccggcgca cagatacgct aggaaggggt gatttggtaa aaattctcat        60
     atcatttgac caagtgaaaa ctagtggttc gtgtggtgtt gtatgtagta gaaagattgg       120
     agagggggca tattagtaat ttggtccgtt cttacctaaa aagaagtggt gcttcgagtg       180
     cttgattttc cacaaagcct tgatgacgtt tccttttcgg acctttctat tattggtaag       240
     gttttttttt aatactaagt tttattgatg cagaggtgga attaaggacc acctgctagt       300
     agtacattta tcgtgtttgt atcccctctt gaagcgagag ggagagtttc acctttcaag       360
     aattggtaga ttgtattaac gtcgttaagt catcatcatc tatgttgtaa attttcaatt       420
     tttaaaagta cggaatattt gagatt                                           446
//
```

```
ID   BOK8A2     standard; RNA; PLN; 780 BP.
XX
AC   Z74921;
XX
NI   e1001207
XX
DT   29-JUL-1996 (Rel. 48, Created)
DT   29-JUL-1996 (Rel. 48, Last updated, Version 1)
XX
DE   B.oleracea mRNA (unknown)
XX
KW   .
XX
OS   Brassica oleracea
OC   Eukaryota; Plantae; Embryobionta; Magnoliophyta; Magnoliopsida;
OC   Dilleniidae; Capparales; Brassicaceae.
XX
RN   [1]
RP   1-780
RA   Booker J.P.;
RT   ;
RL   Thesis (1996), Biological Science, University of Durham
XX
RN   [2]
RP   1-780
RA   Croy R.D.;
RT   ;
RL   Submitted (26-JUL-1996) to the EMBL/GenBank/DDBJ databases.
RL   Ronald R.D. Croy, Biological Sciences, University of Durham, South
RL   Road, Durham, County Durham, DH1 3LE, United Kingdom
XX
FH   Key             Location/Qualifiers
FH
FT   source          1..780
FT                   /organism="Brassica oleracea"
FT                   /strain="alboglabra"
FT                   /clone="C10A"
FT                   /dev_stage="mature"
FT                   /tissue_type="stigma"
FT                   /cell_line="Plant line S29"
FT                   /clone_lib="lambda ExCell"
FT   mRNA            &lt;1..780
FT                   /gene="unknown"
FT                   /note="stigma specific transcript"
FT   polyA_site      780
FT                   /evidence=EXPERIMENTAL
XX
SQ   Sequence 780 BP; 242 A; 137 C; 180 G; 221 T; 0 other;
     gcggccccag gaatttttt tgtggtgaat gcagagttct tgttcagttt tggtgtagtt        60
     atggcactgt gcctctcaat gttattgtca ctcaggttct tcttaacctt tcatatatct       120
     tcaacttcca aaattttaag tctatgaagg tctatgacta atcattgttc atctattgga       180
     ttcacctgaa tgtgaacaga tggggtctcg gtgcaagaag ctgtgatagc agagaatgta       240
     agagactcac ttcacagctg gtgcaagaga gtgaaagaga ggtcaaagca cgcgagatca       300
     gtgtgttcac tagacacagc aacaatagac gagagagacg agatgacagt gggaacattg       360
     tcaaggagct catccatgac ctcgctgaat cagatcacca taaactccac agaccaagca       420
     gagtccatat tcggagcagc agcttcatcg agcagtcctc aagatgaata cactttcgag       480
     agtcgaagga attatctgtc tgaaacattc aataacatcg gttcgatgca acctttaaac       540
     gatgagattg agattgaggt tgaaagtgaa gaagataatg aaaggcgagg gagtgagagt       600
     gaagagacta atggtgaagg tggagaaaca ctttttgatt tgtttaggag gacttgatct       660
     tgtttcttct tcatcttcta cttccttatt gaagcttggt aaagttgtaa ttacatgaca       720
     ataacatcca taaaagcaaa aactactgta acaatgtaac agagtaaatc ttttgttggt       780
//
```

```
ID   BOC10A3UT  standard; RNA; PLN; 508 BP.
XX
AC   Z74963;
XX
NI   e1001210
XX
DT   29-JUL-1996 (Rel. 48, Created)
DT   29-JUL-1996 (Rel. 48, Last updated, Version 1)
XX
DE   B.oleracea mRNA (unknown)
XX
KW   .
XX
OS   Brassica oleracea
OC   Eukaryota; Plantae; Embryobionta; Magnoliophyta; Magnoliopsida;
OC   Dilleniidae; Capparales; Brassicaceae.
XX
RN   [1]
RP   1-508
RA   Booker J.P.;
RT   ;
RL   Thesis (1996), Biological Science, University of Durham
XX
RN   [2]
RP   1-508
RA   Croy R.D.;
RT   ;
RL   Submitted (26-JUL-1996) to the EMBL/GenBank/DDBJ databases.
RL   Ronald R.D. Croy, Biological Sciences, University of Durham, South
RL   Road, Durham, County Durham, DH1 3LE, United Kingdom
XX
FH   Key             Location/Qualifiers
FH
FT   source          1..508
FT                   /organism="Brassica oleracea"
FT                   /strain="alboglabra"
FT                   /clone="C10A"
FT                   /dev_stage="mature"
FT                   /tissue_type="stigma"
FT                   /cell_line="plant line S29"
FT                   /clone_lib="lambda ExCell"
FT   mRNA            &lt;1..508
FT                   /gene="unknown"
FT   polyA_site      508
FT                   /evidence=EXPERIMENTAL
XX
SQ   Sequence 508 BP; 143 A; 87 C; 116 G; 162 T; 0 other;
     caattccaat ggaatttatc tttgaacggg aatctataga tatgaatgga actgaaaagt        60
     aacaacacta tcagcatcac tcgggtttgt ctggactttt ggtccatcgt cgtaagtcgt       120
     agtattgcgt cgtaagtcgt agtattgcac ataaaagaga cacttctcaa agtaggaatc       180
     atctcagggt ttatgagttt gagacaagaa agactccgta agtactgcat ttgaaggctt       240
     ctagcggttc accgtaccta cacccccttt tggcttgtgc tgtgttacag acaagtgggt       300
     gagaagtgag aacttactcc cgagtccaag tataatgaaa taaagtcgag aagtatatag       360
     ggagcaggaa acactgtatc gggttttctt tgggtttctt tggaggtctc ggaacgtgat       420
     tttcaggatt tcaccataat agtttgatta ctgcgaagta atttttttac tgcttaatat       480
     taatctcgtt ggcaagtttg gttatttc                                          508
//
```

```
ID   BOD5A2     standard; RNA; PLN; 1216 BP.
XX
AC   Z74841;
XX
NI   e1001204
XX
DT   29-JUL-1996 (Rel. 48, Created)
DT   29-JUL-1996 (Rel. 48, Last updated, Version 1)
XX
DE   B.oleracea mRNA for hypothetical protein
XX
KW   .
XX
OS   Brassica oleracea
OC   Eukaryota; Plantae; Embryobionta; Magnoliophyta; Magnoliopsida;
OC   Dilleniidae; Capparales; Brassicaceae.
XX
RN   [1]
RP   1-1216
RA   Booker J.P.;
RT   ;
RL   Thesis (1996), Biological Science, University of Durham
XX
RN   [2]
RP   1-1216
RA   Croy R.D.;
RT   ;
RL   Submitted (26-JUL-1996) to the EMBL/GenBank/DDBJ databases.
RL   Ronald R.D. Croy, Biological Sciences, University of Durham, South
RL   Road, Durham, County Durham, DH1 3LE, United Kingdom
XX
FH   Key             Location/Qualifiers
FH
FT   source          1..1216
FT                   /organism="Brassica oleracea"
FT                   /strain="alboglabra"
FT                   /clone="C10A"
FT                   /dev_stage="mature"
FT                   /tissue_type="stigma"
FT                   /cell_line="Plant line S29"
FT                   /clone_lib="lambda ExCell"
FT   CDS             &lt;1..852
FT                   /product="hypothetical protein"
FT                   /codon_start=1
FT                   /db_xref="PID:e256042"
FT                   /translation="RGREKRSLKKVLYLSFIDYGTALSPLLNRLHFNHNALPLSPAPLP
FT                   LLFHRLLPSRSAVDSNVSYYEGTVWHDRLRPVRHSFRYTVRYALFDLDKSLETPPDHLS
FT                   ADEARLLARTTGPIFLLTIPPSVGYEQNPLSLYYCYNLEGSSKRLSKCIAQVTNTHGGE
FT                   RVTFVFDPESDLVAKSLQVSPFMDMLGNWKIRANEPGDELSVSIESQHPHHGNYFSAKL
FT                   KAKRIDQTRVSDPAVFFWLMPHKVAIWIYWHALKLWWRNVPFHPTPQILQTHSYREDSA
FT                   KT"
FT   polyA_site      1216
FT                   /evidence=EXPERIMENTAL
XX
SQ   Sequence 1216 BP; 312 A; 296 C; 260 G; 348 T; 0 other;
     agagggagag agaagagaag cttgaagaag gttctctatc tctcgttcat agattatgga        60
     actgctctat ctcctctgct caatcgtcta cacttcaacc acaacgctct tcctctctct       120
     cctgctcccc ttccgctcct tttccaccgt ctcctacctt cacgctccgc cgtcgattcc       180
     aacgtctctt actacgaagg caccgtctgg cacgaccgtc tccgtcccgt ccgtcactcg       240
     ttccgttaca ccgtccgtta cgcgctcttc gacctcgata aatcccttga aactccgccg       300
     gatcatctct ccgctgacga agctcgcctt ctcgctcgta ccaccggtcc gattttctta       360
     ttgacaatac ctccaagcgt tggatatgag cagaacccgt tgagtttgta ttattgctac       420
     aacttggaag gatcaagcaa acgcttaagt aaatgcattg cacaggttac aaatacgcat       480
     gggggggaac gagtgacatt tgttttcgac cctgaatctg acttggttgc taaatcatta       540
     caagtcagtc ctttcatgga tatgcttggg aattggaaga tcagagcaaa cgaacctgga       600
     gatgagttat ctgtatcaat tgaatcacag catcctcatc acggtaacta cttctctgcc       660
     aagttgaagg ctaaaagaat agaccaaaca cgggtttcag atcccgctgt cttcttctgg       720
     ttgatgcctc ataaggttgc tatatggatc tattggcatg cacttaagct ctggtggagg       780
     aatgtaccct ttcatccaac accccagata ctacaaaccc attcatacag agaggactca       840
     gcgaaaacgt gatcaaaaac ttcgatgtgt tggtttaaga tggatcaaac tctggtgaaa       900
     ccattagctt tgatggctgt agtagtggtt ttggaggatg gtcgctttgc gtggcgagat       960
     gctaattggc cttggtcatg aatatgcaca cacttgagat agcagatcgg catgactcat      1020
     gccgcatgag ttaagaatct acaatcattc atcgagaggt aggagtggtt tacttgttta      1080
     gattttggct taataagcaa agatgaacgt tgtttgatga taagccgacg atgtaacttg      1140
     acttcttgca agttggagaa caatcctgca ttttcaagta ctataccata tgaaatgtta      1200
     gcgtaccatt tgctct                                                      1216
//
```

Appendix D Database accessions for other sequences quoted in this project.

(a) Nucleic acid sequences.

Quoted name	Accession number (EMBL)	Quoted name	Accession number (EMBL)
At0564	T43056	At9541	T21954
At0683	T42068	Atts0356	Z17782
At13934	T13934	Atts4051	Z37586
At21018	N38210	Atts4074	Z37609
At283	T04283	Atts4406	Z47384
At31012	H37316	Atts5265	F14353
At3578	T76357	Atts5748	F15500
At36616	N37366	Bcestc	L35775
At37014	R90370	Bcestq	L35789
At41020	N96410	Hihi0323	L44968
At44114	R30441	Osr27361a	D24899
At4953	T22495	Oss1454	D39835
At5205	T42520	SLG29 (BoS29-2)	X16123
At52221	N96522	SLR1(BoS29-1)	X16122
At6225	T42622	Vp06989	U06949
At70414	R90704	Zm293	T15293
At7042	T21704	Zm638	T18638
At82813	H36828		

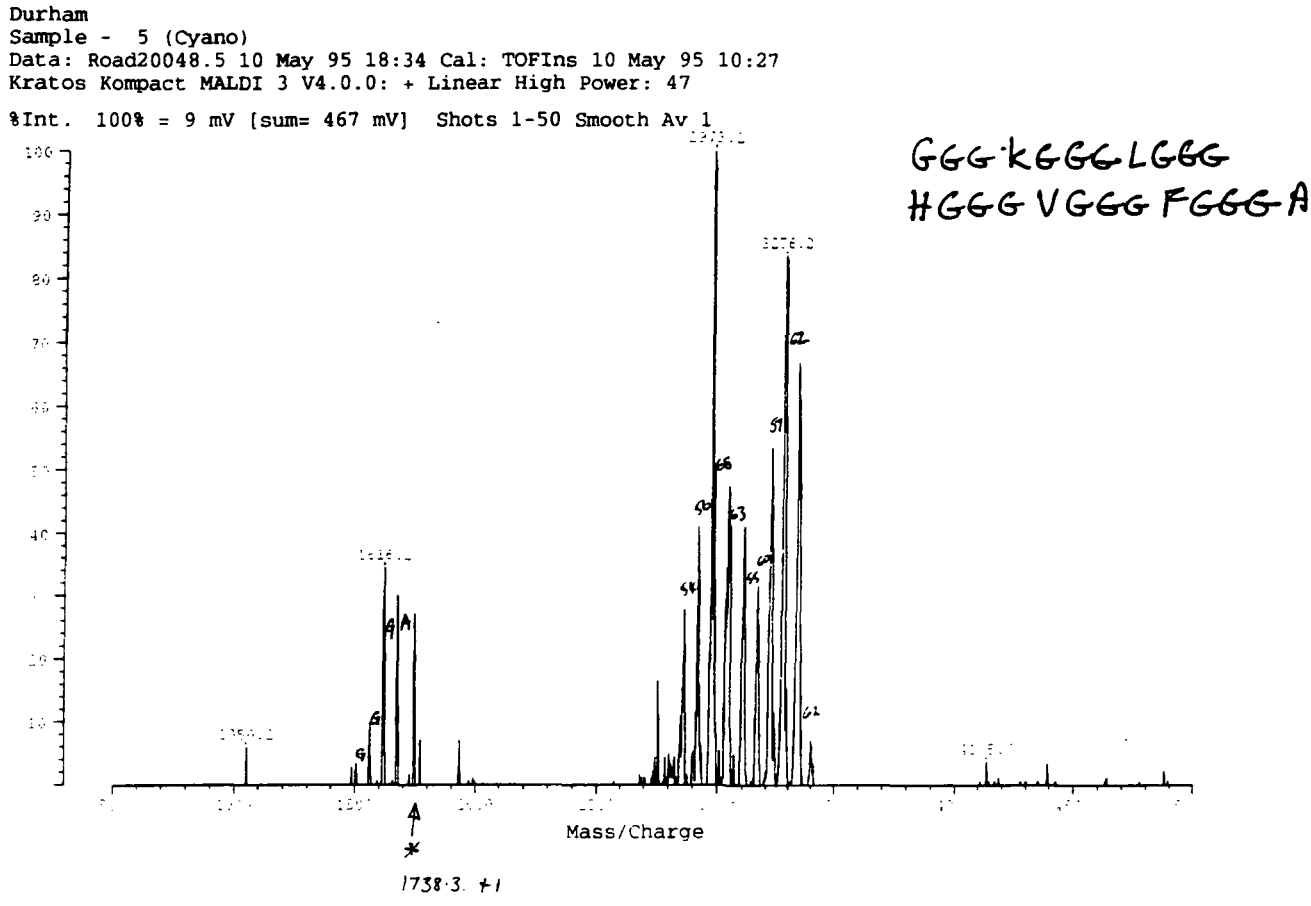
(b) Protein sequences (*=EMBL accession number from which the protein sequence was derived)

Quoted name	Accession number (SWISSPROT)*	Quoted name	Accession number (SWISSPROT)*
A.thal1 (MLP)	X91960	MLP	M37493
A.thal3 (MLP)	X91961	MLP1	L06467
A.thaliana atGRP1	S47405*	*N.sylvestris* GRP	P27484
A.thaliana atGRP2	S47408*	*N.tabacum* GRP	P23137
A.thaliana atGRP3	S47409*	*O.sativa* OsGRP1	P25074
A.thaliana atGRP4	S47413*	*O.sativa* OsGRP2	S29834
A.thaliana atGRP5	S47414*	*P.hybrida* GRP	P09789
A.thaliana genclone	P27483	*P.vulgaris* GRP1.0	P10495
H.vulgare GRP	P17816	*P.vulgaris* GRP1.8	P10496
HI0323	P44638	YRN1	P46235
L.esculentum GRP	Q01157	*Z.mays* GRP	P10979

Appendix E. Further work on the protein product of 15H11.

In order to obtain further information on the localisation of the GRP, predicted to be coded for by 15H11, an immunolocalisation approach was decided upon. In order to raise antibodies for this work it was decided to synthesise a peptide homologous to 15H11. The sequence chosen for the peptide was GGGKGGGLGGGHGGGVGGGFGGGA which corresponds to amino acids 99-123 of the predicted 15H11 protein. This stetch of sequence was selected at random from that available as antigenicity analysis of the predicted 15H11 protein suggested that the level of antigenicity was stable throughout the protein (data not shown).

The peptide was synthesised using a Shimadzu PSSM-8 Peptide Synthesizer that was on loan in the Department of Biological Sciences at the University of Durham. Analysis of the peptide product using a Kratos MALDI-TOF mass-charge analyser revealed that the reaction product contained a large number of peptides with higher molecular weights than the desired peptide (fig below). Despite this it was decided to raise antibodies to this mixture by maleimide coupling to a maleimide carrier and injection into a suitable animal. This work is currently in progress.



MALDI-TOF mass/charge analysis of synthesised peptide. The predicted peptide product is indicated by the arrow and legend 1738.3. 41.

7
References.

Alberts, A., Bray, D., Lewis,J., Raff, M., Roberts, K. and Watson, J.D. (1989) Molecular Biology of the Cell., 2nd edition. (Garland publishing, New York).

Anderson, M.A., Cornish, E.C., Mau, S-L., Williams, E.G., Hoggart, R., Atkinson, A., Bonig, I., Grego, B., Simpson, R., Roche, P.J., Haley, J.D., Penschow, J.D., Niall, H.D., Treager, G.W., Coghlan, J.P., Crawford, R.J. and Clarke, A.E. (1986) Cloning of cDNA for a stylar glycoprotein associated with expression of self-incompatibility in *Nicotiana alata*. Nature 321, 38-44.

Anderson, M.A., McFadden, G.I., Bernatzky, R., Atkinson, A., Orpin, T., Dedman, H., Tregear, G., Fernley, R. and Clarke, A.E. (1989) Sequence variability of three alleles of the self-incompatibility gene of *Nicotiana alata*. Plant Cell 1, 483-491.

Argos,P., Narayana, S.V.L., Nielsen, N.C. (1985) Structural similarity between legumin and vicillin storage proteins from legumes. EMBO Journal 1, 1111-1117

Arumuganathan, K. and Earle, E.D. (1991). Nuclear DNA content of some important plant species. Plant Molecular Biology Reporter 9, 208-218.

Atkinson, A.H., Heath, R.L., Simpson, R.J., Clarke, A.E. and Anderson, M.A. (1993). Proteinase-inhibitors in *Nicotiana alata* stigmas are derived from a precursor protein which is processed into 5 homologous inhibitors. Plant Cell 5, 203-213.

Bacic, A., Gell.. A.C. and Clarke, A.E. (1988). Arabinogalactan proteins from the stigma of *N.alata*. Phytochemistry 27 679-684.

Baldwin, T.C., Coen, E.S. and Dickinson, H.G. (1992). The *ptl1* gene expressed in the transmitting tissue of Antirrhinum encodes an extensin-like protein. Plant Journal 5, 733-739.

Bateman, A.J. (1955) Self-incompatibility systems in angiosperms. III. Cruciferae. Heredity 9, 52-68.

Bäumlein, H., Braun, H., Kakhovskaya, I.A. and Shutov, A.D. (1995) Seed storage proteins of spermatophytes share a common ancestor with desiccation proteins of fungi. Journal of Molecular Evolution 41, 1070-1075.

Bayer, E.A., and Wilchek, M. (1990). Protein biotinylation. Methods in Enzymology 184, 138-162.

Bell, J. and Hicks, G. (1976) Transmitting tissue in the pistil of tobacco: Light and electron microscopic observations. Planta 131, 187-200.

Berger, S.L. and Birkenmeier, C.S. (1979) Inhibition of intractable nucleases with ribonucleoside-vanadyl complexesL isolation of meesenger ribonucleic acid from resting lymphocytes. Biochemistry 18, 5143-

Berridge, M.J. (1988) Inositol lipids and calcium signalling. Proc. Roy. Soc. Lond. ser.B 234, 359-378.

Bown, D.P., Bolwell, G.P. and Gatehouse, J.A. (1993) Characterisation of potato (*Solanum tuberosum* L.) extensins: a novel extensin-like cDNA from dormant tubers Gene 134, 229-233.

Boyes, D.C. and Nasrallah, J.B. (1993) Physical linkage of the SLG and SRK genes at the self-incompatibility locus of *Brassica oleracea* Mol. Gen. Genet. 222, 241-248.

Boyes, D.C. and Nasrallah, J.B. (1995). An anther-specific gene encoded by an S-locus haplotype of Brassica produces complementary and differentially regulated transcripts. Plant Cell 7, 1283-1294.

Boyes, D.C., Chen, C.H., Tantikanjana, T., Esch, J.J., Nasrallah, M.E. and Nasrallah, J.B. (1991) Isolation of a second S-locus related cDNA from *Brassica oleracea*: Genetic relationships between the S locus and two related loci. Genetics 127, 221-228.

Brown, J.W.S. (1986). A catalogue of splice junctions and putative branch point sequences from plant introns. Nucleic Acids Research 14, 9549-9559

Budelier, K.A., Smith, A.G. and Gasser, C.S. (1990). Regulation of a stylar-transmitting tissue-specific gene in wild-type and transgenic tomato and tobacco. Molecular and General Genetics 224, 183-192.

Campbell, W.H. and Gowri, G. (1989) Codon usage in higher plants, green algae and cyanobacteria. Plant Physiology 92, 1-11.

Cappadocia, M., Heizmann, P. and Dumas, C. (1993). Tissue printing and its application in self-incompatibility studies. Plant Molecular Biology 23, 1079-1085.

Carter, A.L. and McNeilly, T. (1975). Effect of increased humidity on pollen growth and seed set following self-pollination in Brussel Sprout (*Brassica oleracea* var *gemifera*). Euphytica 24, 805-813.

Casey, J. and Davison, N. (1977). Rates of formation and thermal stability of RNA:DNA and DNA:DNA duplexes at high concentrations of formamide. Nucleic Acids Research 4, 1539-1557.

Cassab, G.I. and Varner, J.E. (1988). Cell wall proteins. Annual Review of Plant Physiology and Plant Molecular Biology 39, 321-353.

Chen, C-G., Cornish, E.C. and Clarke, A.E. (1992) Specific expression of an extensin-like gene in the style of *Nicotiana alata*. Plant Cell 4, 1053-1062.

Chen, C-G., Mau, S-L. and Clarke, A.E. (1993) Nucleotide sequence and style specific expression of a novel proline-rich protein gene from *Nicotiana alata*. Plant Molecular Biology 21, 391-395.

Chester, K.S. (1933) The problem of acquired physiological immunity in plants. Quarterly review of biology 8, 275-324.

Cheung, A.C.Y. and Cohen, S. (1975) Construction and characterisation of amplifiable, multicopy DNA cloning vectors derived from P15A cryptic miniplasmid. Journal of Bacteriology 134, 1141-1156.

Cheung, A.Y., May, B., Katawa, E.E. and Wu, H-M. (1993) Characterisation of cDNAs for stylar transmitting tissue-specific proline-rich proteins. Plant Journal 3, 151-160.

Cheung, A.Y., Wang, H. and Wu, H-M. (1995) A floral transmitting tissue-specific glycoprotein attracts pollen tubes and stimulates their growth. Cell 82, 383-393.

Churngchow, N., Suntaro, A. and Wititsuwannnakul, R. (1995). ß-1,3-glucanase isozymes from the latex of *Hevea Brasiliensis*. Phytochemistry 39, 505-509.

Clackson, T., Guusow,D. And Jones, P.T. (1991) In PCR A Practical Approach. Ed Mcpherson, M.J., Quirke, P. And Taylor, G.R. Irl Press, p210

Clark, K.R. and Sims, T.L. (1994) The S-Ribonuclease gene of *Petunia hybrida* is expressed in nonstylar tissue, including immature anthers. Plant Physiology 106, 25-36

Clark, K.R., Okuley, J.J., Collins, P.D. and Sims, T.L. (1990). Sequence variability and developmental expression of S-alleles in self-incompatible and pseudo-self-incompatible petunia. Plant Cell 2, 815-826.

Clarke, A., Gleeson, P., Harrison, S. and Knox, R.B. (1979) Pollen-stigma interactions identification and characterisation of surface components with recognition potential. Proc. Natl. Acad. Sci. (USA) 76, 3358-3362.

Coen, E.S. and Carpenter, R. (1993). The metamorphosis of flowers. Plant Cell 5, 1175-1181.

Condit, C.M. and Meagher, R.B.(1987). Expression of a gene encoding a glycine-rich protein in petunia. Molecular Cell Biology 12, 4273-4279.

Condit, C.M. and Meaghler, R.B. (1989). A gene encoding a noval glycine-rich protein of petunia. Nature 323, 178-181.

Constabel, C.P. and Brisson, N. (1995) Stigma- and vascular-specific expression of the PR-10a gene of potato: a novel pattern of expression of a pathogenesis-related gene. Molecular Plant-Microbe Interactions 8, 104-113.

Cooke, R., Raynal, M., Laudie, M. Grellet, F., Delseny, M., Morris, P.C., Guerrier, D., Giraudat, J., Quigley, F., Clabault, G., Li, Y.F., Mache, R, Krivitzky, M., Gy, I.I.J., Kreis, M., Lecharny, A., Parmentier, Y., Marbach, J., Fleck, J., Clement, B., Philipps, G., Herve, C., Bardet, C., Tremousaygue, D., Lescure, B., Lacomme, C., Roby, D., Jourjon, M.F., Chabrier, P., Charpenteau, J.L. Desprez, T., Amselem, J., Chiapello, H. and Hofte, H. (1996). Further progress towards a catalog of all *Arabidopsis* genes: Analysis of a set of 5000 non-redundant ESTs. Plant Journal 9, 101-124.

Corbin, D.R., Sauer, N. and Lamb, C.J. (1987) Differential regulaion of a hydroxyproline-rich glycoprotein gene family in wounded and infected plants. Molecular and Cellular Biology. 7, 4337-4344.

Cresti, M., Keijzer, C.J., Tiezzi, A., Ciampolini, F. and Focardi, S. (1986) Stigma of Nicotiana: Ultrastructural and biochemical studies. American Journal of Botany 73, 1713-1722.

Cresti, M., van Went, J.L., Pacini, E. and Willemse, M.T.M. (1976). Ultrastructure of transmitting tissue of *Lycopersicon peruvianum* style: Development and histochemistry. Planta 132, 305-312.

De Jong, A.J., Cordewener, J., Schiavo, F.L., Terzi, M., Vanderkerckhove, J., van Kammen, A. and De Vries, S.C. (1992) A carrot somatic embryo mutant is rescued by chitinase. Plant Cell 4, 425-433.

De Jong, A.J., Heidstra, R, Spaink, H.P., Hartog, M.V., Meijer, E.A., Hendriks, T, Loschiavo, F, Terzi, M, Bisseling, T, Vankammen, A and De Vries, S.C. (1993) Rhizobium lipooligosaccharides rescue a carrot somatic embryo mutant. Plant Cell 5, 615-620.

De Oliviera, D.E., Seurinck, J., Inzé, D., van Montagu, M. and Botterman, J. (1990). Differential expression of five *Arabidopsis* genes encoding glycine rich proteins. Plant Cell 2, 427-436.

Dean, C., Tamaki, S., Dunsmuir, P., Favreau,M. Katayama, C., Dooner, H. and Bedbrook, J. (1986) mRNA transcripts of several higher plant genes are polyadenylated at multiple sites *in vivo*. Nucleic Acids Reseach 14, 2229-2240.

Del Campillo E. and Lewis, L.N. (1992) Occurrence of 9,5 cellulase and other hydrolases in flower reproductive organs undergoing major cell wall disruption. Plant Physiology 99, 1015-1020.

Dellaporta, S.L., and Calderon-Urrea, A. (1993) Sex determination in flowering plants. Plant Cell 5, 1241-1251.

Delorme, V., Giranton, J.L., Hatzfeld, Y., Friry, A., Heizmann, P., Ariza, M.J., Dumas, C., Gaude, T. and Cock, J.A. (1995) Characterization of the S-locus genes, SLG and SRK, of the *Brassica* S3 haplotype - identification of a membrane-localized protein encoded by the S-locus receptor kinase gene. Plant Journal 7, 429-440.

DeNettancourt, D. (1977) Incompatibility in angiosperms. In Monographs on Theoretical and Applied Genetics. R. Frankel, G.A.E. Gall and H.F. Linskens eds (Springer-Verlag Berlin) pp 28-57.

DeNettancourt, D., Devreux, M., Bozzini, A., Cresti, M., Pacini, E. and Sarfatti, G. (1973). Ultrastructural aspects of the self-incompatibility in *Lycopersicon peruvianum* mill. Journal of Cell Science 12, 403-419.

Desprez, T., Amselem, J., Chiapello, H., Caboche, M. and Hofte, H. (1994) The *Arabidopsis thaliana* transcribed genome: the GDR cDNA program. Unpublished.

Devereux, J., Haeber, P. and Smithies, O. (1984) A complete set of sequence analysis programs for the VAX. Nucleic Acids Research 12, 387-395.

Dickinson, H.G. and Lewis, D. (1973a). Cytochemical and ultrastructural differences between intraspecific compatible and incompatible pollinations in *Raphanus*. Proc. R. Soc. London series B 183, 21-38.

Dickinson, H.G. and Lewis, D. (1973b). The formation of the 'tryphine' coating the pollen grain of *Raphanus* and its properties relating to self-incompatibility. Proc. R. Soc. London B. 184, 149-165.

Dodds, P.N., Bonig, I Du, H, Rodin, J., Anderson, M.A., Newbigin, E. and Clarke, A.E. (1993). S-RNase gene of *Nicotiana alata* is expressed in developing pollen. Plant Cell 5, 1771-1782.

Dodds, P.N., Clarke, A.E., Newbigin, E. (1996). A molecular perspective on pollination in flowering plants. Cell 85, 141-144.

Doughty, J., Henderson, F., McCubbin, A. and Dickinson, H. (1993) Interaction between a coating-borne peptide of the *Brassica* pollen grain and stigmatic S(self-incompatibility)-locus-specific glycoproteins. Proc. Natl. Acad. Sci. (USA) 90, 467-471.

Du, H., Simpson, R.J., Clarke, A.E. and Bacic, A.E. (1996) Molecular characterization of a stigma-specific gene encoding an arabinogalactan protein (AGP) from *Nicotiana alata.* Plant Journal 9, 313-323.

Du, H., Simpson, R.J., Moritz, R.L., Clarke, A.E. and Bacic, A. (1994) Isolation of the protein backbone of an arabinogalactan protein from the styles of *Nicotiana alata* and characterization of a corresponding cDNA. Plant Cell 6, 1643-1653.

Dwyer, K.G., Balent, M.A., Nasrallah, J.B. and Nasrallah, M.E. (1991) DNA-sequences of self-incompatibility genes from *Brassica campestris* and *Brassica oleracea* : Polymorphism predating speciation. Plant Molecular Biology 16, 481-486.

Dwyer, K.G., Kandasamy, M.K. Mahosky, D.I., Acciai, J., Kudish, B.I., Miller, J.E., Nasrallah, M.E. and Nasrallah, J.B. (1994). A superfamily of S-locus-related sequences in *Arabidopsis*:: diverse structures and expression patterns. Plant Cell 6, 1829-1843.

Dwyer, K.G., Lalonde, B.A., Nasrallah, J.B. and Nasrallah, M.E. (1992) Structure and expression of Ats1, an *Arabidopsis thaliana* gene homologous to the S-locus related genes of *Brassica.* Molecular & General Genetics 231, 442-448.

Dzelzkalns, V.A., Thorsness, M.K., Dwyer, K.G., Baxter, J.S., Balent, M.A., Nasrallah, M.E. and Nasrallah, J.B. (1993) Distinct cis-acting elements direct pistil-specific and pollen-specific activity of the *Brassica* S locus glycoprotein gene promoter. Plant Cell 8, 855-863.

Elleman, C.J. and Dickinson, H.G. (1990) The role of the exine coating in pollen-stigma interactions in *Brassica oleracea* L. New Phytologist 114, 511-518.

Elleman, C.J., Franklin-Tong, V. and Dickinson, H.G. (1992). Pollination in species with dry stigmas: the nature of the early stigmatic response and the pathway taken by the pollen tubes. New Phytologist 121, 413-424.

Elleman, C.J., Willson, C., Sarker, R.H. and Dickinson, H.G. (1988). Interaction between the pollen tube and stigmatic cell wall following pollination in *Brassica oleracea.* New Phytologist 109, 111-117.

Fang R-X., Peng Z., Gau D-M., Mang K-Q. and Chou K-H. (1991). cDNA sequence of a virus-inducible glycine-rich protein gene from rice. Plant Molecular Biology 17, 1255-1257.

Fantl, W.J., Johnson, D.E. and Williams, L.T. (1993). Signalling by receptor tyrosine kinases. Annual Review of Biochemistry 62, 453-481.

Ferberg, A.P. and Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Analytical Biochemistry 132, 6-13.

Ferrari, T.E., Bruns, D. and Wallace, D.H. (1981) Isolation of a plant glycoprotein involved with control of intercellular recognition. Plant Physiology 67, 270-277.

Fleischmann R.D., Adams M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A., Merrick, J.M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C.A., Gocayne, J.D., Scott, J.D., Shirley, R., Liu, L.I., Glodek, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E., Cotton, M.D., Utterback, T.R., Hanna, M.C., Nguyen, D.T., Saudek, D.M., Brandon, R.C., Fine, L.D., Fritchman, J.L., Fuhrmann, J.L., Geoghagen, N.S.M., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M., Smith, H.O. and Venter, J.C. (1995) Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science 269, 496-512.

Foote, H.C.C., Ride, J.P., Franklin-Tong, V.E., Walker, E.A., Lawrence, M.J. and Franklin, F.C.H. (1994) Cloning and expression of a distinctive class of self-incompatibility (S) gene from *Papaver rhoeas*. Proc. Natl. Acad. Sci. (USA) 91, 2265-2269.

Franklin-Tong, V.E. and Franklin, F.C.H. (1992). Gametophytic self-incompatibility in *Papaver rhoeas* L. Sexual Plant Reproduction 5, 1-7.

Franklin-Tong, V.E., Atwal, K.K., Howell, E.C., Lawrence, M.J. and Franklin, F.C.H. (1991) Self-incompatibility in *Papaver rhoeas*. There is no evidence for the involvement of stigmatic ribonuclease activity. Plant Cell and Environment 14, 423-429.

Franklin-Tong, V.E., Lawrence, M.J. and Franklin, F.C.H. (1988) An in vitro bioassay for the stigmatic product of the self-incompatibility gene in *Papaver rhoeas*. New Phytologist 110, 109-118.

Franklin-Tong, V.E., Lawrence, M.J. and Franklin, F.C.H. (1990) Self-incompatibility in *Papaver rhoeas*: inhibition of incompatible pollen tube growth is dependent on pollen gene expression. New Phytologist 116, 319-324.

Franklin-Tong, V.E., Ride, J.P. and Franklin, F.C.H. (1995). Recombinant stigmatic self-incompatibility (S-) protein elicits a Ca^{2+} transient in pollen of *Papaver rhoeas*. Plant Journal 8, 299-307.

Franklin-Tong, V.E., Ride, J.P., Read, N.D., Trewavas, A.J. and Franklin, F.C.H. (1993). The self-incompatibility response in *Papaver rhoeas* is mediated by cytosolic-free calcium. Plant Journal 4, 163-177.

Franklin-Tong, V.E., Ruuth, E., Marmey, P., Lawrence, M.J. and Franklin, F.C.H. (1989) Characterisation of a stigmatic component from *Papaver rhoeas* L. which exhibits the specific activity of a self-incompatibility (S-)gene product. New Phytologist 112, 307-315.

Fry, S.C. (1986) Cross-linking of matrix polymers in the growing cell walls of angiosperms. Annual Review of Plant Physiology and Plant Molecular Biology 37, 165-186.

Fry, J., Barnason, A. and Horsch, R.B. (1987) Transformation of *Brassica napus* with *Agrobacterium tumefaciens* based vectors. Plant Cell Reporter 6, 6-13.

Fuchs, R. (1984). Predicting protein function: a versatile tool for the Apple Macintosh. Comput. Appl. Biosci. 10, 171-178.

Gane, A.M., Weinhandl, A., Bacic, A. and Harris, P.J. (1994) Histochemistry and Composition of the cell walls of styles of *Nicotina alata* Link et Otto. Planta 195, 217-225.

Garcia-Olmedo, F., Rodriguez-Palenzuela, P., Hernandez-Lucas, C., Ponz, F., Marana, C., Carmona, M.J., Lopez-Fando, J., Fernandez, J.A. and Carbenero, P. (1989) The thionins: a protein family that includes purothionins, viscotoxins and crambins. Oxford Surveys of Plant Molecular Biology 6, 31-60/

Garnier, J., Osguthorpe, D.J. and Robson, B. (1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. Journal of molecular Biology 120, 97-120.

Gasser, C.S. (1991) Molecular studies on the differentiation of floral organs. Annual Review Of Plant Physiology And Plant Molecular Biology 42, 621-649.

Gasser, C.S. and Robinson-Beers, K. (1993). Pistil development. Plant Cell 5, 1231-1239.

Gaude, T., Friry, A., Heizmann, P., Mariac, C., Rougier, M., Fobis, I and Dumas, C. (1993). Expression of a self-incompatibility gene in a self-compatible line of *Brassica oleracea*. Plant Cell 5, 75-86.

Gell.. A.C., Bacic, A., and Clarke, A.E. (1986) Arabinogalactan proteins of the female sexual tissue of *Nicotiana alata*. Changes during floral development and pollination Plant Physiology 82, 885-889.

Gibbs, P.E.M., Strongin, K.B. and McPherson, A. (1990) Evolution of legume seed storage proteins-a domain common to legumins and vicillins is duplicated in vicillins. Molecular Biology and Evolution 6, 614-623.

Gleeson, P.A. and Clarke, A.E. (1980) Arabinogalactans of sexual and somatic tissue of *Gladiolus* and *Lilium*. Phytochemistry 19, 777-782.

Gleeson, P.A.. and Clarke, A.E. (1979) Structural studies on the major component of *Gladiolus* style mucilage-an arabinogalactan protein. Biochemical Journal 181, 607-621.

Goldman, M.H.D., Pezzotti, M., Seurinck, J. and Mariani, C. (1992) Developmental expression of tobacco pistil-specific genes encoding novel extensin-like proteins. Plant Cell 4, 1041-1051.

Gomez, J., Sanchez-Martinez, D., Stiefel, V., Rigau, J., Puigdomenach, P. and Pagès, M. (1988). A gene induced by the plant hormone abiscisic acid in response to water stress encodes a glycine-rich protein. Nature 334, 262-264.

Goodall, G.J. and Filipowicz, W. (1989). The AU-rich sequences present in the introns of plant nuclear pre-mRNAs are required foe splicing. Cell 58,473-483.

Goodall, G.J. and Filipowicz, W. (1990). The minimum functional length of pre-mRNA introns in monocots and dicots. Plant Molecular Biology 14, 727-733.

Goring, D.R., Banks, P., Beversdorf, W.D., Rothstein, S.J. (1992) Use of the polymerase chain-reaction to isolate an S locus glycoprotein cDNA introgressed from *Brassica campestris* into *B.napus* ssp *oleifera*. Molecular & General Genetics 234, 185-

Goring, D.R., Banks, P., Fallis, L., Baszczynski, C.L., Beversdorf, W.D. and Rothstein, S.J. (1992). Identification of an S-locus glycoprotein allele introgressed from *Brassica napus* ssp *rapifera* to *Brassica napus* ssp *oleifera* Plant Journal 6, 983-989.

Goring, D.R., Glavin, T.L., Schafer, U. and Rothstein, S.J. (1993). An S receptor kinase gene in self-compatible *Brassica napus* has a 1-bp deletion. Plant Cell 5, 531-539.

Goring, D.R., Thomson, L. and Rothstein, S.J. (1991) Transformation of a partial nopaline synthase gene into tobacco suppresses the expression of a resident wild-type gene. Proceedings Of The National Academy Of Sciences Of The United States Of America 88, 1770-1774.

Gray, J.E., Mcclure, B.A., Bonig, I., Anderson, M.A. and Clarke, A.E. (1991). Action of the style product of the self-incompatibility gene of *Nicotiana-alata* (S-RNase) on in vitro-grown pollen tubes. Plant Cell 3, 271-283.

Gu, Q., Kawata, E., Morse, M-J., Wu, H-M. and Cheung, A.Y. (1992) A flower-specific cDNA encoding a novel thionin in tobacco. Molecular & General Genetics 234, 89-96.

Guiltinan, M.J. and Xiping, N. (1996) cDNA encoding a wheat (*Triticum aestivum* cv Chinese spring) glycine-rich RNA-binding protein. Plant Molecular Biology 30, 1301-1306.

Hackett, R.M., Lawrence, M.J. and Franklin, F.C.H. (1992) A *Brassica* S-locus related gene promoter directs expression in both pollen and pistil of tobacco. Plant Journal 2, 613-617.

Hanahan, D. (1985). in DNA Cloning, a Practical Approach (Glover, D.M., ed). IRL Press.

Harikrishna, K., Jampates-Beale, R., Milligan, S.B. and Gasser, C.S. (1996) An endochitinase gene expressed at high levels in the stylar transmitting tissue of tomatoes. Plant Molecular Biology 30, 899-911.

Heizmann, P., Friry, A., Guilluy, C-M. and Dumas, C. (1991) Expression of SLG self-incompatibility genes in the anthers of *Brassica oleracea*: new members in the SLG gene family. Science 80, 193-199.

Herrero, M and Dickinson, H.G. (1979) Pollen-pistil incompatibility in *Petunia hybrida*: Changes in the pistil following compatible and incompatible intraspecific crosses. Journal of Cell Science 36, 1-18.

Heslop-Harrison, Y. (1981) Stigma characteristics and angiosperm taxonomy. Nordic Journal of Botany 1, 401-420.

Heslop-Harrison, J. (1975). Incompatibility and the pollen-stigma interaction. Annual Review of Plant Physiology 26, 403-425.

Heslop-Harrison, J., Heslop-Harrison, Y. and Barber, J. (1975) The stigma surface in incompatibility responses. Proc. Royal soc. London series B 188, 287-297.

Heslop-Harrison, Y. and Shivanna K.R. (1977)The receptive surface of the angiosperm stigma. Annals of Botany 41, 1233-1258.

Hill, J.P. and Lord, E.M. (1987) Dynamics of pollen tube growth in the wild radish *Raphanus raphinastrum* (Brassicaceae). II. Morphology, cytochemistry and ultrastructure of transmitting tissues and path of pollen tube growth. American Journal of Botany 74, 988-997.

Hiscock, S.J., Dewey, F.M., Coleman, J.O.D. and Dickinson, H.G. (1994). Identification and localization of an active cutinase in the pollen of *Brassica napus*. planta 193, 377-384.

Hiscock, S.J., Doughty, J., Willis, A.C. and Dickinson, H.G. (1995) A 7kDa pollen coating-borne peptide from *Brassica napus* interacts with S-locus glycoprotein and S-locus-related glycoprotein. Planta 96, 367-374.

Hodge, R., Paul, W., Draper, J. and Scott, R.L.(1992). Cold-plaque screening - a simple technique for the isolation of low abundance, differentially expressed transcripts from conventional cDNA libraries. Plant Journal 2, 257-260.

Hodgkin, T, Lyon, G.D. and Dickinson, H.G. (1988) Recognition in flowering plants: A comparison of the *Brassica* self-incompatibility system and plant pathogen interactions. New Phytologist 110, 557-569.

Hoggart, R,M. and Clarke, A.E. (1984) Arabinogalactans are common components of angiosperm styles. Phytochemistry 23, 1571-1573.

Howley, P.M., Israel, M.F., Law, M-F. and Martin, M.A. (1979). A rapid method for detecting and mapping homology between heterologous DNAs. Journal of Biological Chemistry 254, 4876-4883.

Huang, S. Lee, H.S., Karunanandaa, B., Kao, T.H. (1994). Ribonuclease-activity of *Petunia inflata* S proteins is essential for rejection of self-pollen. Plant Cell 6, 1021-1028.

Hunt, A.G. (1994) Messenger RNA 3' end formation in plants. Annual Review of Plant Physiology and Plant Molecular Biology 45, 47-60.

Hynes, R.O. and Landers, A.D. (1992). Contact and adhesive specificities in the associations,migrations and targetting of cells and axons. Cell 68, 303-322.

Jahnen, W., Batterham, M.P., Clarke, A.E., Moritz, R.L. and Simpson, R.J(1989) Identification, isolation, and N-terminal sequencing of style glycoproteins associated with self-incompatibility in *Nicotiana-alata*. Plant Cell 2, 493-499.

Jahnen, W., Lush,W.M. and, Clarke, A.E. (1989) Inhibition of in vitro pollen-tube growth by isolated S-glycoproteins of *Nicotiana-alata*. Plant Cell 5, 501-510.

Jauh, G.Y., Lord, E.M. (1996) Localization of pectins and arabinogalactan-proteins in lily (*Lilium longiflorum* L) pollen-tube and style, and their possible roles in pollination. Planta 199, 251-261.

Jermyn, M.A. and Yeo, W. (1975) A class of lectin present in the tissue of seed plants. Australian Journal of Plant Physiology 2, 501-503.

Joshi, C.P. (1987). An inspection of the domain between putative TATA box and translation start site in 79 plant genes. Nucleic Acids Research 15 6643-6653.

Jung, J. (1956) Sind narbe und griffel eintrittsporforten fur pilzin testtionen? Phytopathol Z. 27, 752-756

Kamalay, J.C. and Goldberg, R.B. (1980) Regulation of structural gene expression in tobacco. Cell 19, 935-946.

Kandasamy, M.K., Dwyer, K.G., Paolillo, D.J., Doney, R.C., Nasrallah, J.B. and Nasrallah, M.E. (1990). *Brassica* S-proteins accumulate in the intercellular matrix along the path of pollen tubes in transgenic tobacco pistils. Plant Cell 2, 39-46.

Kandasamy, M.K., Thorsness, M.K., Rundle, S.J., Goldberg, M.L., Nasrallah, J.B. and Nasrallah, M.E. (1993). Ablation of papillar cell-function in *Brassica* flowers results in the loss of stigma receptivity to pollination. Plant Cell 5, 263-275.

Kandasamy, M.K. and Kristen, U. (1987) Developmental aspects of ultrastructure, histochemistry and receptivity of the stigma of *Nicotiana sylvestris*. Annals of Botany 60, 427-437.

Kandasamy, M.K., Paolillo, D.J., Faraday, C.D., Nasrallah, J.B. and Nasrallah, M.E. (1989) The S-locus specific glycoproteins of *Brassica* accumulate in the cell wall of developing stigma papillae. Developmental Biology 134, 462-472.

Kandasamy, M.K., Nasrallah, J.B. and Nasrallah, M.E. (1994). Pollen pistil interactions and developmental regulation of pollen tube growth in *Arabidopsis*. Development 120, 3405-3418.

Kaufmann, H., Salamini, F. and Thompson, R.D. (1991) Sequence variability and gene structure at the self- incompatibility locus of *Solanum tuberosum*. Molecular & General Genetics 226, 457-466.

Kawata, Y., Sakiyama, F., Hayashi, F., and Kyogoku, Y. (1990) Identification of two identical histidine residues of ribonuclease T2 in *Aspergillus oryzae*. European Journal of Biochemistry. 176, 683-697.

Keller B., Sauer N. and Lamb C.J. (1988). Glycine-rich cell wall protins in bean: gene structure and association of the protein with the vascular system. EMBO Journal 7, 3625-3633.

Keller, B., Schmid, J. and Lamb, C.J. (1989a). Vascular expression of a bean cell wall glycine-rich protein-ß-glucuronidase gene fusion inb transgenic tobacco. EMBO Journal 8, 1309-1314.

Keller, B., Templeton, M.D. and Lamb, C.J. (1989b). Specific localisation of a plant cell wall glycine-rich protein in protoxylem cells of the vascular system. Proc. Natl. Acad. Sci. (USA) 86, 1529-1533.

Khandijan, E.W. (1987). Optimised hybridisation of DNA blotted and fixed to nitrocellulose and nylon membranes. Biotechnology 5, 165-167.

Klee, H., Horsch, R. and Rogers, S. (1987) Agrobacterium-mediated plant transformation and its further applications to plant biology. Annual Review Of Plant Physiology And Plant Molecular Biology 38, 467-486.

Knox, R.B. (1984) Pollen-pistil interactions. In Encyclopedia of Plant Physiology vol.17: Cellular interactions. H.F. Linskens and J. Heslop-Harrison eds. (Springer-Verlag, Berlin) pp 508-608.

Knox, R.B., Harrison, S., Smith, P. and Marchalonis, J.J. (1976) Cell recognition in plants. Determints of the stigma surface and their pollen interactions.Proc. Natl. Acad. Sci. (USA) 76, 2788-2792

Konar, R.N. and Linskens, H.F. (1966a). The morphology and anatomy of the stigma of *Petunia hybrida*. Planta 71, 372-387.

Konar, R.N. and Linskens, H.F. (1966b). Physiology and Biochemistry of the stigmatic fluid of *Petunia hybrida*. Planta 71, 356-371.

Kowyama, Y., Kunz, C., Lewis, I., Newbigin, E., Clarke, A.E. and Anderson, M.A. (1994). Self incompatibility in a *Lycopersicon peruvianum* varient (LA2157) is associated with a lack of style S-RNase activity. Theoretical and Applied Genetics 88, 859-864.

Krivitzky, M., Jean-Jacques, I., Kreis, M. and Lecharny, A. (1995) The *Arabidopsis thaliana* transcribed genome: the GDR cDNA program. Unpublished.

Kumar, V., and Trick, M. (1994). Expression of the S-locus receptor kinase multigene family in *Brassica oleracea*. Plant Journal 6, 807-813.

Kyte, J. and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. Journal of Molecular Biology. 157, 105-132.

Laemelli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680-685

Lalonde, B.A., Nasrallah, M.E., Dwyer, K.G., Chen, C-H., Barlow, B. and Nasrallah, J.B. (1989) A highly conserved *Brassica* gene with homology to the S-locus-specific glycoprotein structural gene. Plant Cell 1, 249-258.

Lamport, D.T.A. (1969). The isolation and partial characterisation of hydoxyproline glycopeptides obtained by enzymic of primary cell walls. Biochemistry 8, 1155-1163.

Lawrence, M.C., Izard, T., Beuchat, M., Blagrove, R.J. and Colman, P.M. (1994) Structure of phaseolin at 2.2Å resolution. Journal of Molecular Biology 238, 748-776.

Lawrence, M.J. (1975) The genetics of self-incompatibility in *Papaver rhoeas*. Proc. Roy. Soc. London 188 275-28.

Lawrence, M.J., Afzal, M. and Kenrick, J. (1976) The genetical control of self-incompatibility in *Papaver rhoeas*. Heredity 40, 239-253.

Lee, H.S., Huang, S.S and Kao, T.H. (1994) S-proteins control rejection of incompatible pollen in *Petunia inflata*. Nature, 367, 560-563.

Lei, M. and Wu R. (1991). A novel glycine-rich cell wall protein gene in rice. Plant Molecular Biology 16, 187-198.

Leung, D.W. (1992) Involvement of plant chitinase in sexual reproduction of higher plants. Phytochemistry 6, 1899-1900.

Li, Q., Chen, F., Linskens, H.F. and Cresti, M. (1994) Distribution of unesterified and esterified pectinsin cell-walls of pollen tubes of flowering plants. Sexual Plant Reproduction 7, 145-152.

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

Lim, C.O., Hwang, I., and Cho, M.J. (1994) Generation of expressed sequence tags from random sequencing of chinese cabbage (*Brassica campestris* L. ssp. pekinensis) flower bud cDNA clones. Unpublished.

Lind, J.L., Bacic, A., Clarke, A.E. and Anderson, M.A. (1994) A style-specific hydroxyproline-rich glycoprotein with properties of both extensins and arabinogalactan proteins. Plant Journal 6, 491-502.

Linskens, H.F. and Heinen, W. (1962).Cutinase Nachweis in Pollen. Z.Bot. 50, 338-347.

Logemann, J., Schell, J. and Willmitzer, L. (1987) Improved method for the isolation of RNA from plant tissue. Analytical biochemistry 163, 16-20.

Lord, E.M. and Kohorn, L.U. (1986) Gynoecial development, pollination and the path of the pollen tube growth in the tepary bean *Phaseolus acutifolius*. American Journal of Botany 73, 70-78.

Lord, E.M. and Sanders, L.C. (1992) Roles for the extracellular matrix in plant development and polination-a special case of cell movement in plants. Developmental Biology 153, 15-28.

Lotan, T., Ori, N. and Fluhr, R. (1989) Pathogenesis-related proteins are developmentally regulated in tobacco flowers. Plant Cell 1, 881-887.

Lutcke, H.A., Chow, K.C., Mickel, F.S., Moss, K.A., Kern, H.F. and Scheele, G.A. (1987) Selection of AUG initiation codons differs in plants and animals. EMBO Journal 6, 43-48.

Marck, C. (1988) 'DNA Strider': a 'C' program for the fast analysisi of DNA and protein sequences on the Apple Macintosh family of computers. Nucleic Acids Research 16, 1829-1837.

Marton, L., Wullems, G.J., Molendijk, L. and Shilperoot, R.A. (1979) In vitro transformation of cultured cells from *Nicotiana tabacum* by *Agrobacterium tumiifaciens*. Nature 277, 129-131.

Mascarenhas, J.P. (1993). Molecular Mechanisms Of Pollen-Tube Growth And Differentiation. Plant Cell 5, 1303-1314.

Mattson, O., Knox, R.B., Heslop-Harrison, J. and Heslop-Harrison, Y. (1974). Protein pellicle of stigmatic papillae as a probable recognition site in incompatibility reactions. Nature 247, 298-300.

McCarter, L.L. (1994) MotY, a component of the sodium-type flagellar motor. Journal of Bacteriology 176, 4219-4225.

McClure, B.A., Gray, J.E., Anderson, M.A. and Clarke, A.E. (1990) Self-incompatibility in *Nicotiana alata* involves degradation of pollen ribosomal RNA. Nature, 347, 757-760.

McClure, B.A., Haring, V, Ebert, P.R., Anderson, M.A., Simpson, R.J., Sakiyama, F. and Clarke, A.E. (1989). Style self-incompatibility gene-products of *Nicotiana alata* are ribonucleases. Nature 342, 955-957.

Meikle, P.J., Bonig, I., Hoogenraad, N.J., Clarke, A.E. and Stone, B.A. (1991) The location of (1-3)-beta-glucans in the walls of pollen tubes of *Nicotiana alata* using a (1-3)-beta-glucan-specific monoclonal-antibody. Planta 185, 1-8.

Memelink, J., Linthorst, J.M.H., Schilperoort, R.A. and Hoge, J.H.C. (1990) Tobacco genes encoding acidic and basic isoforms of pathogenesis-related proteins display differential expression patterns. Plant Molecular Biology 14, 119-126.

Mendez, E., Moreno, A., Colilla, F., Pelaez, F., Limas, G.G., Mendez, R., Soriano, F., Salinas, M. and deHaro, C. (1990) Primary structure and inhibition of protein synthesis in eukaryotic cell-free systems of a novel thionin γ-hordothionin from barley endosperm. European Journal of Biochemistry 194, 533-539.

Milligan, S.B. and Gasser, C.S. (1995) Nature and regulation of pistil-expressed genes in tomato. Plant Molecular Biology 28, 691-711.

Minobe, Y., and Sasaki, T. (1995). Rice cDNA from root. Unpublished.

Mogen, B.D., MacDonald, M.H., Gratbosch, R. and Hunt, A.G. (1990) Upstream sequences other than AAUAAA are required for efficient messenger RNA 3' end formation in plants. Plant Cell 2, 1261-1272.

Mulcahy, G.B. and Mulcahy, D.L. (1982). The two phases of growth of *Petunia hybrida(Hort. Vilm-Andz.)* pollen tubes through compatible styles. Journal of Palynol. 18, 61-64.

Murfett, J., Atherton, T.L., Mou, B., Gasser, C.S. and McClure, B.A. (1994) S-RNase expressed in transgenic *Nicotiana* causes S-allele-specific pollen rejection. Nature 367, 563-566.

Murray, M.G. and Thomopson, W.F. (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Research 8, 4321-4327

Nasrallah, J.B. and Nasrallah, M.E. (1984) Electrophoretic heterogeneity exhibited by the S-allele specific glycoproteins of *Brassica*. Planta 165, 100-107.

Nasrallah, J.B., Kao, T.H., Chen, C.H., Goldberg, M.L. and Nasrallah, M.E. (1987) Amino acid sequence of glycoproteins encoded by 3 alleles of the S-locus of *Brassica oleracea*. Nature 326, 617-619.

Nasrallah, J.B., Kao, T.H., Goldberg, M.L. and Nasrallah, M.E. (1985) A cDNA clone encoding an S-locus specific glycoprotein from *Brassica oleracea*. Nature 318, 263-267.

Nasrallah, J.B., Nasrallah, M.E. (1993). Pollen stigma signalling in the sporophytic self-incompatibility response. Plant Cell 5, 1325-1335

Nasrallah, J.B., Nishio, T. and Nasrallah, M.E. (1991) The self-incompatibility genes of *Brassica*: Expression and use in genetic ablation of floral tissues. Annual Review of Plant Physiology and Plant Molecular Biology 42, 393-422.

Nasrallah, J.B., Rundle, S.J. and Nasrallah, M.E. (1994). Genetic-evidence for the requirement of the *Brassica* S-locus receptor kinase gene in the self-incompatibility response. Plant Journal 5, 373-384.

Nasrallah, J.B., Yu, S.M. and Nasrallah, M.E. (1988) Self-incompatibility genes of *Brassica oleracea*: Expression, isolation and structure. Proc. Natl. Acad. Sci. 85, 5551-5555.

Nasrallah, M.E. (1974). Genetic control of quantitative variation in self-incompatibility proteins detected by immundiffusion.. Genetics 76, 45-50.

Nasrallah, M.E., Kandasamy, M.K. and Nasrallah, J.B. (1992). A genetically defined trans-acting locus regulates S-locus function in *Brassica*. Plant Journal 4, 497-506.

Nessler, C.L. and Burnett, R.J. (1992). Organization of the major latex protein gene family in opium poppy. Plant Mol. Biol. 20, 749-752.

Nessler, C.L. and Mahlberg, P.G. (1976) Lactifers in stamens of *Papaver somniferum* L. Planta 129, 83-85.

Nessler, C.L. and Vonder Harr, R.A. (1990). Cloning and expression analysis of DNA sequences for the major latex protein of opium poppy. Planta 180, 487-491.

Nessler, C.L., Kurz, W.G.W and Pelcher, L.E. (1990) Isolation and analysis of the major latex protein genes of opium poppy. Plant Molecular Biology 15, 951-953.

Newbigin, E., Anderson, M.A. and Clarke, A.E. (1993) Gametophytic self-incompatibility systems. Plant Cell 5, 1315-1324.

Newman, T., de Bruijn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh , L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, E. and Somerville, C. (1994). Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous *Arabidopsis* cDNA clones. Plant Physiology 106,1241-1255

Nishio, T. and Hinata, K. (1977). Analysis of S-specific proteins in stigma of *Brassica oleracea* L. by isoelectic focusing, Heredity 38 391-396.

Nishio, T., Toriyama, K., Sato, T., Kandasamy, M.K., Paolillo, D.J., Nasrallah, J.B. and Nasrallah, M.E. (1992) Expression of S-locus glycoprotein genes from *Brassica oleracea* and *Brassica campestris* in transgenic plants of self-compatible *B.napus* cv westar. Sexual Plant Reproduction 5, 101-109.

Obokota J., Ohme M. and Hayashida N. (1991) Nucleotide sequence of a cDNA encoding a putative glycine-rich protein of 19.7kD in *Nicotiana sylvestris*. Plant Molecular Biology 17, 953-955.

Ockendon, D.J. (1972). Pollen tube growth and the site of the incompatibility reaction in *Brassica oleracea*. New Phytologist. 71, 519-522.

Ockendon, D.J. (1974) Distribution of self-incompatibility alleles and breeding structure in open-pollinated cultivars of Brussels sprouts. Heredity 33, 159-171.

Okamuro, K., den Boer, B.G.W. and Jofuku, K.D. (1993). Regulation of Arabidopsis flower development. Plant Cell 5, 1183-1193.

Ori, N., Sessa, G., Lotan, T., Himmelhoch, S. and Fluhr, R. (1990) A major stylar matrix protein (sp41) is a member of the pathogenesis-related proteins superclass. EMBO Journal 9, 3429-3436.

Pearce, G., Sy., L., Russell, C., Ryan, C.A. and Hess, G.M. (1982) Isolation and characterisation of two polypeptide inhibitors of serine proteinases. Arch. Biochem. Biophys. 213, 456-482.

Pellé, P.G. and Murphy, N.B. (1993). Northern hybridisation: Rapid and simple electrophoretic conditions. Nucleic Acids Research 21, 2783-2784.

Pena-Cortés, H., Willmitzer, L. and Sanchez-Serrano, J.J. (1991) Abscisic acid mediates wound induction but not developmental specific expression of the proteinase inhibitor II gene family. Plant Cell 3, 963-972.

Philipps, G. and Gigot C. (1994) The *Arabidopsis thaliana* transcribed genome: the GDR cDNA program. Unpublished.

Preuss, D., Lemieux, B, Yen, G., Davis, R.W. (1993) A conditional sterile mutation eliminates surface components from *Arabidopsis* pollenand disrupts cell signalling during fertilization. Genes and Development 7, 974-985.

Pruitt, R.E. and Meyerowitz, E.M. (1986). Characterisation of the genome of *Arabidopsis thaliana* Journal of Molecular Biology 187, 169-183.

Quigley F., Villot M-L. and Mache R. (1991) Nucleotide sequence and expression of a novel glycine-rich protein gene from *Arabidopsis thaliana*. Plant Molecular Biology 17, 949-952.

Rae, A.L., Harris, P.J., Bacic, A. and Clarke, A.E. (1985) Composition of the cell-walls of *Nicotiana Alata* Link et Otto pollen tubes. Planta, 166, 128-133.

Ranjeva, R. and Boudet, A.M. (1987) Phosphorylation of proteins in plants: Regulatory effects and potential involvement in stimulus/response coupling. Annual Review of Plant Physiology 38, 73-93.

Roberts, I.N. and Dickinson, H.G. (1983) Intraspecific incompatibility on the stigma of *Brassica*. Phytomorphology 31, 165-174

Roberts, D.M. and Harmon, A.C. (1992). Calcium-modulated proteins: Targets of intracellular calcium signals in higher plants. Annual Review of Plant Physiology and Plant Molecular Biology 43, 375-414.

Roberts, I.N., Harrod, G. and Dickinson, H.G. (1984a). Pollen-stigma interactions in *Brassica oleracea*. I Ultrastructure and physiology of the stigmatic papilar cells. Journal of Cell Science 66, 241-253.

Roberts, I.N., Harrod, G. and Dickinson, H.G. (1984b). Pollen-stigma interactions in *Brassica oleracea*. II The fate of stigma surface proteins following pollination and their role in the self-incompatibility response. Journal of Cell Science 66, 255-264.

Roberts, I.N., Stead, A.D., Ockendon, D.J. and Dickinson, H.G. (1979) A glycoprotein associated with the acquisition of the self-incompatibility system by maturing stigmas of *Brassica oleracea*. Planta, 146 179-183.

Rohde W., Rosch K., Kröger K. and Salamini F. (1990). Nucleotide sequence of a *Hordeum vulgare* gene encoding a glycine-rich protein with homology to vertebrate cytokeratins. Plant Molecular Biology 14, 1057-1059.

Rosen, W.G., (1961) Studies on pollen tube chemotropism. American Journal of Botany 48, 889-895.

Rundle, S.J. and Nasrallah, J.B. (1992) Molecular charaterisation of type-I serine/threonine phosphatases from *Brassica oleracea*. Plant Molecular Biology 20, 1189-1193.

Rundle, S.J., Nasrallah, M.E. and Nasrallah, J.B. (1993) Effects of inhibitors of protein serine threonine phosphatase s on pollination in *Brassica*. Plant Physiology 103, 1165-1171.

Ryan, C.A. (1984) Defense responses of plants. In PlantGene Research E.S. Dennis, B. Hohn, T.Hohn, P. King, J. Schell and D.P.S. Verma, eds (Springer-Verlag, New York) pp 375-386.

Sambrook ,J., Maniatis, T. and Fritsh, E.F. (1989) Molecular cloning , a Laboratory Manual, 2nd Editiom. Cold Spring Harbour Press.

Sanders L.C. and Lord E.M. (1989). Directed movement of latex particles in the gynoecia of three species of flowering plants. Science 243, 1606-1608.

Sanders, L.C., Wang, C.S., Walling, L.L. and Lord, E.M. (1991) A homologue of the substrate adhesion molecule vitronectin occurs in 4 species of flowering plants. Plant Cell 3, 629-635.

Sanfoçan, H., Brodmann, P. and Hohn, T. (1991) A dissection of the cauliflower mosaic virus polyadenylation signal. Genes and Development 5, 141-149.

Sarker, R.H., Elleman, C.J. and Dickinson, H.G. (1988) Control of pollen hydration in *Brassica* requires continued protein synthesis, and glycosylation is necessary for intraspecific incompatibility. Proc. Natl. Acad. Sci. 85, 4340-4344.

Sasaki, T., Miyao, A. and Yamamoto K. (1994) Rice cDNA from shoot. Unpublished.

Sato, T., Thorsness, M.K., Kandasamy, M.K., Nishio, T., Hirai, M. Nasrallah, J.B. and Nasrallah, M.E.. (1991). Activity of an S-locus gene promoter in pistils and anthers of transgenic Brassica. Plant Cell 3, 867-876.

Scutt, C.P., (1990) Molecular biology of self-incompatibility in *Brassica*. PhD thesis, University of Durham.

Sears, M. (1937). Cytological phenomenon connected with self-sterility in the flowering plants. Genetics 22, 130-181.

Shen, B., Carneiro, N., Torres-Jerez, I., Stevenson, R., Helentjaris, T., Habben, J., Larkins, B., Almira, E., Ferl, R. and Baysdorfer, C. (1994) Single-pass sequencing and mapping of clones from two maize cDNA libraries. Unpublished.

Shivanna, K.R. and Sastri, D.C. (1981). Stigma surface esterase activity and stigma receptivity in some taxa characterised by wet stigmas. Annals of botany 47, 53-64.

Showalter A.M., Zhou J., Rumeau D., Worst S.G. and Varner J.C. (1991). Tomato extensin and extensin-like cDNAs: structure and expression in response to wounding. Plant Molecular Biology 16, 547-565.

Showalter, A.M. (1993) Structure and function of plant cell wall proteins. Plant Cell 5, 9-23.

Shutov, A.D., Kakhovskaya, I.A., Braun, H., Bäumlein, H. and Müntz, K. (1995) Legumin-like and vicilin-like seed storage proteins: evidence for a common single-domain ancestral gene. Journal of Molecular Evolution 41, 1057-1069.

Singh, A., Perdue, T.D. and Polillo, D.J. (1989). Pollen-pistil interactions in *Brassica oleracea*: Cell calcium in self and cross pollen grains. Protoplasma 151, 57-61.

Smith, C.J.S., Watson, C.F., Ray, J., Bird, C.R., Morris, Schuch, W. and Grierson, D. (1988). Antisense RNA inhibition of polygalacturonase gene expression in transgenic tomatoes. Nature 334, 724-726.

Somerville, C.J. (1989). Arabidopsis blooms. Plant Cell 1, 1131-135.

Sommer-Knudsen, J., Clarke, A.E. and Bacic, A. (1996) A galactose-rich cell wall glycoprotein from the styles of *Nicotiana alata*. Plant Journal 9, 71-83.

Southern, E. (1979). Gel electrophoresis of restriction fragments. Methods in Enzymology 68, 152-163.

Spaink, H.P. (1992) Rhizobial lipo-oligosaccharides : answers and questions. Plant Molecular Biology 20, 977-986.

Spaink, H.P., Wijfjes, A.H.M., Vanvliet, T.B., Kijne, J.W., Lugtenberg, B.J.J. (1993) Rhizobial lipo-oligosaccharide signals and their role in plant morphogenesis - are analogous lipophilic chitin derivatives produced by the plant. Australian Journal Of Plant Physiology 20, 381-392.

Stanchev, B.S., Doughty, J., Scutt, C.P., Dickinson, H. and Croy, R.R.D. (1996). Cloning of PCP1, a member of a family of pollen coat protein (pcp) genes from *Brassica oleracea* encoding novel cysteine rich proteins involved in pollen stigma interactions. Plant Journal 10, 303-314.

Stead, A.D., Roberts, I.N. and Dickinson, H.G. (1980) Pollen stigma interaction in *Brassica oleracea*: the role of stigmatic proteins in polen grain adhesion. Journal of Cell Science 42, 417-423.

Steer, M.W. and Steer, J.M. (1989) Pollen tube tip growth. New Phytologist 111, 323-358.

Stein, J.C., Howlett, B., Boyes, D.C., Nasrallah, M.E. and Nasrallah, J.B. (1991) Molecular cloning of a putative receptor protein kinase gene encoded at the self-incompatibility locus of *Brassica oleracea*. Proc. Natl. Acad. Sci. 88, 8816-8820.

Stryer, L. and Bourne, H.R. (1986) G proteins: a family of signal tansducers. Annual Review of Cell Biology 2, 391-419.

Tantikanjana, T., Nasrallah, M.E., Stein, J.C., Chen, C.H. and Nasrallah, J.B. (1993) An alternative transcript of the S-locus glycoprotein gene in a class-ii pollen-recessive self-incompatibility haplotype of *Brassica oleracea* encodes a membrane-anchored protein. Plant Cell 5, 657-666.

Thomas, P.S. and Farquar, M.N. (1978). Specific measurement of DNA in nuclei and nucleic acids useing diaminobenzoic acid. Analytical Biochemistry 89, 35-44

Thompson, J.D., Higgins, D.G. and Gibson T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Research 22, 4673-4680

Thompson,K.F. and Taylor, J.P. (1966). Non-linear dominance relationships between S-alleles. Heredity 27, 459-471.

Thorsness, M.K., Kandasamy, M.K., Nasrallah, M.E. and Nasrallah, J.B. (1993) Genetic ablation of floral cells in *Arabidopsis*. Plant Cell 5, 53-261.

Tobias, C.M., Howlett, B. and Nasrallah, J.B. (1992) An *Arabidopsis thaliana* gene with sequence similarity to the S-locus receptor kinase of *Brassica oleracea*: Sequence and expression. Plant Physiology 99, 284-290.

Topping,, T.F., Agyeman, F., Henricot, B., and Lindsey, K (1994) Identification of molecular markers of embryogenesis in *Arabidopsis thaliana* by promoter trapping. Plant Journal 5, 895-904.

Trick, M. and Flavell, R.B. (1989) A homozygous S genotype of *Brassica oleracea* expresses two S-like genes. Mol. Gen. Genet. 218, 112-117.

Tsai, D-S., Lee, H.S., Post, L.C., Kreiling, K.M. and Kao, T-H. (1992) Sequence of an S-protein of *Lycopersicon peruvianum* and comparison with other solanaceous S-sequences. Sexual Plant Reproduction 5, 256-263.

Uknes, S. Dincher, S., Friedrich, L., Negrotto, D., Williams, S., Thompson-Taylor, H., Potter, S., Ward, E. and Ryals, J. (1993). Regulation of pathogenesis-related protein 1a gene expression in tobacco. Plant Cell 5, 159-169.

Uknes, S.J., Mauch-Mani, B., Moyer, M. Potter, S., Williams, S. Dincher, S., Chandler, D., Slusarenko, A., Ward, E. and Ryals, J. (1992). Acquired Resistance in Arabidopsis. Plant Cell 4, 645-656.

Umbach, A.L., Lalonde, B.A., Kandasamy, M.K., Nasrallah, J.B. and Nasrallah, M.E. (1990) Immunodetection and post-translational modification of two products encoded by two independent genes of the self-incompatibility multigene family of *Brassica*. Plant Physiology 93, 739-747.

Underwood, P.A. an Bennet, F.A. (1989). A comparison of the biological activities of the cell adhesion proteins vitronectin and fibronectin. Journal of Cell Science 93, 641-649.

van Engelen, F.A., Hartog, M.V., Thomas, T.L., Taylor, B., Sturm, A., van Kammen, A. and de Vries S.C. (1993) The carrot secreted glycoprotein EP1 is expressed in the epidermis and has sequence homology to *Brassica* S-locus glycoproteins. Plant Journal 4, 855-862.

van Kan, J.A.L., Cornelissen, B.J.C. and Bol, J.F. (1988). A virus inducible tobacco gene encoding a glycine-rich protein shares putative regulatory elements with the ribulose bisphosphate carboxylase small subunit gene. Molecular Plant-Microbe Interactions. 1, 107-112

van Parijs, J., Broekaert, W.F., Goldstein, I.J. and Peumans, W.J. (1991) Hevein- an antifungal protein from rubber tree (*Hevea brasiliensis*) latex. Planta 183-258-264.

Vithanage, H.I.M.V. and Knox, R.B. (1976). Pollen Wall proteins: quantitative cytochemistry of the origins of the intine and exine enzymes in *Brassica oleracea*. Journal of Cell Science 21, 423-435.

Walker, J. (1993) Receptor-like protein kinase genes of *Arabidopsis thaliana*. Plant Journal 3, 451-456.

Walker, J.C. and Zhang, R. (1990) Relationship of a putative receptor protein-kinase from maize to the S locus glycoproteins of *Brassica*. Nature 345, 743-746.

Walterscheid, J. and Milburn, S. (1993) The use of LiCl precipitation for RNA purification. Ambion TechNotes 1 (no.3), 1-4.

Wang, C.S., Walling, L.L., Gu, Y.Q., Ware, C.F. and Lord, E.M. (1994). 2 classes of proteins and messenger RNAs in *Liliun longiflorum* L identified by human vitronectin probes. Plant Physiology 104, 711-717.

Wang, H., Wu, H-M. and Cheung, A.Y. (1993) Development and pollination regulated accumulation and glycosylation of a style transmitting tissue-specific proline-rich protein. Plant Cell 5, 1639-1650.

Ward, E.R., Uknes, S.J., Williams, S.C., Dincher, S.S., Wiederhold, D.L., Alexander, D.C., Ahi-Goy, P., Metraux, J-P. and Ryals, J.A. (1991). Coordinate gene activity in reponse to agents that induce systemic aquired resistence. Plant Cell 3, 1085-1094.

Wemmer, T., Kaufmann, H., Kirch, H.H., Schneider, K., Lottspeich, F. and Thompson, R.D. (1994) The most abundant soluble basic-protein of the stylar transmitting tract in potato (*Solanum tuberosum* L) is an endochitinase. Planta, 194, 264-273.

White, R.F. (1979) Acatylsalicylic acid (asprin) induces resistance to tobacco mosaic virus in tobacco Virology 99, 410-412.

Wu, H-M., Cheung, A.Y.and Wang, H. (1995) A pollen tube growth-stimulatory glycoprotein is deglycosylated by pollen tubes and displays a glycosyltion gradient in the flower. Cell 82, 395-403.

Wu, H-M., Zou, J., May, B., Gu, Q. and Cheung, A.Y. (1993) A tobacco gene family for flower cell wall proteins with a proline-rich domain and a cysteine rich domain. Proc. Natl. Acad. Sci. (USA) 90, 6829-6833.

Wycoff, K.L., Powell, P.A., Dixon, R.A. and Lamb, C.J. (1990). Tissue specific and wound-responsive expression of a hydroxyproline rich glycoprotein gene promoter in transgenic tobacco. Journal of cellular Biochemistry 14E 344.

Xu, D.P., Lei, M. and Wu, R. Expression of the rice OsGRP1 promoter- GUS reporter gene is specifically associated with cell elongation, expansion and differentiation. (1995). Plant Molecular Biology 28, 455-471.

Yang, F., Demma M., Dharmawardhane, S., Warren, V. and Candeelis, J. (1990). Identification of an actin binding protein from Dictyostelium as elongation factor 1α. Nature 347, 494-496.

Ye, Z.H. and Varner, J.E. (1991) Tissue-specific expression of cell wall proteins in developing soybean tissue. Plant Cell 3, 23-37.

Zaki, M.A.M. and Dickinson, H.G. (1990) Microspore derived embryos in *Brassica*-The significance of division symmetry in pollen mitosis-I to embryogenic development. Sexual Plant Reproduction 4, 48-55.

Zhang, R and Walker, J.C. (1993) Structure and expression of the S-locus-related genes of maize. Plant Molecular Biology 21, 1171-1174.

Zhu, J.K., Damsz, B., Kononowicz, A.K., Bressan, R.A. and Hasegawa, P.M. (1994). A higher plant extracellular vitronectin-like adhesion protein is related to the translation elongation factor 1-a. Plant Cell 6 393-404.

Zhu, Q., Dabi, T. and Lamb, C. (1995). TATA box and initiator functions in the accurate transcription of a plant minimal promoter in vivo. Plant Cell 7, 1681-1689.

Zhu, Q., Doerner, P.W. and Lamb, C.J. (1993). Stress induction and developmental regulation of a rice chitinase promoter in transgenic tobacco. Plant Journal 3, 203-212.

Zimmerman, C.R., Orr, W.C., Leclerc, R.F., Barnard, E,C. and Timberlake, W.E. (1980) Molecular cloning and selection of genes regulated in *Aspergillus* development. Cell 21, 709-715.

Zuberi, M.I. and Dickinson, H.G. (1985). Pollen-stigma interactions in *Brassica*. III. Hydration of the pollen grains. Jounal of Cell Science 76, 321-336.