GLYCOSYLHYDROLASES AND THE CONTROL OF MANNOSE/GALACTOSE RATIO IN LEGUME-SEED GALACTOMANNAN

DAVID HOLLAND

This thesis is presented for

the degree of

Doctor of Philosophy

University of Stirling

1999

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

CONTENTS.

Table of Figures

Table of tables.

Acknowledgements

I would like to acknowledge the guidance and support given to me by numerous members of the Departments of Biological Science during the period of my PhD.

I would particularly like to acknowledge Hazel Crombie and John Davis for their advice.

I must thank Grant Reid for his support throughout the course of this project.

Acknowledgement must also be given to all those at Unilever, Colworth for amino acid and DNA sequencing and also for primer synthesis.

I would like to acknowledge the BBSRC for financial support.

Finally I would like to thank my wife Tricia and my family in Sheffield for all the support and encouragement they have given me throughout my studies.

Abbreviations

A	Adenosine
ATP	Adenosine-5'-triphosphate
Bps	Base pairs
C	Cytosine
CTP	Cytosine triphosphate
cDNA	Complimentary deoxyribonucleic acid
DAB	Diaminobenzamide
DEAE	Diethylaminoethyl
DEPC	Diethyl pyrocarbonate
DDT	Dithiothreitol
DNA	Deoxyribonucleic acid
EDTA	Ethylene diamine tetra acetic acid
G	Guanosine
GDP	Guanosine diphosphate
GTP	Guanosine triphosphate
HPLC	High pressure liquid chromatography
IAA	Isoamyl alcohol
IEF	Isoelectric Focusing
I-PCR	Inverse PCR
PNP-α-D-Gal	p-nitrophenyl-α-D-galactosylpyranose
Mr	Molecular weight
MMLVRT	Molonie murine leukaemia virus reverse transcriptase
mRNA	Messenger ribonucleic acid
NTP	Nucleotide triphosphate
PAGE	Polyacrylamide gel electrophoresis
RACE	Rapid amplification of cDNA ends
RNA	Ribonucleic acid
RT	Reverse transcription
PCR	Polymerase chain reaction
SDS	Sodium dodecyl sulphate
SS-PCR	Single sided PCR
T	Thymine
TTP	Thymine triphosphate
UDP	Uridine Diphosphate

Abstract

This thesis describes the work carried out and results obtained from an investigation into the relationship between the two forms of α-galactosidase found within the maturing and germinating S*enna occidentalis* endosperms. These two enzymes are responsible for the final mannose/galactose ratio of galactomannan within the maturing senna endosperm and the mobilisation of the galactomannan storage reserve during germination.

The methods used to characterise the two α-galactosidase enzymes and the results describing the slight but significant dissimilarities are described.

The methods used to isolate RNA from the maturing and germinating endosperms are shown along with a description of the RT-PCR and cloning strategies then employed.

The analysis of the relationship between the cDNAs obtained from the maturing and germinating endosperms used southern blotting, 5'RACE-PCR, DNA sequencing reverse-translation and multiple sequence alignments are described.

Finally the similarities between the cloned senna α-galactosidases and those obtained from the swiss-prot database are shown and the commercial uses of the α-galactosidase are discussed along with the future chances of producing commercially viable transgenic galactomannan yielding crops.

Introduction.

General seed biology

Seeds have evolved to fulfil a number of roles. The seed must protect the embryo from mechanical damage or predation. The seed must also be able to remain viable until favourable conditions for germination occur; this can be until the next growing season or many years hence. The seed must also be able to provide food reserves to support the growth and development of the embryo until the seedling can provide these for itself.

A hardened seed coat or testa provides the mechanical protection of the seed. This testa also prevents the absorption or loss of water by the seed maintaining a water content of between 5 and 10%. The low water levels within the seeds curtails normal metabolic activity although potentially active enzymes are present (Bewley and Black 1978). Most seeds are durable enough to withstand a year or two of "suspended animation" although others are more or less fragile as shown in Table 1 *(Salisbury and Ross 1992)*

Table 1. Representative life spans for seeds.

Species	Viability (%), Initial	Final	Age at Test	Storage Conditions
Sugar maple (Acersaccharinum)			<1 week	
English elm (Ulmus campestris)			6 months	
American elm (Ulmus americana)	70	28	10 months	dry storage
Heavea, Boehea, Thea, sugarcane,	85	-	<1 y	buried 20 cm in soil
Wild oats (Avena fatua)	56	9	1y	buried 20 cm in soil
Alfalfa (Medicago sativa)	50	1	6y	buried 20 cm in soil
Yellow foxtail (Setaria lutescens)	57	4	10y	buried 20 cm in soil
Cocklebur (Xanthium strumarium)	91	15	16y	buried 20 cm in soil
Canada thistle (Cirsium arvense)	90	1	21y	buried 20 cm in soil
Kentucky bluegrass (Poa pratensis)	89	1	30y	buried 20 cm in soil
Red clover (Trifolium pratense)		1	30y	buried 20 cm in soil
Tobacco (Nicotiana tabacum)		13	30y	buried 20 cm in soil
Button clover (Medicago orbicularis)			78y	herbarium
Clover (Trifolium striatum)			90y	herbarium
Big trifoil (Lotus uliginosus)		1	100y	dry storage
Red clover (Trifolium pratense)		1	100y	dry storage
Locoweed (Astragalus massiliensis)			100-150y	herbarium
Sensitive plant (Mimosa glomerate)			221y	herbarium
Indian lotus (Nelumbo nucifera)			1040y	peat bog
Arctic Lupine (Lupinus arcticus)			10000?	frozen silt,

				lemming burrows

Note: blank cells are unmeasured

Germination takes place after impaction, scarification or other means causes a breach of the testas integrity (unless a dormancy mechanism is in operation). Water can then be absorbed and the seed expands, splitting off the testa. As long as certain requirements, such as temperature and the availability of O_2 are met, metabolic activity is initiated; the embryo resumes cellular division and elongation, storage reserves are mobilised and eventually the radical (embryonic root) emerges from the seed. After radical emergence has occurred, germination is officially over and seedling growth takes place, however until the shoot is above ground and photosynthesis starts, the young seedling is still dependent on the seed reserves to support growth.

The storage material on which almost all seeds depend is usually lipid or carbohydrate (with lipid being the most common). Proteins are also stored within the seed but are usually not the most abundant storage material present. Table 2 gives some idea of the proportions of lipid, carbohydrate and protein commonly found within seeds and also in which tissue they located (Bewley and Black 1978)

Table 2. Storage material composition and localisation.

Species		Average % composition dry wt		Storage Organ
	Protein	Lipid	CHO major type	
Corn	11	5	75 (Starch)	Endosperm
Sweet corn	12	9	70 (Starch)	Endosperm
Oats	13	8	66 (Starch)	Endosperm
Wheat	12	2	75 (Starch)	Endosperm
Rye	12	2	76 (Starch)	Endosperm
Barley	12	3	76 (Starch)	Endosperm
Broad bean	23	1	56 (Starch)	Cotyledon
Flax	24	36	24 (Starch)	Cotyledon
Field pea	24	6	56 (Starch)	Cotyledon
Garden pea	25	6	52 (Starch)	Cotyledon
Peanut	31	48	12 (Starch)	Cotyledon
Soybean	37	17	26 (Starch)	Cotyledon
Cotton	39	33	15	Cotyledon
Rape	21	48	19 (Starch)	Cotyledon
Watermelon	38	48	5	Cotyledon
Brazil nut	18	68	6	Radicle/Hypocotyl
Oil palm	9	49	28	Endosperm
Ivory nut	5	1	79 (Mannan)	Endosperm
Date	6	9	58 (Mannan)	Endosperm
Castor bean	18	64	Trace	Endosperm
Pine	35	48	6	Megagametophyte

The types of storage carbohydrates.

The most common and commercially important storage carbohydrate is the α-1,4 and 1,6 linked glucose polymer, starch. Starch is stored as grains located in amyloplasts within the cytoplasm of the storage cells of the seed. Less common than starch (although more significant in some plants) and located within the cell walls rather than within the protoplast are a variety of other storage polysaccharides. These cell wall storage carbohydrates can be classified as galactan, xyloglucan and mannan based compounds. Galactans, the rarest of the cell wall storage compounds consist of a β-1,4 linked galactan with arabinose, xylose and uronic acid residues associated (Crawshaw and Reid 1984). In tissues other than storage tissues these compounds are classed as belonging to the pectin family. Xyloglucan structure is described as a β-1,4 linked glucose chain with xylose residues linked α-1,6 to this backbone. The xyloses are further substituted by β-1,2 linked galactosyl units. As with the galactans the xyloglucans are not restricted to the role of seed storage reserve but are also found in the plants primary cell walls. The mannan family of polysaccharides consists of pure mannans, glucomannans and the galactomannans. The pure mannans consist of a linear β-1,4 linked mannan chain with less than 2% of the mannan residues being linked to via their 6 position to a galactosyl residue (Aspinall et al 1953). The structure of the glucomannans is that of a linear β-1,4 linked chain consisting of both mannosyl and glucosyl residues in approximately equal amounts. Like the pure mannans, glucomannans can have a very low level of galactosyl substitution (Goldberg 1969). Galactomannans are described below.

The source of galactomannans:

Galactomannans obtained from the seeds of higher plants, come mainly from the Leguminosae family. E. Anderson (1949) examined the seeds of 163 legume species and found that 119 contained "mucilages". When investigated further, these "mucilages" were identified as galactomannan. In these plants the galactomannans are located within the endosperms (the non-endospermic Leguminosae do not contain galactomannans).

The primary structure of galactomannans:

Galactomannans from various sources have been structurally analysed. This has usually been carried out by an initial hot or cold water extraction from ground seed or endosperm flour. The soluble polysaccharides are isolated from the aqueous mixture by centrifugation and recovered by addition of ethanol. Galactomannans are then isolated from this polysaccharide mix by the addition of Cu^{2+} or Ba^{2+} ions which complex to the galactomannans, causing their selective precipitation (Hui and Neukom, 1964; Dea and Morrison, 1975). The isolation of galactomannan by these means has been shown to maintain its native chemical composition; as shown by the ratios of the galactose to mannose in the galactomannan and from the optical rotation caused by the galactomannan (Unrau and Choy, 1970), (Rao, Chouldhury, and Bagchi 1961)

The ratio of mannosyl to galactosyl residues in the isolated galactomannans was obtained by total acid hydrolysis. The man:gal ratio varies with the source of the

endospermic tissue (Reid and Meier, 1970; Dea and Morrison, 1975); the primitive Leguminosae-caesalpiniodeae have a high man:gal ratio whereas the more advanced Leguminosae-faboideae have a lower man:gal ratio. Examples of the man:gal ratios are compiled in Tables 3a and 3b.

Table 3. The composition of galactomannan isolated from the Tribe Trifolium (Reid and Meier, 1970; Dea and Morrison, 1975).

Species	Section	%Galactose	%Mannose
Trifolium incarnatam -C. presl.		46	54
Trifolium dubium Sibth		46	54
Medicago sativa L.		45.5	54.5
Medicago lupulina L		46	54
Medicago radiata (L.) Heyn		46	54
Melilotus alba medicus		47	53
Melilotus offiinalis (L.) pallas		45.5	54.5
Trigonella calliceras Fisch	Callicerates Boiss	47	53
Trigonella corniculata L	Falcatulae Boiss.	46	54
Trigonella hamosa L.	Falcatulae Boiss	46	54
Trigonella caerulea (L.) Ser	Capitatae Boiss.	47	53
Trigonella monspeliaca L.	Reflexae (Sirj) vass.	48	52
Trigonella polytcerata Bieb.	Bucerates Boiss	47	53
Trigonella foenum-graecum	Foenum-graecum Ser	47	53
Triqonella cretica (L.) Boiss.	Samaroideae	39	61

Table 3b. The composition of galactomannan isolated from the Tribe Genisteae (Reid and Meier, 1970; Dea and Morrison, 1975).

Species	%Galactose	%Mannose
Spartium junceum L.	31	69
Genista ovata W.K.	33	67
Genista tinctoria L.	32	68
Genista monosperma Lam	32	68
Petteria ramentacea (Sieb) Presl.	32	68
Labernum alpinum (Mill) Presl.	32	68
Labernum anagyroides medicus	31	69
Ulex europaeus L.	34	66
Cytisus Hirsutus L.	30	70
Cytisus supinus L.	33	67
Sarothamnus scoparius L.	30	70

The linkage types found in the galactomannan have been analysed by methylation analysis, peroxidate oxidation and partial hydrolysis (this work being carried out initially on the commercially important guar and locust-bean galactomannan (Smith, 1948) (Ahmed and Whistler 1951) (Whistler and Durso 1951). The mode of linkage (α or β) was elucidated through the use of optical rotation analysis. (Kooiman, 1972) (Leschziner and Cerezo,1969)

Bringing together the data from the structural analysis has led to the generally accepted proposition that galactomannans are all comprised of a linear 1→4 β-linked D-mannan backbone with single D-galactopyranosyl units attached via 1→6 α-links (fig. 1)(Dea and Morrison, 1975).

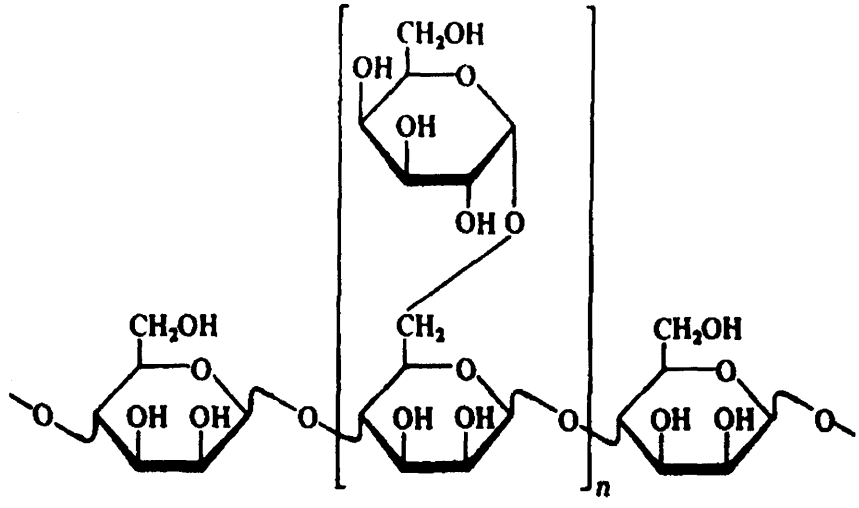


Figure 1. The primary structure of galactomannan.

The only ambiguity that remains in the model above is that of where the galactosyl residues are attached along the mannan backbone. The possible arrangements of the galactosyl units are shown below (*Fig2).*

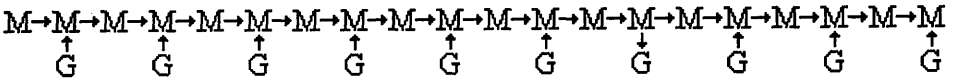


Figure 2A

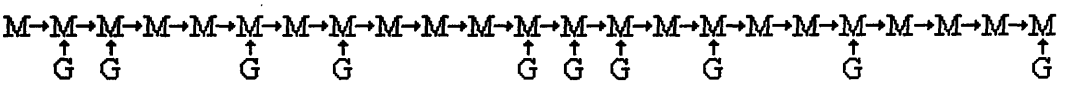


Figure 2B

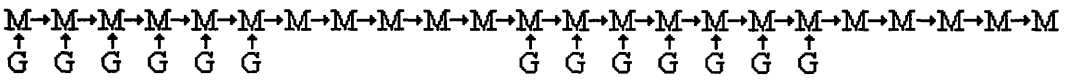


Figure 2C

Figure 2A, B and C. Shows the possible arrangements of the galactosyl residues along the mannan backbone.

A shows the arrangement along the backbone if substitution was regular. 2B is the structure of galactomannan when galactosyl substitution is random. 3C is the product of non-random yet non-regular galactosyl substitution, giving rise to block structures.

Digestion of galactomannans with endo-β-mannanases gave rise to large segments of polysaccharide with almost total galactosyl substitution of the mannan backbone (Courtois and Le Dizet, 1968) (Baker and Whistler, 1975) leading to the supposition that galactosyl substitution is not random but is also not regular (Dea and Morrison, 1975). Native galactomannans therefore seem to conform to the block structure shown in fig. 2C.

Galactomannan synthesis.

It was over one hundred years ago that the first investigations into galactomannan formation were undertaken. In 1890 Nadelmann documented how "mucilages" were seen forming in cytoplasmic vacuoles of four leguminous species. These mucilages were then deposited in the cell walls of the endosperm cells. More modern investigations using periodate-shiff staining combined with interference microscopy and electron-microscopy have confirmed that Nadelmann was correct; galactomannans are formed inside the cisternae of the endoplasmic reticulum within the cytoplasm. The endoplasmic reticulum becomes so swollen with galactomannan that it can be observed through a light microscope. The galactomannan accumulations are subsequently discharged into the cell wall. In fenugreek the filling of the cell walls of the endosperm begins adjacent to the embryo and continues, filling each cell completely occupying the space previously taken up by the cytoplasm (which degrades), until only a single outer layer of cells (the aleurone layer) remains unfilled (Reid and Meier, 1970. Reid and Meier, 1973. Meier and Reid, 1977.).

The first study into the biochemistry of galactomannan formation was carried out by Campbell and Reid in 1982. This investigation showed that during the period of galactomannan accumulation, endosperm homogenates contain a membrane bound enzyme activity that is able to transfer D-mannosyl residues from GDP-mannose to an acceptor molecule. The enzyme activity is associated with the endoplasmic reticulum (Campbell, 1978). The acceptor for this mannosyl-transferase was shown, through the

use of selective precipitations with borate, transition metal ions and the galactosyl binding lectin from *Ricinus communis,* to be galactomannan.

When the enzyme preparation was incubated with GDP-mannose, the acceptor molecule and a Mg^{2+}, Ca^{2+} or Mn^{2+} co-factor (in the absence of UDP-galactose) a 1→4 β linked mannan was produced. If UDP-galactose was also present in the reaction mixture, high rates of galactose transfer were also observed. The product thus produced was galactomannan. Conversely if the membrane preparation was incubated with UDP-(^{14}C)-D-galactose, co-factors and acceptor molecules (in the absence of GDP-mannose) no radiolabelled galactose was incorporated into a polysaccharide. The mannosyl-transferase activity can therefore operate in the absence of UDP-galactose; whereas the galactosyl-transferase is dependent on the transfer of mannose from GDP-mannan to a growing mannan backbone, as catalysed by the mannosyl-transferase. It was also shown that only simultaneous mannan transfer enabled the galactosyl-transferase to add the galactose side chains; pre-formed mannan chains would not accept galactosyl substitutions (Reid et al, 1987. Edwards et al 1989).

The results of these experiments enabled a model of galactomannan biosynthesis to be proposed, this is shown below:-

I. Two membrane bound enzymes are involved: a GDP-mannose-dependent mannosyltransferase and a UDP-galactose-dependent galactosyltransferase.

II. The product of the mannosyltransferase when UDP-galactose is absent is linear (1→4)-β-D-mannan.

III. The glycosylacceptor for the galactosyltransferase is the product of the mannosyltransferase. The final product of the two enzymes is a galactomannan of the same structural type as that found in the plants (*in vivo)* from where the enzymes have been prepared (fenugreek and guar).

IV. Galactosyl transfer can only occur onto a nascent or newly transferred D-mannosyl residue.

The control of the ratio of mannose to galactose was investigated *in vitro*. The man:gal ratio was manipulated by varying the concentrations of the sugar- nucleotide precursors. With both substrates set at their saturating concentrations, high man:gal galactomannan was produced in both fenugreek and guar (which *in vivo* produce low man:gal galactomannan). By reducing the GDP-mannan concentration whilst maintaining the UDP-galactose at a saturating concentration, galactomannans with degrees of substitution approaching (but not exceeding) those found *in vivo* could be produced (Edwards et al, 1989).

In 1992 Edwards et al stated that the levels of mannosyl- and galactosyltransferases within the developing endosperms of fenugreek, guar and senna correlate closely with the deposition of galactomannan in each of these plants, thus confirming the role of the two enzymes in galactomannan biosynthesis. The relative *in vitro* activities of the mannosyl- and galactosyltransferases in fenugreek and guar are similar and constant throughout the period of galactomannan synthesis, however the ratio of mannosyl to galactosyl-transferase was slightly higher in guar than fenugreek, mirroring the final mannose:galactose ratio of the galactomannan of these two species.

Further investigations using enzymatic fingerprinting techniques (Reid et al 1995) were needed in order to establish whether the mechanism behind the distribution of galactose on the newly formed galactomannan was (A), via the *in vivo* control of the sugar-nucleotide levels (as suggested in Edwards 1989) or (B), via the transfer specificities of the transferase enzymes (also suggested by Edwards because of the limited effects of *in vitro* manipulation of the sugar-nucleotide concentrations). Mannosyl and galactosyl-transferases were isolated from seed endosperms and used in vitro to produce radio–labelled galactomannan. The galactomannan was subjected to an exhaustive digestion by endo-(1→4)-β-D-mannanase from *Aspergillus niger*. The subsite specificities of this enzyme are well known. The *Aspergillus niger* (1→4)-β-D-mannanase has an optimum substrate binding requirement of five (1→4)-β-linked D-mannosyl residues, however unsubstituted mannotetraose is digested; mannotriose on the other hand, is only hydrolysed very slowly. Substitutions on the mannan chain being hydrolysed affects the activity of the mannanase. Galactosyl substitutions on the second and/or forth mannosyl residue prevents hydrolysis; substitutions elsewhere have no effect (this is shown in fig 3 McCleary and Matheson 1983, McCleary 1979)

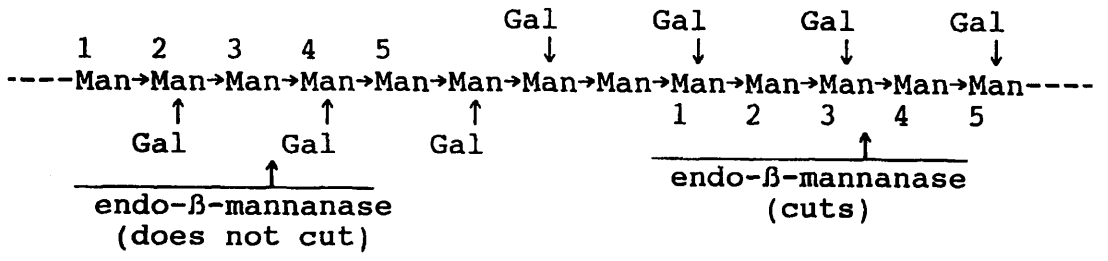


Fig 3. The action of the Aspergillus niger mannanase on galactomannan.

Galactomannans hydrolysed by this mannanase are therefore split into well-defined fragments. These fragments were separated by thin layer chromatography and the relative amounts of each of these fragments, as measured by digital autoradiography were used to calculate the statistical distribution of the galactosyl residues along the mannan backbone (Reid et al 1995). A computer program was then set up to simulate the galactomannan biosynthetic process. It modelled the elongation of the mannan backbone and the transfer of galactosyl residues onto the newly transferred (terminal) mannosyl residue. Whether or not a galactosyl residue would be added to the growing terminus of the galactomannan was, in the model, dependent on the terminal mannosyl's nearest neighbour and next nearest neighbour. The state of substitution of only the two nearest mannosyl residues to the terminal mannosyl residue affected the probability of galactosyl transfer occurring onto that terminal mannosyl residue. This assumption is known as a second-order Markov chain assumption. After many simulated galactomannans were "synthesised" (using different values for the effects of the degrees of substitution of the two neighbouring mannosyl residues), a synthetic *Aspergillus niger* (1→4)-β-D-mannanase hydrolysis was carried out. On comparison of the simulations with real life, *in vitro* galactomannan biosynthesis and degradation, three points became apparent.

1) The second order Markov chain assumption built into the simulation was correct. *In vivo* the galactosyl-transferase transfers a galactosyl residue onto the terminal mannosyl residue depending on the pattern of substitution of the two preceding ones.
2) The differing patterns of substitution of the two crucial neighbouring residues affect the galactosyl-transferases of fenugreek, guar and senna differently. The

different specificities of these three enzymes give rise to different statistical pattern of galactosyl substitution along the mannan backbone.

3) The statistical substitution rules for all three species used in the investigation gave maximum degrees of substitution which are those seen in the primary galactomannan products *in vivo*.

In guar and fenugreek the man:gal ratio of galactomannans synthesised remains constant throughout the stages of seed maturation. In senna however, the man:gal ratio increases during maturation from approximately 2.3 to about 3.3 at maturity. The question then asked was, could this increase in ratio be due to the increase in mannosyltransferase activity relative to galactosyl transferase activity? After charting the changing man:gal ratio during maturation it was obvious that the change in the ratio of man:gal immediately before the desiccation stage of the seeds maturation was too abrupt to be explained by the slow linear increase in activity of the mannosyltransferase relative to galactosyltransferase. On examination of the activity of α-galactosidase (an enzyme that removes galactose from galactomannan) within the senna endosperm, a strong correlation between increased α-galactosidase activity and the rapid increase in senna man:gal ratio was established. Fig 4 and 5 (Edwards et al 1992).

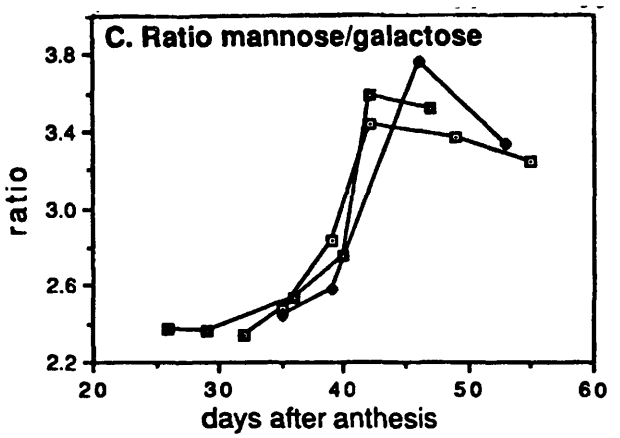


Figure 4. The rapid change in man:gal ratio during maturation of the senna endosperm.

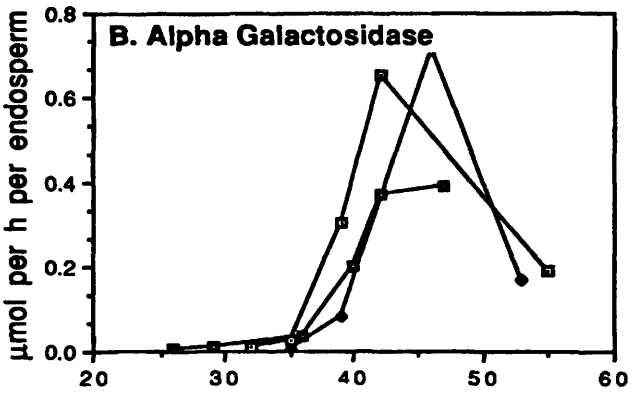


Figure 5. A senna α-galactosidase activity that appears at the same time as the man:gal ratio increases.

In summary one can state that the process of galactomannan biosynthesis is directed by two enzymes, the mannosyltransferase and the galactosyltransferase. In some plants the primary product of these enzymes can be modified by a third enzyme, the α-galactosidase. The biosynthetic process can only be fully modelled when all three enzymes working in concert are considered.

The Catabolism of galactomannans.

In addition to his pioneering work on galactomannan anabolism, Nadelmann was the first person to describe the disappearance of galactomannan (which he describes as "mucilages") after germination.

The next work carried out on galactomannan catabolism was eighty years later. In fenugreek, germination is complete in 10 hours (Reid and Bewley, 1979), however the amount of galactomannan is stable until 14 hours after the onset of germination at which time there is a rapid decrease in its amount until there is none left. When the degradation of galactomannan is compared to the breakdown of storage proteins and lipids (found in the cotyledons) (Leung *et al.*, 1981) it is obvious that the galactomannan breakdown is very rapid. As the galactomannan is broken down in the endosperm, sucrose (and transitory starch) levels in the embryo increase in amount proportionally (Reid, 1971). Galactose and mannose are not detected within the embryo (Reid, 1971; Uebelmann, 1978) which suggests that they must be phosphorylated quickly and converted into sucrose and starch. Within the endosperm however, there is a low level of galactose, mannose and manno-oligosaccharides. If uptake into the embryo and conversion into sucrose and starch is prevented these levels quantitatively mirror the amount of galactomannan that is broken down (Reid and Meier, 1972).

The breakdown of galactomannan into galactose and mannose is accomplished through the employment of three enzymes; *endo*-β-mannanase, β-mannosidase and an α-galactosidase (which in senna may or not be the same α-galactosidase observed in the biosynthesis of the galactomannan). The pathway of galactomannan mobilisation, derived from fenugreek data but applicable to all endosperm galactomannan utilisation is shown in fig 6.

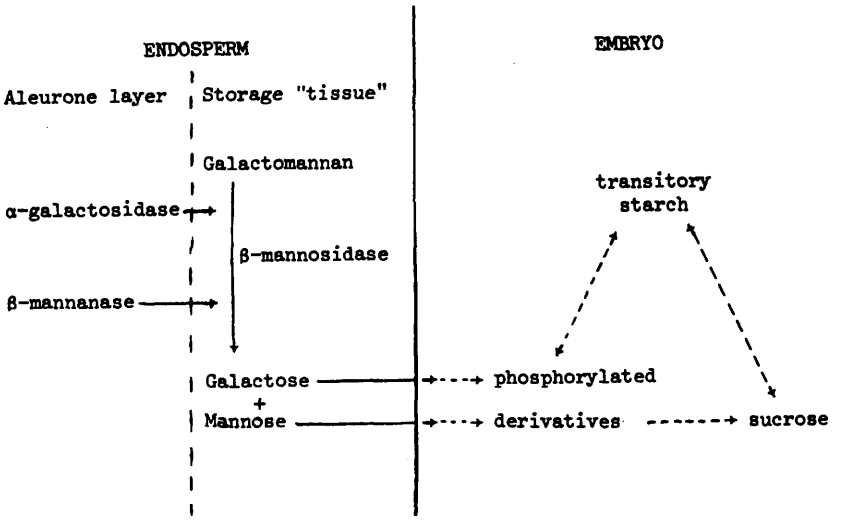


Fig 6. A systematic diagram of galactomannan mobilisation in the fenugreek seed.

The only differences between the fenugreek, guar and carob seeds galactomannan mobilisation are the result of the seeds morphology. The fenugreek and guar seeds have an endosperm totally filled with galactomannan; the only living part of which is the alurone layer. It is from this that the three catabolic enzymes are produced, the result of which is a zone immediately within the alurone layer of galactomannan depolymerisation. This zone moves inwards towards the embryo until all the galactomannan has been broken down (Reid, 1971; Reid and Meier, 1972). In carob

the cells of the endosperm are not entirely filled with galactomannan and so retain a living protoplast. On germination the all the cells of the endosperm produce the galactomannan hydrolysis enzymes, leading to a small level of galactomannan breakdown throughout the endosperm. This however rapidly stops and the majority of galactomannan mobilisation occurs from the embryo outwards (Seiler, 1977). This is probably the result of the breakdown products inhibiting the hydrolytic enzymes, it is only through the embryo that these products are removed from the endosperm.

The biological importance of galactomannans

De Vries in 1877 came to the conclusion that the "mucilages" found within red clover endosperms were used to store water during germination. Nadelmann (1890) however disputed this and concluded that the "mucilages" are reserves. Following Nadelmann it was taken for granted that the galactomannan "mucilages" are carbohydrate reserves, however almost one hundred years later the multifunctionality of the endospermic galactomannan was demonstrated.

On comparison of whole fenugreek seeds and those stripped of their endosperms it was concluded that there is no reason that 30% of the seeds storage material should take the form of the complex hydrophilic galactomannan rather than any other storage material such as protein, lipid or starch (Reid and Bewley, 1979). The benefit of investing the seeds reserves in galactomannan seemed to be that being hydrophilic it absorbs large amounts of water which it then distributes through the endosperm, entirely surrounding and buffering the embryo against water loss. The buffering is due to the galactomannan giving the endosperm the ability to lose (or gain) water without its water potential changing; therefore an embryo is only in contact with the water potential of the endosperm and will not lose water in a desiccating environment until the endosperm has first lost all of its water content. It is only after germination that the galactomannan is mobilised and used as a source of carbohydrate for the growing seedling.

The galactomannan therefore clearly has two functions,

1. The galactomannan maintains the water potential around the embryo, buffering it against desiccation.
2. The galactomannan acts as a carbohydrate reserve, utilised by the growing seedling after germination.

Role 1 above is particularly important for these plants as they are thought to have originated in the semi-arid region of the Eastern Mediterranean (Hegi, 1935)

The commercial importance of galactomannans

Galactomannans are exploited commercially mainly because of three properties they have;

1. At very low concentrations (1%) galactomannans form highly viscous aqueous solutions that are stable within the pH range of 1-10.5 (due to its non-ionic make-up). The viscous galactomannan solutions are also stable to heating and cooling and also the effects of the ionic properties of the solution. This property makes galactomannans ideal for use in the food industry as a thickener for soups, desserts, sauces, mayonnaise, et cetera.
2. Purified galactomannans are able to imbibe water (as they do in germinating seeds), forming a thick paste that is impervious to further wetting. This makes galactomannans an ideal coating for waterproofing explosives as a plugging agent in leaking oil wells and as an additive to increase the wet strength of paper.
3. Small amounts of galactomannan will interact with certain polysaccharides, promoting their gel forming abilities. This means that a small amount of inexpensive galactomannan can replace a large proportion of expensive agarose or carrageenan in a gel mixture, thus producing a cost saving. The addition of galactomannan to gels also results in gels with different properties to the standard gels; galactomannan mixed with xanthan from *Xanthomonas campestris* (which on its own does not gel at any concentration) produces a gel when the galactomannan and xanthan concentrations are as low as 0.5%. This gel has a distinctive mouth-feel that is useful in the food industry (Dea et al, 1977; Rocks, 1973).

The ability of a polysaccharide to promote the formation of a gel (or a highly viscous solution) depends on its ability to behave like an intimately packed, structurally regular, closely interacting polysaccharide, such as cellulose whilst at the same time displaying properties of random coil polysaccharides, where chain-chain interactions do not occur. This demands that a gel-forming molecule should have:

1. Inter-chain association areas formed by long structurally- and conformationally-regular regions.
2. Conformationally-irregular non-associating regions where chain-solvent interactions predominate.

In galactomannans the inter-chain association areas equate with the "smooth" regions of unsubstituted mannan which aggregate in a regular, ribbon like conformation. The highly substituted "hairy" regions of the galactomannan chains interact with water making the network soluble (fig 7) (Dea et al, 1977).

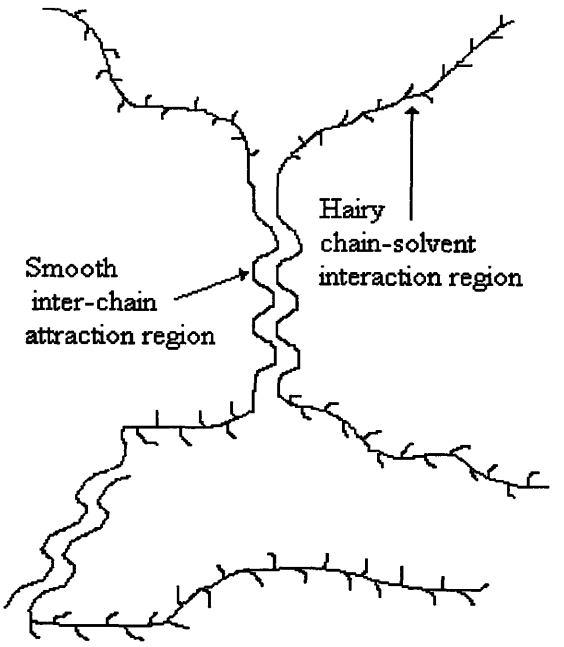


Figure 7. The structure of a galactomannan gel.

Investigations were carried out to discover the importance of different degrees of block type substitutions of the galactomannan on its ability to form inter-chain self associations. Galactomannans from three sources were used:

1. Locust bean gum (from *Ceratonia siliqua*) with a Man:Gal ratio of 3.5:1 and a block structure;
2. Tara gum (from *Caesalpinia spinosa)* with a Man:Gal ratio of 3:1 and a semi block structure;
3. Guar gum (from *Cyamopsis tetragonoloba)* with a Man Gal ratio of 1.6:1 and little if any block-like pattern of substitution.

The results showed that guar gum galactomannan chains were not self-associating, the tara gum at 0.75% formed some associations that produced a weak gel that broke down on heating to 30°C, the locust bean gum produced a cohesive (though weak) gel at concentrations as low as 0.5%. The degree of block structure possessed by a galactomannan was therefore proved to determine its ability to form viscous solutions and gels through inter-chain self interactions.

The ability of galactomannan to promote the gelling of other polysaccharides is also dependent on its ability to present unsubstituted mannan backbone to the polysaccharide it is mixed with. McCleary in 1979 showed this by comparing the effects of the addition of galactomannan from either *Leucaena leucocephala* or *Cyamopsis tetragonoloba* to xanthan. *Leucaena leucocephala* galactomannan-xanthan mixtures formed cohesive gels down to concentrations of 0.1% galactomannan plus 0.1% xanthan. *Cyamopsis tetragonoloba* galactomannan-xanthan mixtures would not gel at any concentration (the only effect that could be obtained was a raising of the solutions viscosity). This difference in action of the two galactomannans was not

caused through different amounts of galactosyl substitution (both were 40% substituted) but by the different pattern of substitution. *Leucaena leucocephala* galactomannan has galactosyl residues arrayed only down one side of the mannan backbone (assuming an alternating conformation), making galactomannan-xanthan interactions possible, whereas *Cyamopsis tetragonolobus* galactomannan has its galactosyl residues randomly distributed around the mannan backbone, thus preventing galactomannan-xanthan interactions occurring (fig 8).

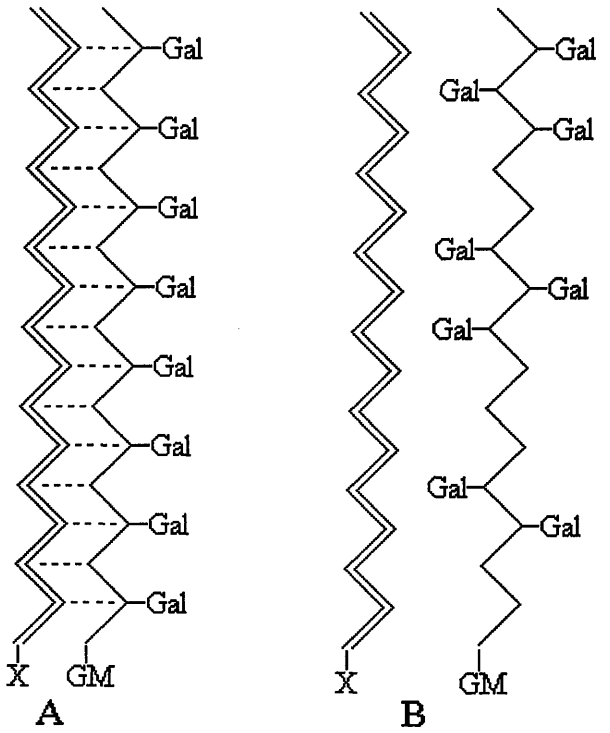


Figure 8. The interaction (dotted lines) of galactomannans (GM) from (A) Leucaena leucocephala and (B) Cyamopsis tetragonolobus with xanthan (X).

The commercial importance of galactomannans with low levels of galactosyl-substitution in blocky arrangements makes a working knowledge of the α-galactosidase that strips the locust bean and senna nacent galactomannan of many of its galactosyl stubs very important. In the future it may be possible to engineer an easily grown species, such as guar, to contain the α-galactosidase enzyme found in the maturing senna endosperm so that it will express a more valuable, low galactose galactomannan.

Aims

The initial aim of this project was to compare the α-galactosidase involved in the biosynthesis of *Senna* galactomannan with that found in the germination of the S*enna* seeds. The project was designed to answer the question of whether *in vivo* these two enzymes are,

A, Identical,

B, From the same gene but different due to post-translational modifications,

Or C, Different enzymes produced from different genes.

Following on from the initial aim was the intention to obtain a full clone of the maturing α-galactosidase gene as this would be important in future attempts to engineer a transgenic guar which would then produce the commercially valuable low galactose galactomannan.

Materials and Methods.

Growing conditions of the *Senna occidentalis.*

The *Senna occidentalis* were obtained as seeds from the Royal Botanic Gardens, Kew, England and were used to build up a plant population from which seeds for the next generation were collected. The plants were grown under the conditions: day/night = 11h at 22°C / 13h at 20°C; photon flux rate at table level = 200 μmol m^{-2} s^{-1}; mineral nutrition was with Phostrogen fertiliser from Phostrogen, Corwen, Clwyd, UK. The concentration used was 0.5g l^{-1} applied continually through capillary matting. Hand pollination was required for fertilisation. Anthers were removed from the open flowers, cut open and the pollen removed. The pollen was transferred to stamen by paintbrush.

Isolation of α-galactosidase from senna seeds.

α-galactosidase was isolated from maturing senna seeds (35 days after pollination) and germinating senna seeds (3 days after imbibition). The endosperms were hand dissected from the seeds and 50 were added to 2.5ml of ice cold $^{1}/_{10}$ th strength McIlvaines buffer pH 5.0 (Dawson et al, 1982). The endosperms were ground with 2g acid washed sand and incubated with 12.5μl of stock *Aspergillus niger* endo-(1-4)-β-mannanase at 30°C for 30 minutes. The homogenate was centrifuged at 5,000g for 5 minutes to remove the sand. A further centrifugation at 20,000g for 20 minutes removed cellular debris and the supernatant was stored in 500μl aliquots at -20°C.

α-Galactosidase assay.

25µl of enzyme sample was added to 375µl of 1x McIlvains buffer (pH 5.0) and incubated at 30°C for 2 minutes. 100µl of 50mM p-nitrophenyl-α-D-galactosylpyranose (PNP-α-D-Gal) was added to this mixture and incubated at 30°C for a further 15 minutes. The reaction was then stopped by the addition of 0.1M Na_2CO_3. The absorbance of the released p-nitrophenol was measured at 400nm.

Polyacrylamide Gel Electrophoresis (SDS - PAGE)

SDS-polyacrylamide electrophoresis was carried out using the Neuhoff et. al (1986) modifications of the original Laemmli (1970) method.

The following solutions were prepared:

Separating gel buffer: 1.5 M Tris, 0.4% SDS, pH adjusted to 8.7 with conc. HCl.

Stacking gel buffer: 0.5 M Tris, 0.4% SDS, pH adjusted to 6.8 with conc. HCl.

Sample buffer: 0.27 M Tris, 10.7% SDS, 27% glycerol, bromophenol blue, pH adjusted to 6.8 with phosphoric acid

Reservoir buffer: 0.125 M Tris, 0.960 M Glycine, 0.5% SDS pH 8.3 when diluted.

Acrylamide stock solution : 30% Acrylamide, 0.8% Methylenebis-acrylamide.

Polymerising agent: 15µl TEMED, 210µl 10% freshly made Ammonium persulphate

Coomassie blue stain: 0.1% Brilliant Blue in acetic acid, methanol and distilled water in a ratio of 2:5:5.

The gels were prepared by mixing-
5ml of separating gel buffer,
6.65ml of Acrylamide stock solution
8.35ml dH_2O
The mix was deaerated.
15µl of TEMED and 210µl 10% fresh Ammonium persulphate were added and....

1mm thick separating acrylamide gels (10 x 12 cm) were cast using the gel pouring equipment. The gels were rinsed with distilled water and stored in cling film at 4°C for no more than 14 days. When required stacking gels (6.1ml dH_2O, 1.34ml acrylamide stock, 2.5ml stacking buffer, 10µl TEMED and 100µl ammonium persulphate) were poured onto the separating gels and run in reservoir buffer using Hoefer Mighty Tall gel apparatus.

Samples were prepared by adding sample buffer/DTT mix (containing 3 parts sample buffer to 1 part 1M DTT) in a ratio of 3:1 followed by boiling for 3 min. The boiled samples were then injected using a Hamilton syringe into the gel wells and each gel ran at a constant current of 25mA. The gel was stopped and the apparatus disassembled when the dye front reached the bottom of the gel (this took approximately 1hour and 30 minutes.

Gels for staining were agitated in Coomassie stain for 30 min then destained with 10% acetic acid.

Western Blotting.

The following solutions were prepared:

CAPS buffer: 100mM 3-cyclohexylamino-1-propanesulphonic acid solution pH11

Blocking buffer: 4g Marvel powdered milk in 200ml Phosphate buffered saline (10mM Potassium phosphate, 150mM Sodium chloride, pH 7.4) + 200µl Triton X-100.

A SDS-PAGE gel was run as described above. The gel was then blotted onto nitro-cellulose membrane in CAPS buffer for 2 hours at 50V in a Hoefer Transphor electrophoresis unit. The blot was then incubated over night at room temperature gently shaking in 100ml of blocking buffer. A second incubation was prepared. The blotting membrane was placed in 20ml of blocking buffer containing 7.5µl of rabbit anti α-galactosidase antibody raised by Unilever against guar germinating α-galactosidase (Overbeek et al 1989). An incubation for 2 hours shaking in the dark was then carried out. After the second incubation the membrane was washed three times (for 10 minutes per wash) with 10ml blocking buffer. The membrane was then incubated in 10ml of blocking buffer containing 10µl of goat anti-rabbit antibody (which was coupled to horse-radish peroxidase), in the dark, shaking for one hour. The membrane was then washed five times (for ten minutes per time) with 10ml of blocking buffer. The bound peroxidase was localised by the addition of a 10mg diaminobenzamide (DAB) tablet in 20ml of 0.1M Tris pH 8 with 25µl 3% hydrogen peroxide. The reaction was stopped by rinsing the membrane thoroughly with

distilled water. All DAB-containing solutions and washes were collected for safe disposal. The immunoblots were photographed and stored in the dark.

Isoelectric Focusing (IEF)

Isoelectric focusing was carried out using the Hoefer "Tall Mighty Small" equipment. The following solutions were made up.

Gel Mixture (12ml: 1% agarose, 0.66M sorbitol, 0.5% ampholyte, 0.5% detergent)

IEF grade Agarose	(Gibco Brl)	0.12g
Sorbitol	(Sigma)	1.44g
Double distilled water		10.36ml
Ampholytes pH 5-8	(Sigma)	0.60ml
Triton X100	(Sigma)	0.60ml

Cathode Solution: 0.02 M NaOH (Sigma).

Anode Solution: 0.02M Phosphoric Acid (Sigma).

Fixative: 48% ethanol, 10% trichloroacetic acid (Sigma).

Stain Solution: 0.025% Coomassie R250, 40% methanol (Sigma).

Destain Solution: 40% methanol, 7% acetic acid (Sigma).

Sample buffer

glycerol	(BDH)	0.3ml
ampholytes	(Sigma)	0.1ml
bromophenol blue	(Sigma)	10µl

For a single gel 750µl of sample is normally mixed with 250µl of sample buffer. An overlay buffer is also prepared, 400µl distilled water, 100µl sample buffer.

A sheet of Gelbond film (Flowgen instruments), premarked into 1cm vertical strips each with subdivided into 30 2mm horizontal rows was adhered by its hydrophobic side to a glass gel plate by applying a small amount of water. The apparatus was assembled as for SDS-PAGE with in this case an alumina plate replacing the rear glass plate. The apparatus was then placed in a 60°C oven to preheat. The gel mixture was boiled (without ampholytes) then cooled to 65°C. Ampholytes were prepared by mixing pH 3-10 and 5-8 ampholytes from sigma in a ratio of 1:4. The ampholytes were then added to the gel mixture, the mixture was injected into the preheated apparatus using a 10ml syringe and once filled a comb was added and the gel left to solidify in a horizontal position for 1 hour. Once set, the comb was removed and the gel was inserted into the cooled running apparatus. The samples were prepared for running by the addition of sample buffer (600µl of glycerol, 150µl of the prepared ampholytes and 15µl bromophenol blue). The samples were injected into their wells and overlain with a 1:4 dilution of the sample buffer. The upper cathode reservoir was then filled with sodium hydroxide solution and the lower anode reservoir was filled with phosphoric acid. The gel was run for 30 minutes at 200V and a further hour at 600V.

Once the gel had completed its run it was removed (still bound to the gel bond) from all the apparatus and cut into the marked strips. The strips were then treated in several ways.

To stain the strips they were soaked for 15min in the fixative solution then washed in methanol. The strips of gel on their backing gelbond were then covered in pre-methanol soaked filter paper which was then covered with paper towel and finally a weighted glass plate. After 15 minutes the gel/gelbond was removed and dried with a hairdryer. The gel/gelbond was then stained in 0.05% Coomassie blue/ 40% methanol/ 10% acetic acid for 15 minutes. The stained gel was rinsed in 40% methanol/ 10% acetic acid until the background staining had faded and the stained bands were obvious. The gel/gelbond was then air-dried.

To measure the pH gradient produced across the gel during the run, the strips of gel were separated into thirty 2mm horizontal sections. Each section was then soaked in 200µl of 1M potassium chloride. After 2 hours incubation at room temperature the pH was measured using a pH electrode.

To localise the α-galactosidase activity on the IEF gel the strips of gel were washed with McIlvaines buffer and then separated into thirty 2mm horizontal sections. Each section was soaked in 300µl McIlvaines buffer. After soaking the buffer was diluted ten times and then used in a PNP-α-D-Gal assay.

Peptide mapping

In gel peptide digests were carried out to compare the primary structures of the maturing and germinating α-galactosidases (Rosenfeld et al 1992, Cleaveland et al 1977) To perform peptide mapping two stacking gels were poured on to 10% SDS gels and Preparative combs (one small well and one very large well) were inserted. 400µl of total sample buffer was prepared per gel (see the SDS PAGE section). 500µl of each crude protein preparation was added to a total sample buffer aliquot and then boiled for 3 minutes. Each 10% gel was loaded with a boiled sample in the large well and a 5µl aliquot of Dalton 7 (molecular weight markers) in the small well. The preparative SDS Page's gels were run at 25mA until the dye fronts had reached the bottom of the gels. The gels were removed from their supporting gels plates and lightly stained in 0.2% Coomasie brilliant blue, 20% methanol and 0.5% acetic acid for 20 minutes. After staining the gels were destained in 30% methanol for 30 minutes. Once destained the bands on the gels that corresponded to the α-galactosidases were excised.

Meanwhile a stacker was poured on an 18% SDS gel and a five well comb inserted. While the stacker was polymerising the stock protease was hydrated with deionised H_2O (according to the manufacturers instructions). The dilution gave a 1µg μl^{-1} solution in the case of GluC (cuts C-terminally at glutamic acid) and trypsin (cuts C-terminally at lysine and arginine) and a 0.1µg μl^{-1} solution in the case of LysC (cuts C-terminally at lysine). The working solution of the protease to be used was prepared by adding 8µl of stock protease to 144µl of Cleaveland/EDTA ($^1/_4$ X stacking gel buffer,

1mM ethlyenediaminetetraacetic acid (EDTA), Bromophenol blue to colour) and 144µl of Cleaveland/EDTA/20%Glycerol.

The 18% SDS gel was loaded. Three strips of each of the α-galactosidases excised from the 10% were loaded into wells. The wells were filled with Cleaveland/EDTA solution and 10µl of the working protease solution was layered under this and on top of the gel strips in the bottom of the wells. 5µl of Dalton 7 marker were layered into the standards lane and the running buffer poured very carefully into the equipment so as not to disturb the layerings. The gel was run at 25mA until the dye front was 1-2mm into the separating gel. The power was then turned off and the gel incubated for 1 hour. After the incubation the running was resumed until the dye front was 1 cm from the bottom of the gel. The gel was removed from the glass supports and either coomasie stained (as previously described) or blotted onto Problot membrane for peptide sequencing.

Blotting was carried out by initially wetting the Problot membrane with HPLC grade methanol for 5 seconds, then both the membrane and gel was soaked in CAPS buffer for 10 seconds before the blotting apparatus was assembled and run in the same manner as the previously described western blots.

After blotting the membrane was removed from the apparatus and stained by swirling it in in 0.1% Coomasie Brilliant Blue, 40% HPLC methanol and 1% acetic acid for 1 minute. Excess stain was removed by washing the membrane in 50% HPLC methanol which was it's self removed by washing in deionised H_2O. The membrane was

allowed to air dry, during which time the protein bands became visible and was then packaged and sent to Unilever for N-terminus amino acid sequencing.

mRNA Purification methods.

All solutions other than Sigma molecular biology chemicals were diethyl pyrocarbonate (DEPC) treated to remove any potentially contaminating RNAses.

The senna seeds (maturing or germinating) were dissected and the endosperms (or endosperms + testas) were frozen in liquid nitrogen. The frozen tissue was then ground to a flour in a ice cold mortar and pestle. The flour was used immediately in RNA purification methods.

Method 1.

unmodified from López-Gómez and Gómez-Lim (1992)

Solutions:

Lysis buffer: 150mM Tris Borate pH 7.5, 2% SDS, 50mM EDTA, 1% freshly added Mercaptoethanol. Sigma Molecular Biology

Molecular biology grade ethanol. Sigma Molecular Biology

5M potassium acetate. Sigma Molecular Biology

Chloroform/isoamyl alcohol (IAA) (ratio 24:1) Sigma Molecular Biology

12M lithium chloride Sigma Molecular Biology

To 40 endosperms worth of frozen tissue flour, 5ml of lysis buffer, 1.25ml of ethanol and 0.55ml of potassium acetate were added. The mix was then vortexed for 1 minute. 6.8ml of chloroform/IAA was added and then vortexed for a further minute. The tube was then centrifuged at 4000rpm for 10 minutes at 4°C. To the upper phase

of the tube an equal volume of phenol/chloroform/IAA was added (ratio 25:24:1) and the tube was vortexed for 1 minute then spun at 4000rpm, 4°C for 10 minutes. To the resulting upper phase an equal volume of chloroform/IAA was added, vortexed again for 1 minute and spun again as before. To the final upper phase $^1/_3$ rd of a volume of lithium chloride was added, mixed and incubated at -20°C overnight.

The next day the mixture was decanted into microcentrifuge tubes and centrifuged at 20,000g, 4°C for 90 minutes. The pellets formed were washed with ice cold 70% ethanol and respun at the top speed in a bench centrifuge for 20 minutes. The pellets in the microcentrifuge tubes were pooled in one tube and respun as before. The supernatant was removed and the pellet dried for one hour. The RNA pellet was resuspended in DEPC H_2O; the concentration measured spectrophotometrically and then stored at –70°C

Qiagen kit preparation of RNA

All solutions used were provided in the Qiagen kit.

60mg of endosperm flour was added to the Qiagen provided lysis buffer (either buffer RLT or RLC). The mixes were vortexed for 2 minutes. The lysed solutions were centrifuged (for 2 minutes) through Qiagen shredder columns. 225µl of ethanol was added and mixed. The samples were spun through Qiagen RNeasy spin column. The liquid flow through was discarded and the RNeasy spin columns were washed with 700µl of RW1 solution. The flow through was again discarded and the columns washed twice with 500µl RPE solution, (with the flow through being discarded). The RNeasy spin columns were placed in fresh collection tubes and 30µl aliquots of depc H_2O were added to the tops of the columns. The columns were spun for one minute.

The RNA flow through was collected and its concentration measured spectrophotometrically. The RNA solution was stored at -70°C.

In an attempt to solve to problem of the lysed sample gelling and blocking the spin columns two modifications were made to the standard Qiagen preparation:

1. A. After the addition of the 225µl of Ethanol to the RLC and RLT shredded samples, the viscous gel was spun through a glass fibre filter to remove the precipitate. The liquid and gelatinous flow-through was then taken through the rest of the standard Qiagen method until the elution stage at which point the spin column was allowed to air dry to remove any traces of remaining ethanol. The elution was then carried out in accordance with the Qiagen method.
2. Ethanol was added to the RLC and RLT lysis buffers (to give a 20% solution) in the hope of keeping carbohydrates in solution. Again flow-through after the addition of the 225µl of ethanol was then taken through the rest of the standard Qiagen method until the elution stage at which point the spin column was allowed to air dry to remove any traces of remaining ethanol. The elution was again carried out in accordance with the Qiagen method.

In order to solve the problem of the Qiagen samples floating out of the wells of agarose gels two approaches were taken:

1. Two drops of glycerol were added to each 10µl of RNA solution and 2µl of sample buffer. The mix was vortexed, spun down and loaded onto a gel.
2. 100% Ethanol was added to the RNA solutions to give a final concentration of 70%. The mixes were then incubated at -70°C for 2 hour to allow the RNA to

precipitate. The RNA was then pelleted by spinning for 20 minutes at top speed in a bench-top centrifuge. The supernatant was removed and the pellet was washed with ice cold 70% ethanol. After a brief 5 minute spin the wash was removed and the pellet was air dried for 1 hour. The pellet was then resuspended in depc H_2O, mixed with sample buffer and loaded into an agarose gel.

Dynabead extraction of mRNA from lysed tissue.

The solutions used were those provided by the Dynabead kit and also those of the Lopez-Gomez method.

0.1g of tissue flour was added to 1ml of lysis buffer (150mM Tris borate pH 7.5, 2% SDS, 50mM EDTA, 1% freshly added mercaptoethanol). 0.25ml of ethanol and 0.1ml of 5M potassium acetate was then added. The tissue was then homogenised, transferred to microcentrifuge tubes and spun for 30 seconds to remove cellular debris. 0.25ml of Dynabeads was prepared by removing the buffer in which the beads were stored and washing with 1 X lysis/ binding buffer. The washed beads were then added to the lysed tissue, mixed and incubated at room temperature for 5 minutes for hybridisation to take place. The supernatant was removed and the beads washed with 3 X 500µl of the kits washing buffer (+LIDS) and then 3 X washing buffer (-LIDS). 20µl of elution solution was then added to the beads and incubated for 65°C for 2 minutes. The supernatant containing RNA was frozen and stored.

Synthesis of a solid phase Dynabead cDNA library.

The storage supernatant was removed from 100µl of Dynabeads and they were washed once with 50µl of 2X binding buffer. Low concentration RNA previously prepared by the Lopez-Gomez method was added to an equal volume (120µl) of 2X binding buffer, heated to 65°C for 2 minutes and then placed on ice. The RNA/binding buffer mix was then added to the prepared Dynabeads and incubated at room temperature for 5 minutes. After the incubation the supernatant was removed. The Dynabeads were washed twice with 100µl washing buffer (no LIDS) and then washed three times in 125µl of 1X RT buffer. A reverse transcription mix was made up as follows:

10µl 5X 1st strand buffer.

2.5µl 0.1M DTT.

10µl NTP's (2.5µl of each)

0.625µl RNAsin

26.25µl DEPC H_2O

The RT mix was added to the Dynabeads, mixed and then split into two. One aliquot was incubated at 65°C for 3 minutes and then placed on ice and the other aliquot was placed directly on ice without the heating step. 1.25µl Molonie murine leukaemia virus reverse transcriptase (MMLVRT) was added to each aliquot and they were both incubated on ice for 2 hours. After the incubation the RT mixes were removed and 50µl of elution solution was added to each aliquot before incubating at 95°C for 60 seconds. The elution solution was discarded and the Dynabeads were washed and then stored in TE solution.

Guanadine HCl method of RNA preparation

Unmodified from Logemann et al (1987).

0.25g of endosperms were ground in liquid Nitrogen and added to a 15ml falcon tube. Guanadine HCl buffer was added to the tissue flour to give a 15% w/v mixture. The mix was vortexed, split into 2 microcentrifuge tubes and centrifuged at top speed in a bench-top centrifuge in a cold room (4°C). The supernatant was removed and added to an equal volume of phenol/chloroform (1:1 ratio) and vortexed. This mix was spun at top speed at 4°C for 45 minutes and the upper phase was placed into fresh microcentrifuge tubes. A 0.7 volume aliquot of cold ethanol and a 0.2 volume aliquot of 1M acetic acid were added followed by mixing and incubating at -70°C for 1 hour. After the incubation the microcentrifuge tubes were spun at top speed at 4°C for 10 minutes and the supernatant was discarded. 1ml of 3M sodium acetate pH5.3 was added and each microcentrifuge tube was vortexed until the pellets were fragmented. The tubes were spun at top speed for 10 minutes and the supernatant was discarded. The pellets were then washed three times in 0.5ml of 3M NaAc pH5.3 and then once in 70% cold ethanol. The ethanol and the soft white polysaccharide pellets were removed leaving the hard RNA pellets in the microcentrifuge tubes, which were allowed to air dry. The pellets were then resuspended in 20µl of DEPC H_2O.

DNAse Treatment of RNA

To ensure the RNA prepared was free from contaminating DNA, the RNA was treated with DNAseI, an enzyme which degrades DNA.

Solutions:

500mM Tris pH 7.5	Sigma Molecular Biology
1M Magnesium chloride	Sigma Molecular Biology
DNAse 1 (232u/µl)	Gibco BRL
Phenol/chloroform/IAA (25:24:1)	Sigma Molecular Biology
3M Sodium acetate pH 4.8	Sigma Molecular Biology
70% ethanol	Sigma HPLC grade

To 5µg of RNA, 10µl of Tris pH7.5, 1ml MgCl, and 1µl DNAseI were added, the volume made up to 100ml with DEPC H_2O and incubated at 37°C for 15 minutes. One volume of phenol/chloroform/IAA pH 4.5 was added, the tube vortexed for 1 minute and spun for 2 minutes in a bench centrifuge. The upper aqueous phase was isolated and one volume of chloroform/ IAA was added, vortexed and spun. The top phase was then ethanol-precipitated on ice for 5 minutes with 0.1 volumes 3M sodium acetate pH4.8 and 2.5 volumes ethanol then spun for 10 minutes. The resulting very small pellet was washed with cold 70% ethanol, dried in air under pierced parafilm for 30 minutes then dissolved in 10µl DEPC H_2O ready for use in the RT-PCR reaction.

Preparation of first Strand cDNA for use in 3' RACE RT-PCR

The first strand of cDNA for use in 3' Rapid amplification of cDNA ends PCR (RACE-PCR Frohman et al 1988) was produced by annealing the oligo-dT-$R_1R_0T_7$ primer to the polyA tails of the mRNA present (primer sequences are shown in the appendix). Reverse transcription was then carried out using MMLV reverse transcriptase as described below. Figure 9 shows this procedure graphically.

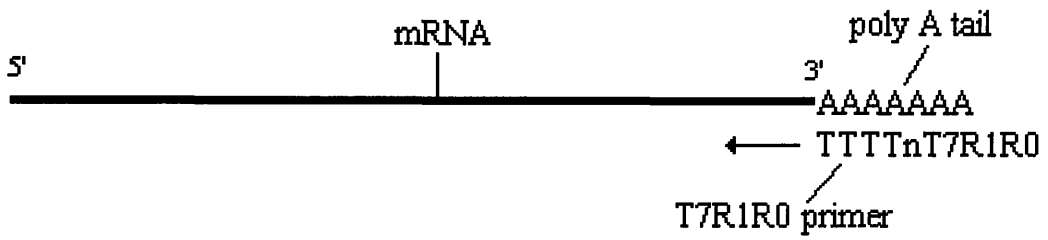


Figure 9. The arrangement of RNA and Primer and the direction of DNA polymerisation during reverse transcription for 3'RACE.

Solutions:

MMLVRT.	Gibco BRL
5X 1st strand buffer	Gibco BRL
Sterile 1 X TE buffer	Sigma molecular biology.
0.1M DTT	Sigma
dNTP's (at 10mM concentration)	Sigma molecular biology
RNasin	Gibco BRL

DT_{17} R_1R_0 primer (5pmol/µl)

The 10µl RNA sample was incubated at 65°C for 3 minutes then placed on ice. Mean while the reverse transcription reaction mixture was prepared (4µl 5X 1st strand buffer, 1µl 0.1M DTT, 4µl dNTP's (4X 1µl of each dNTP), 0.25µl RNasin, 0.5µl R_1R_0 primer). The RT mix was then added to the RNA and 1µl (200 units) of MMLV reverse transcriptase was added. After mixing, the reaction was incubated at 41°C for 2 hours. The volume was then made up to 500µl with 1X TE buffer and stored at -20°C.

Clean up of the first strand of cDNA.

Promega wizard method.

The Promega kit was used for this proceedure provided the solutions and equipment used.

A syringe barrel was attached to a mini-column. 1ml of the promega resin was added to the 500µl of first strand cDNA and mixed by inverting. The resin/first strand mix was added to the syringe barrel and pushed through using the plunger. The syringe was detached from the column, the plunger was removed and the barrel reattaced to the column before 2ml of 80% isopropanol was added to it. The isopropanol was pushed through the column using the plunger. The column was detached from the syringe, placed in an microcentrifuge tube and spun for 20 seconds to dry. To make sure the column was dry it was then placed under pierced parafilm for 15 minutes. When dry the column was placed in a new ependorf and 50µl of 70°C dH_2O was added. After a 1 minute wait the column was centrifuged for 20 seconds and the flow through collected and used in PCR or stored at -20°C.

RNA Degradation method.

$^{1}/_{25\ th}$ of a volume of 0.5M EDTA was added to the cDNA in TE buffer and then $^{1}/_{2}$ a volume of 150mM NaOH was added. The reaction was mixed and incubated at 65°C for 1 hour to allow the RNA degradation to reach completion. The reaction was neutralised by the addition of a $^{1}/_{3}$ rd of a volume of 1M Tris HCl pH8 and a $^{1}/_{3}$ rd of a volume of 1M HCl. The final cDNA solution was diluted 10 times by the addition of 9 volumes of 1X TE solution.

Polymerase chain reactions (PCR).

Standard PCR (Saiki et al 1985) and RACE-PCR was performed to isolate the α-galactosidase cDNA. The materials used were provided by Sigma (PCR buffer, dNTP's,) and Gibco (Taq polymerase, T_7 primer). Gene specific primers were custom synthesised by Unilever and are described in the appendix.

The PCR's carried out utilised either a gene specific primer (obtained from a "back translation" of the obtained amino acid sequences) and the T_7 primer (which would hybridise to a site on the oligo-dT-$R_1R_0T_7$ that had been incorporated into the first stand cDNA) or two gene specific primers. The arrangements of primers and cDNA are shown in figure 10

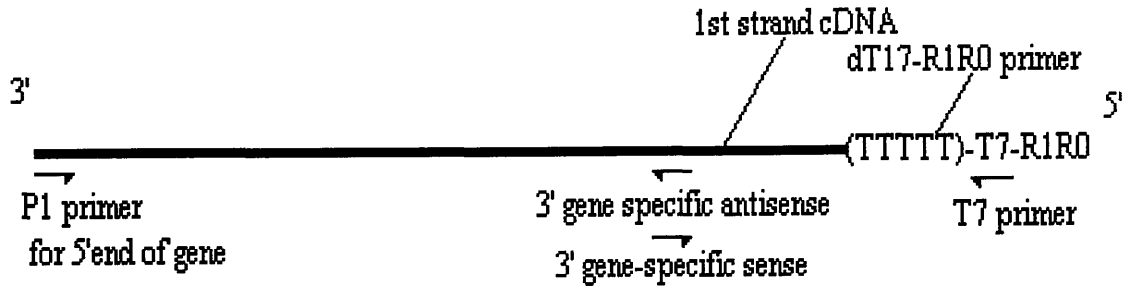


Figure 10. The arrangement of gene specific primers (P1, 3'GSP antisense and 3'GSP sense) and T7 in standard or 3' RACE PCRs.

A PCR cocktail was made up:

10X PCR buffer	5µl
dNTPs (1µl of each at 10mM)	4µl
T_7 primer (5pmol/µl)	5µl
Gene specific primer (5pmol/µl)	5µl
Template (from RT reaction)	10µl
MilliQ	20.5µl
	Total = 49.5µl

The template was denatured by incubating the cocktail (in the PCR Block) at 95°C for 7 minutes then cooling to 75°C. 2.5 units of Taq (0.5µl) was then added to the reaction mixture, mixed and spun down. 50µl of mineral oil was layered on the reaction mix to prevent evaporation and the reaction was thermocycled:-

1 cycle of:-

50°C (annealing)	2 minutes
72°C (cDNA extension)	10 minutes

then 30 cycles of:-

94°C (melting of strands)	1 minute
50°C	1 minute
72°C	1.5 minutes

then 1 cycle of:-

72°C	15 minutes

After the reaction was complete the PCR product was stored frozen (-20°C) after a 5µl aliquot was removed for running on an agarose gel.

Agarose gels.

1% Agarose gels were prepared to run samples of the PCR products. 200mg of agarose was added to 20ml 1xTBE buffer (to give 1% w/v) and melted in the microwave for 2 minutes on medium high setting then poured into a 50ml Falcon tube. 1µl ethidium bromide was added into the Falcon tube and gently inverted. The 20ml was then poured into the casting/running apparatus (assembled with a comb in place) and the gels left to set in the fume cupboard before being wrapped in cling film and stored in the fridge. If the gel was to be run for screening purposes then 5µl of the PCR products (with 1 µl sample buffer) were run on the gel, if the purpose of the gel

was to enable bands to be cut out to used further then 35µl samples were loaded. The standards used in the gels were 1kb ladder and HindIII standard. Gels were run at a constant voltage of 100V each for approximately 35 minutes in a running buffer of 1X TBE with 0.5% v/v ethidium bromide. After the run the gel was photographed on a transillumiator and any bands to be kept were excised using a scalpel.

Production of competent cells.

Solutions:

L-Broth (5g Luria broth base (Gibco) in 200ml of deionised H_2O, then autoclaved).

L-broth with supplements (5g Luria broth base (Gibco), 0.4g Maltose, 0.5g $MgSO_4.7H_2O$ (Sigma) in 200ml of deionised H_2O, then autoclaved).

L-Agar (7.4g Luria agar (Gibco) in 200ml of deionised H_2O, then autoclaved).

0.1M $MgCl_2$ (Sigma molecular biology grade).

0.1M $CaCl_2$ (Sigma molecular biology grade).

L-Agar plates were streaked with stock cells and were incubated at 37°C overnight. After the cells have grown into colonies, one was picked and grown up in 5ml of L-broth + supplements overnight at 37°C shaking at 200rpm in an orbital incubator. 50ml of L-broth was inoculated with 0.5ml of the overnight culture and incubated at 37°C at 200rpm until the $OD_{550} = 0.5$. The cells were then spun down at 4,000 rpm for 5 minutes and the resuspended in 50ml of ice cold 0.1M $MgCl_2$. The cells were again spun down at 4,000rpm for 5 minutes and then resuspended in ice cold 0.1M $CaCl_2$ and incubated on ice for 30 minutes. The cells were spun for a final time at 4,000rpm for 5 minutes and resuspended in 5ml 0.1M $CaCl_2$ 15% glycerol before being split into 50µl aliquots and frozen at -70°C.

Ligations and transformations.

The bands on agarose gels corresponding to the PCR products to be cloned were excised and stored at -20°C. When required the bands were thawed and spun down.

Novogen Cloning.

A, ligation.

A reaction mix was prepared:-

T vector	1µl
T_4 DNA ligase	3µl
Ligation buffer	2µl
PCR product from gel	4µl

The reaction mixture was incubated at room temperature overnight.

B, transformation.

2µl of the ligation mixture from the previous day was added to 20µl of freshly thawed Novablue competent cells, mixed and incubated on ice for 30 minutes. The cells were heat-shocked at 42°C for 40 seconds and then returned to ice for 2 minutes. 80µl of SOC buffer was added to the competent cells and they were then incubated at 37°C for 1 hour in an orbital shaker. 50µl aliquots of cells were spread on L. agar plates containing:-

per 200ml L. Agar

IPTG (0.5M)	200µl
X-gal (80µg/ml)	40µl

Ampicillin (50µg/ml) 250µl

Tetracycline (50µg/ml)240µl

The plates were allowed to dry and were incubated at 37°C overnight.

pBluescript Blunt ended Cloning.

Blunt ended cloning was based on the Stratagene pCR-script SK(+) cloning method. pBluescript vector was prepared for blunt-ended ligations by an initial restriction digest:-

10µl	Stock pBluescript	
2µl	One-Phor-All buffer or React-4 buffer	Gibco BRL
2µl	SMA1	Gibco BRL
6µl	MilliQ	

The reaction mixture was incubated over-night at 37°C. The product of the digestion was run on an agarose gel and excised. The gel slice was cleaned up using the hybaid PCR clean up kit (as described below) as used in the ligation reactions.

PCR products were also excised from gels then immediately cleaned-up using the Hybaid PCR clean-up kit.

A, Hybaid PCR clean-up.

The bands of gel were placed in a spin column (no more than 300mg/column) and 400µl of resuspended binding buffer was added. The column was incubated at 55°C for 5 minutes and flicked to mix. The column was spun for 30 seconds and the flow-through discarded. 2X washes of 500µl wash solution were added to the spin column

and each spun for 50 seconds, the flow-through again being discarded. The column was spun for 1 minute to dry it and then placed in a new collection microcentrifuge tube. 25µl of elution solution was added to the column, it was flicked to mix and then spun down for 30 seconds. The PCR product was collected in the flow-through.

B, Ligation.

The ligation was set up using the following mix:-

4µl PCR product from hybaid clean-up.

1µl one-phor-all buffer.	Gibco BRL
1µl rATP.	Gibco BRL
3µl ligase.	Gibco BRL
1µl SMA1 (to open up self-ligated pBluescript).	Gibco BRL

1µl pBluescript from clean-up.

Also set up were a positive control (with no SMA1 to give self ligated bluescripts) and a negative control (with no PCR product).

The ligations were incubated at room temperature overnight.

C, Transformations.

Each transformation mix consisted of 20µl of competent XL1-Blue and 2µl ligation mix from the previous day. The mix was incubated on ice for 30 minutes, heat shocked at 42°C for 50 seconds and returned to ice for 2 minutes. 80 µl SOC medium (2.0% tryptone, 0.5% yeast extract, 10.0mM NaCl, 2.5mM KCl, 10.0mM $MgCl_2$, 20.0mM glucose) was added and the transformation mix was incubated shaking at 37°C for 1 hour. The transformation mix was spread on Ampicillin plates and grown up at 37°C overnight.

Invitrogen Cloning.

A, Ligation.

As before the PCR products were excised from gels and were immediately cleaned-up using the Hybaid PCR clean-up kit.

The PCR product was immediately used in the ligation reaction after adjusting its concentration so that 1µl contained X ng of PCR calculated using the formula:

$$X = \text{size of product in bps} \times 50 \text{ (the mass of 1 base pair in ng)} \Big/ \text{Vector size in bps}$$

The ligation mix was set up as follows:

10X ligation buffer	1µl
PCR II vector	2µl
Ligase	1µl
MilliQ	5µl
PCR product	1µl

The ligation was incubated at 14°C overnight.

If the PCR product was more than two days old before the T/A ligation and transformation reaction was carried out then it was important to ensure that the Adenosine overhangs were in place. This was carried out by mixing the original PCR reaction mix with 0.5µl fresh Taq and 1µl 10mM dATP. The reaction was incubated at 72°C for 10 minutes and then loaded onto an agarose gel and run and cleaned up as before.

B, Transformation.

50µl aliquots Invitrogen "one shot cells" (INVαF') were thawed on ice. 2µl of 0.5M β-mercaptoethanol was added and stirred using a pipette tip (INVαF' cells should not be vortexed). The ligation mix from the previous day was pulsed down (briefly centrifuged) and 2µl added to the one shot cells. After stirring the cells were incubated on ice for 30 minutes. The cells were then heat-shocked for 30 seconds at 42°C and then placed on ice for 2 minutes. 450µl of SOC medium was added to the cells and they were incubated horizontally at 37°C for 1 hour at 200rpm in an orbital shaker. After the incubation aliquots of 50µl and 200µl of cells were spread on AMP-XGAL plates (200ml of L-Agar had 250µl Ampicillin (40µg/µl) and 64µl XGAL (250µg/µl) added). Once the plates had dried they were incubated at 37°C overnight.

Promega Wizard Miniprep of plasmid DNA.

For each miniprep a colony was picked from the transformation plates. The colony was grown up shaking in 5ml of LB media containing 6.25µl of Ampicillin (at 40µg/µl) at 37°C overnight. 5ml of the bacterial culture was spun down at 5,000g for 5 minutes, the supernatant was poured off and the tubes were blotted upside-down on a paper towel. The bacterial cells were resuspended by vortexing the pellet in 250µl of Wizard *Plus-SV* cell resuspension solution and then transferred into an microcentrifuge tube. 250µl of Wizard *Plus-SV* cell lysis solution was added to each miniprep and the microcentrifuge tubes were mixed by inverting four times (to prevent shearing of chromosomal DNA). The cells were left to lyse for no more than 5 minutes before 10µl of alkaline protease solution was added. The cells were then incubated for 5 minutes. 350µl of Wizard *Plus-SV* neutralisation solution was added to each microcentrifuge tube and mixed by four inversions. The lysed cells were then spun at top speed in a bench-top centrifuge for 10 minutes. The cleared lysate was transferred to the Wizard *Plus-SV* miniprep spin column by decanting (avoiding transfering any of the white genomic precipitate). The spin column was then centrifuged at top speed for 1 minute. The flow through was discarded. 750µl of Wizard *Plus-SV* column wash solution was added and spun through the spin column as before. The flow through was discarded and a second wash of 250µl of column wash solution was performed. The spin column was transferred to a new, sterile microcentrifuge tube and the plasmid DNA was eluted from the column by the

addition of 100µl of nuclease-free water and a 1 minute, top speed centrifugation. The DNA was stored at -20°C.

Hybaid Miniprep of plasmid DNA.

For each miniprep a colony was picked from the transformation plates. The colony was grown up shaking in 3ml of LB media containing 3.75µl of ampicillin (at 40µg/µl) at 37°C overnight. 1.5ml of the bacterial culture was then spun down for 30 seconds and the media was poured off. The pellet of bacterial cells was resuspended in 50µl of Hybaid pre-lysis buffer. Lysis was carried out by the addition of 100µl of the supplied lysis solution followed by pipetting up and down until the cell suspension became clear and viscous. 75µl of neutralising solution was then added and the mix was mixed by vortexing. The resulting precipitate was removed by centrifuging for 2 minutes and transferring the supernatant to a Hybaid spin filter. The supplied binding buffer was shaken to resuspend the silica gel matrix and 250µl was added to the spin filter. This was mixed with the product of the cell lysis (already in the filter) by pipetting up and down. The spin filter was spun for 1 minute to collect the liquid in the bottom of the vial. 350µl of wash solution was added to the spin filter and spun for 1 minute. The collection vial was emptied and the spin filter spun again for 1 minute to dry it. The spin filter was then transferred into a new collection vial, 50µl of sterile MilliQ was added and the spinfilter vortexed briefly. The filter was spun for 30 seconds. An aliquot of the DNA collected in the collection vial was restriction digested and run on an agarose gel; the rest was stored at -20°C.

Restriction Digests.

Restriction digests were used to excise the inserts from the isolated ligated plasmids before running the ligations on agarose gels. The digests were set up as follows:

3µl minipreped plamid.

1µl one-phor-all buffer

1µl of each restriction enzyme required for insert excision (consultation of the plasmid map was required at this point).

Xµl of sterile MilliQ to give 10µl total volume.

The reactions were incubated for 2 hours at 37°C.

Manual Sequencing using the Pharmacia kit.

Preparation of the apparatus.

The sequencing gel pouring apparatus (anachem) was cooled at 4°C for 30 minutes. 100ml of Easigel (Scotlab), 500µl of 10% freshly made ammonium persulphate and 100 µl of TEMED were mixed and quickly poured into the gel apparatus. After the gel had been allowed to polymerise for 1 hour excess gel material was removed from its ends, combs were positioned and it was placed in the running apparatus with exposed gel temporarily covered with Clingfilm.

Preparation of the template.

The template to be sequenced was prepared next. 1µl of it was diluted in 99µl of deionised H_2O and its absorbance at 260nm was measured. From this information the concentration of the template can be calculated (A_{260} X 50 = µg ml^{-1}). The template was then diluted with deionised H_2O to produce a 32µl solution containing 2µg of template.

Annealing of the primer to the double stranded template.

The 32µl of template was denatured by incubating with 8µl of 2M NaOH at room temperature for 20 minutes. After the incubation 7µl of 3M sodium acetate (pH4.8), 4µl of distilled water and 120µl 100% ethanol were added. The mix was precipitated at -70°C for 1 hour. The precipitated denatured template was spun down by centrifuging at top speed on a bench-top centrifuge for 30 minutes. The supernatant was removed and the pellet washed gently with ice-cold 70% ethanol. After a further spin of 10 minutes the ethanol wash was removed, the pellet dried in a desicator under vacuum for 10 minutes and then resuspended in 10µl of distilled water. 2µl of the desired primer (at a concentration of 5pmol μl^{-1}) and 2 µl of annealing buffer was added to the 10µl of template and the mix was incubated at 65°C for 5 minutes. The mix was then incubated at 37°C for 10 minutes and then placed at room temperature for 5 minutes. The annealed primer-template was then used immediately in the sequencing reactions.

Sequencing reactions.

Into each of four microcentrifuge tubes (marked A,C,G and T) 2.5µl of the corresponding "Read Short" mixes were pipetted. The supplied T7 DNA polymerase was diluted with it's dilution buffer. Each template to be sequenced required 0.5µl of polymerase to be diluted with 2µl of buffer. The labelling reactions were then performed. The labelling mix was prepared by adding, 3µl of supplied labelling mix, 1µl α-^{35}S dATP and 2µl diluted T7 DNA polymerase To the 14µl of annealed primer-template. The reaction was mixed by gentle pipetting and briefly spun down before incubating at room temperature for 5 minutes. Meanwhile the four marked microfuge tubes were warmed to 37°C in a water bath. After the 5 minute labelling incubation was complete 4.5µl of the mix was transferred to each of the warmed "Read Short"microfuge tubes, mixed by pipetting and incubated at 37°C for 5 minutes. The reaction was stopped by the addition of 5µl of stop solution into each microfuge tube followed by a gentle mixing and a quick centrifugation.

Running the gel.

The sequencing reaction were prepared for running on the gel by taking 3µl aliquots from each of the four microfuge tubes (A, C, G and T), heating them to 80°C for 2 minutes and the immediately loading them onto the sequencing gel. The sequencing gel was run at 2000V for 2 to 4 hours. After running the gel was removed from the glass plates and dried onto 3M paper. The gel was placed into contact with x-ray film and stored in an x-ray cassette whilst the film was exposed (which could take from overnight to a week depending on the "freshness" of the α-^{35}S dATP). After exposure the x-ray film was developed and the sequence read.

Genomic DNA preparation.

5.34g of young senna leaves were harvested and placed in an autoclaved foil packet which was then plunged into liquid nitrogen. The frozen leaves were ground in a pre-cooled mortar and pestle to a fine powder. The leaf flour was added to 20ml of extraction buffer into which 1.2ml of 20% SDS was added, after mixing this was incubated at 65°C for 10 minutes. After the incubation 5ml potassium acetate was added, the tube was shaken gently and placed on ice for 20 minutes. The tube was spun at 5,000 rpm for 20 minutes. The supernatant was poured through an autoclaved glass-wool filter. 10ml of isopropanol was added to the supernatant and gently mixed. Using a flamed pasteur pipette the precipitated DNA was spooled out and placed into a clean tube. 3ml of TE buffer (10mM TrisHCl pH8, 1mM EDTA)was used to dissolve the DNA and 20µl RNase (1723U/µl Gibco BRL) was added and incubated at 37°C for 30 minutes. After the incubation 0.5ml of phenol and 0.5ml of chloroform was added mixed and centrifuged at 5,000rpm for 15 minutes. The upper aqueous phase was collected and 3ml chloroform was added mixed and spun as before. The upper phase was again collected and 0.8 volumes of isopropanol and 0.2 volumes of 3M sodium acetate were added and gently mixed. The DNA was again spooled out with a flamed pasteur pipette. The DNA was washed on the pipette with 1ml 70% ethanol. The residual ethanol was squeezed out and air dried. The DNA was then dissolved in TE buffer (1ml per 5g of tissue). The DNA was stored frozen at -20°C.

Southern Blotting.

Southern blotting was carried out to attempt to discover if there was one or more than one α-galactosidase gene present in the senna genome. This was performed by probing the genome with a highly conserved PCR product (PCR product P1GA1) which would localise any α-galactosidase like genes and then probing with a less conserved PCR product (PCR product GS1T7) which would only highlight the germinating α-galactosidase gene. If only one gene was localised in both probes then it would be possible to state that there is only one α-galactosidase like gene in the senna genome.

A, Genomic DNA Digest.

Three genomic digestion mix was set up:-

Genomic DNA	20µl (8µg)
One-phor-all buffer	3µl
High concentration Restriction enzyme (EcoRI, HindIII, BamHI)	3µl
MilliQ	4µl

The digest was incubated at 37°C for 24 hours.

B, Agarose gel.

A 0.8% Agarose midi size gel was prepared and loaded with all three digests. The gel was then run at 20V for 18 hours. After running the gel was photographed on a transilluminator.

C, Blotting

The run agarose gel was placed face down in a tray with 500ml 0.25M HCl and incubated shaking for 15 minutes. The blotting trays and glass support was assembled. 3 sheets of Whatman's 3M paper was placed on the support and wetted with freshly made-up 0.4M NaOH (any bubbles were removed). The blotting tray was filled with 600ml of the 0.4M NaOH and the agarose was then placed face down on the 3M paper (and again any air bubbles were removed). Hybaid nylon N+ membrane pre-wetted in sterile deionised H_2O was then placed on top of the gel and bubbles were removed. The wicks of 3M paper were covered with clingfilm. One sheet of 3M paper soaked in 0.4M NaOH was placed on the N+ membrane and 2 dry sheets were placed on this. A wad of tissues were placed on the 3M paper, a glass plate on the wad and a 1kg weight on the glass plate. The southern blotting equipment is shown in figure 11.

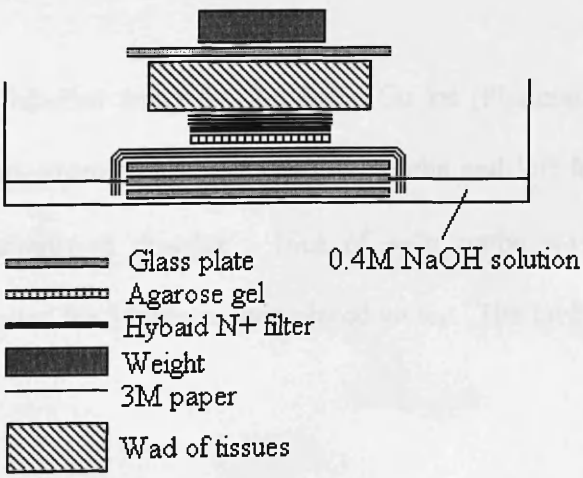


Figure 11. The arrangement of equipment in a southern blot.

After assembling the apparatus it was left overnight at room temperature to allow transfer to take place.

The next day a photograph of the gel was taken under U.V. irradiation and compared to the gel before blotting so as to assess the efficiency of transfer. The N+ membrane was washed in 2X SSC for 2-3 minutes, blotted on 3M paper and stored at -20°C in a plastic bag.

Probe preparation.

The probes used in the southern blotting were PCR products previously cloned. 200ng of oligonucleotide was required for the preparation of each probe and so the amount of plasmid required to give this amount after restriction digest was calculated and then doubled to take into account losses. Standard restriction digests were then carried out. When complete the digests were run on an agarose gel, the fragments to be used for probing were excised and cleaned up using the Hybaid clean up kit. The concentration of the probes were measured and adjusted to 200ng/10µl with sterile deionised H_2O.

The probes were labelled using the Ready-To-Go kit (Pharmacia biotec). 20µl of deionised H_2O was added to each Ready-To-Go tube and left for 5 minutes for the Ready-To-Go reactants to dissolve. 10µl of each probe was added to 15µl of deionised H_2O, boiled for 3 minutes then placed on ice. The probe labelling reactions were set up:-

20µl of the Ready-To Go solution, (Pharmacia)

25µl boiled DNA solutions, (Sigma)

5µl ^{32}P-ATP (3000µCi/mol). (NEN)

The reactions were then incubated at 37°C for 30 minutes. After the incubation 10µl of 0.1M EDTA was added to each reaction and the labelled probes were cleaned up using the hybaid clean-up kit.

Prehybridisation

Prehybridisation buffer was prepared from stock solutions:-

Stock Solution	Amount used	Final Concentration
20X SSC Buffer	12ml	6X
10X SDS	2ml	0.5%
1M NaH_2PO_4	0.8ml	0.02M
50X Denhardt's solution	4ml	5X
Deionised H_2O	20.8ml	

The prehybridisation buffer was heated to 60°C and 0.4ml of boiled sonicated salmon sperm DNA was added. Two hybridisation tubes were heated to 55°C and 20ml of the hybridisation added to each tube and they were kept at 55°C for 30 minutes. The blotted membranes were then warmed and placed, one in each tube (with the blotted face pointed inwards). The membranes were then incubated at 55°C for 4 hours.

Hybridisation.

The prepared probes were boiled for 3 minutes in order to denature them. The probes were then placed on ice. The hybridisation tubes were opened and the probe solutions were added directly to the prehybridisation buffers (avoiding direct contact with the membranes). The hybridisations were incubated at 55°C overnight.

Washing and Imaging.

400ml of wash solution was prepared (2X SSC, 0.1% SDS) and heated to 55°C.

The radioactive hybridisation solution was poured off and 100ml of wash solution was added to each tube. The tubes were placed back into the hybridisation oven and incubated for 10 minutes. The first wash was poured off and a second wash performed. The final wash solution was poured off and the membranes dried on 3M paper. The membranes were then placed in saranwrap and placed in an intensifier cassette with x-ray film and stored at -70°C until the film had been exposed. The x-ray film was developed and imaged.

Inverse PCR.

Inverse PCR was carried out using a method based on the work of Ochman et al 1988 and Ochman et al 1990. Senna genomic DNA was prepared using the previous method. The DNA was then digested as described below:-

25µl (1µg)	Senna DNA
3µl	IPA buffer
1µl	restriction enzyme (TaqI, MboI or EcoRI)
1µl	deionised H_2O

The digest was incubated at 37°C overnight.

After the incubation the digest was heated to 65°C (to inactivate the restriction enzymes) and then vortexed with an equal volume of phenol/chloroform (1:1). After centrifugation the DNA containing aqueous phase was removed and vortexed with an equal volume of chloroform. After another centrifugation the aqueous phase was mixed with 0.8 volumes of isopropanol and 0.2 volumes of 3M NaAc pH5.2 and incubated at -20°C for 1 hour. The sample was centrifuged for 15 minutes at top speed in a bench top centrifuge, the supernatant was removed and the pellet washed in ice cold 70% ethanol. The wash was removed and the pellet dried. The DNA was resuspended in 20µl of deionised H_2O.

The digested Genomic DNA was self ligated at 14°C overnight using the mix below.

10µl	Senna genomic DNA.
20µl	5X ligation buffer.
5µl	Ligase.
65µl	Deionised H_2O.

After the ligation reaction the resulting self ligated (circularised) genomic DNA fragments were used in standard PCR reactions. The primers used in these reactions (GA2 and GS3) were designed to be close to the 5' UTR by the use of the known sequence of the senna α-galactosidase. The primers were designed to point in opposite directions, meaning that they would only produce a PCR product from the α-galactosidase gene if self ligation had taken place; this is shown in figure 12.

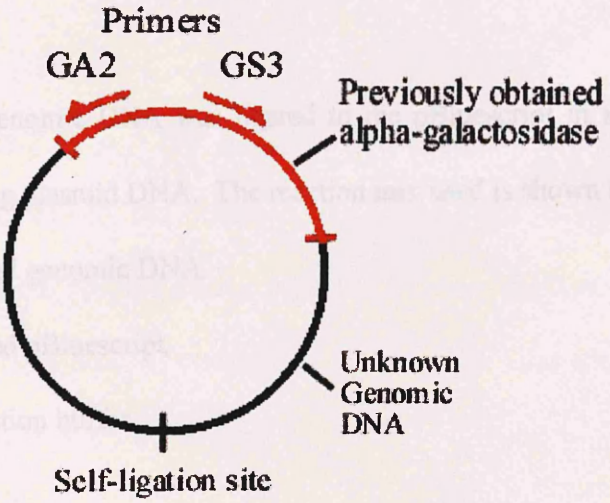


Figure 12 Inverse PCR. The diagram shows the previously obtained α-galactosidase sequence (in red),the GA2 and GS3 primer sites and directions and the unsequenced genomic DNA (incorporating the 5' UTR upstream from the previously sequenced region).

Single sided PCR.

Single sided PCR was carried out by initially isolating senna genomic DNA as before. The DNA was digested using EcoRI, HindIII and BamHI. pBluescript was digested using the same restriction enzymes as the genomic DNA and then all the digests were heated to 65°C for 15 minutes to inactivate the restriction enzymes and purified using the Hybaid purification kit.

The digested senna genomic DNA was ligated to the pBluescript in a ratio of 1μg genomic DNA to 1/5μg plasmid DNA. The reaction mix used is shown below:-

0.5μg	Digested genomic DNA.
0.1μg	Digested pBluescript.
6μl	5X ligation buffer.
1.5μl	Ligase.

Deionised H_2O purified to bring the final volume to 30μl.

The ligation reactions were carried out at 14°C overnight.

After the ligation standard PCRs were carried out using the plasmid/genomic DNA constructs as templates. The primers used in these reactions (T7 and KS on the vector and GA1 and GA3 on the α-galactosidase gene) were designed to be close to the 5' UTR and to be used in nested PCRs (thus increasing the stringency). A diagrammatic representation of the single sided PCR is shown in figure13.

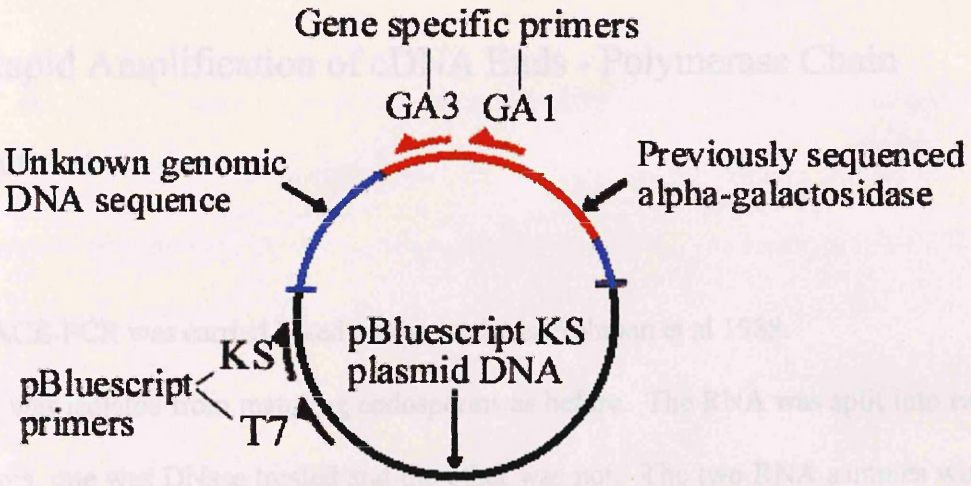


Figure 13. Single sided PCR. The diagram shows the previously obtained α-galactosidase sequence (in red), the unsequenced genomic DNA (incorporating the 5' UTR upstream region) and the pBluescript plasmid. Primers GA1 and T7 were used for the first round of PCR and the products of this reaction were used in a second round of PCR with the primers GA3 and KS.

5' Rapid Amplification of cDNA Ends - Polymerase Chain Reactions.

5' RACE-PCR was carried based on the methods Frohman et al 1988.

RNA was isolated from maturing endosperms as before. The RNA was split into two aliquots, one was DNase treated and the other was not. The two RNA samples were used to produce 1st strand cDNA using either random hexamer primers or an α-galactosidase gene specific primer (GA1) using the reaction mix:-

4µl	first strand buffer	(Gibco BRL)
1µl	DTT	(Gibco BRL)
4µl	dNTPs (2.5mM of each)	(Sigma)
0.25µl	RNasin	(Gibco BRL)
0.5µl	Primer (hexamer or gene specific)	(Gibco BRL or Unilever)
5ng in 10µl	Prepared RNA	

The RNA was heated at 65°C for 3 minutes before being added to the rest of the reaction mix which was then incubated at 42°C for 2 hours.

The 1st strands were cleaned up using the Hybaid clean-up kit and then poly-A tailed using the TdT enzyme using the reaction:-

20µl	cDNA first strand.	
8µl	Tailing buffer	(Gibco BRL)
4µl	ATP	(Sigma)
2µl	TdT enzyme	(Gibco BRL)
6µl	Deionised H_2O.	

A 2nd strand of cDNA was produced using the oligo-dT-RiRoT7 primer as before. PCRs were then carried out using primers Ro and GA1 for the random primed cDNA and primers Ro and GA3 for the gene specific primed cDNA the result of this PCR is shown in figure 58.

Results.

The identification of the Mr of *senna* α-galactosidases.

After the preparation of a crude extract of enzymes from germinating *senna* endosperms was serially diluted, two SDS PAGE gels were run. The first gel was coomassie stained to give an estimate of the amount and relative proportions of soluble proteins present in germinating senna seeds.

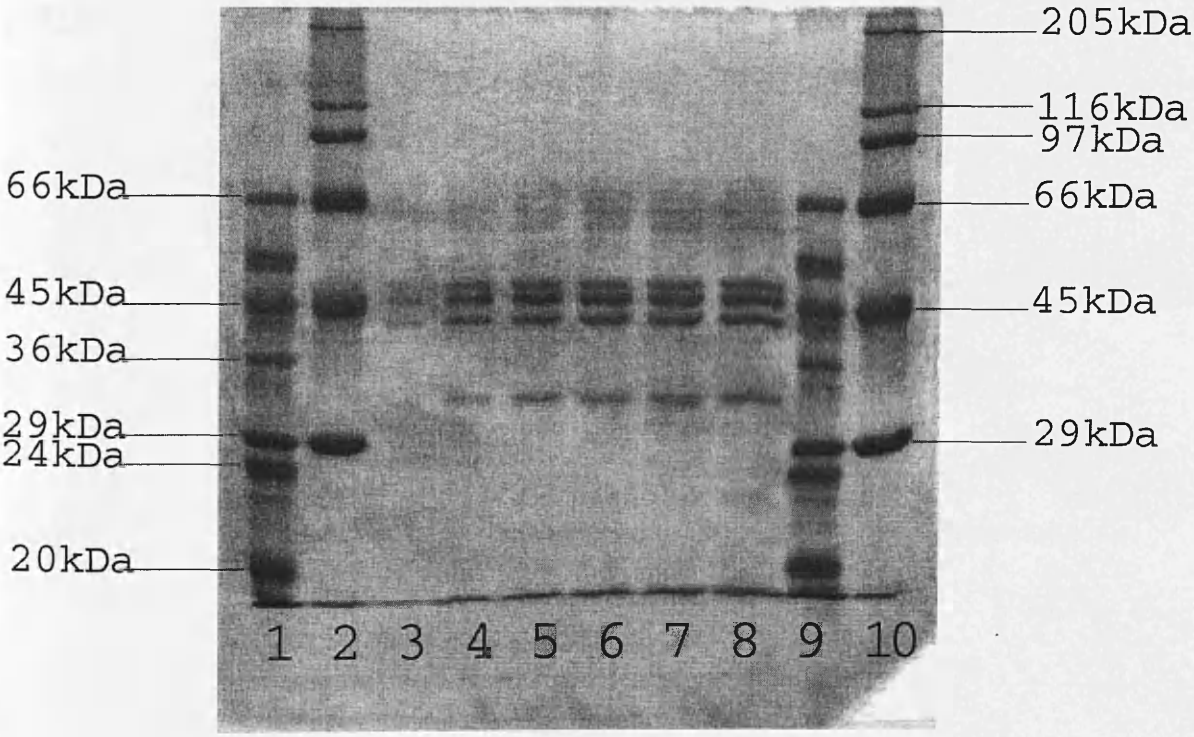


Figure 14. The proteins present in germinating senna seeds. As with all photographs lanes are numbered from left to right. Lanes 1 and 9 contained Dalton 7 markers, lanes 2 and 10 High Dalton markers (Appendix B). Lanes 3-8 contained a series of dilutions (5μl to 30μl per well) of a crude protein preparation from germinating senna seeds.

Figure 14 shows that there are few soluble proteins present during germination with three similarly sized ones being prominent.

The second SDS gel was western blotted as described in the methods section. After blotting, the lanes containing the high and low Dalton molecular weight markers were cut from the blot. The marker lanes were amido black stained to visualise all the protein bands present and the rest of the blot was probed with the anti Guar α-galactosidase antibody to locate the α-galactosidase present. The result is shown in figure 15

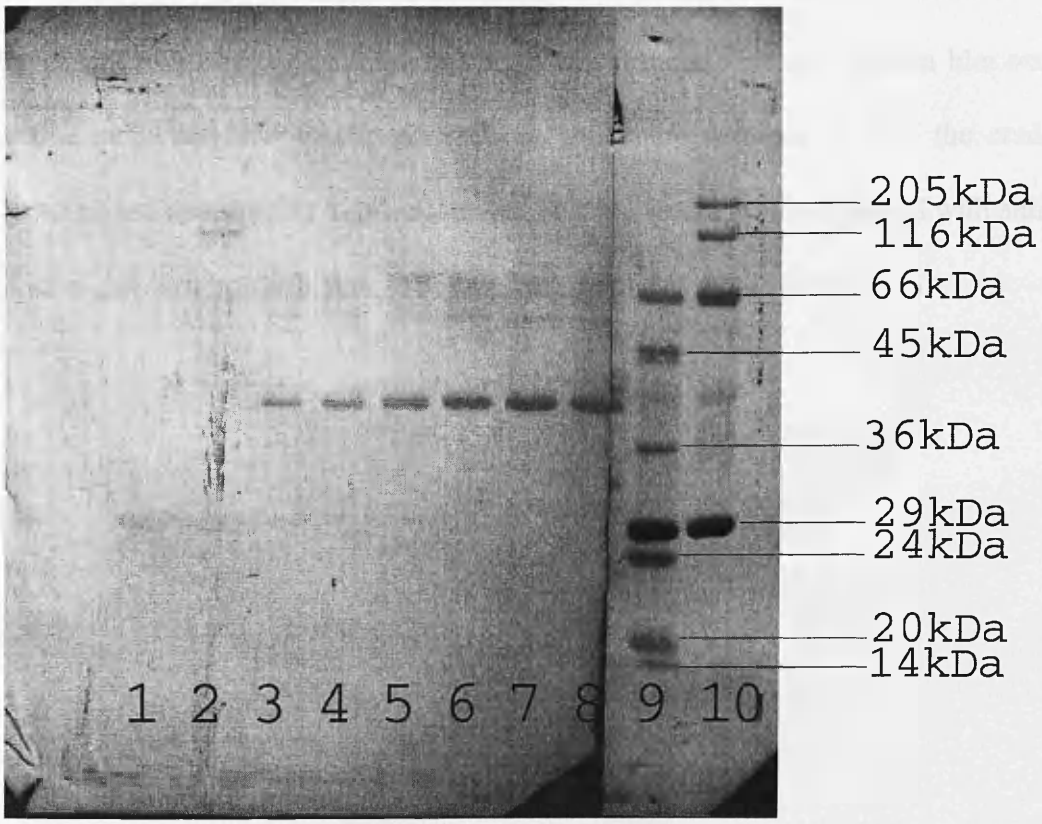


Figure 15 The localisation of α-galactosidase on a western blot. Lane 1 contained Dalton 7 markers, lane 2 High Dalton markers. Lanes 3-8 contained a series of dilutions (5μl to 30μl per well) of a crude protein preparation from germinating senna seeds. Some of lane 8 and Lanes 9 and 10 were amido black stained. Lane 9 contained Dalton 7 and lane 10 contained high Dalton markers.

Figure 15 shows that germinating senna α-galactosidase has an approximate mass of 40 kDa and is the smallest of the three most prominent proteins found in the germinating senna seeds.

In order to identify which protein in the crude preparation of maturing senna seeds corresponded to the α-galactosidase activity found there, a further western blot was carried out. An SDS PAGE gel was run containing dilutions of both the crude maturing and germinating α-galactosidase. This gel was blotted and probed with anti-Guar α-galactosidase as before. The result of this blot is shown in figure 16

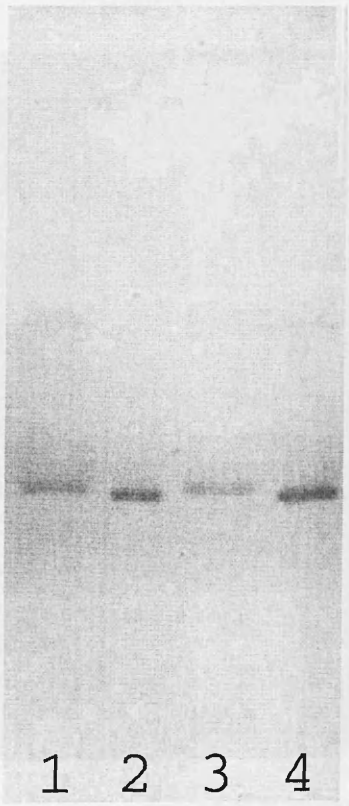


Figure 16 A western blot showing the distinction in size of the two forms of α-galactosidase. Lane 1 contains 20μl of maturing α-galactosidase, lane 2 10μl of germinating α-galactosidase, lane 3 30μl of maturing α-galactosidase and lane 4 contains 30μl of germinating α-galactosidase.

The maturing α-galactosidase is obviously less mobile through the gel than the germinating enzyme indicating that it has a greater Mr, however an accurate estimation of both Mr's can not be achieved on a western blot because molecular weight markers cannot be observed on the blot. For the final accurate measurement of both Mr's an SDS gel containing both enzyme preparations along with molecular weight markers was run and then coomassie blue stained as shown in figure 17

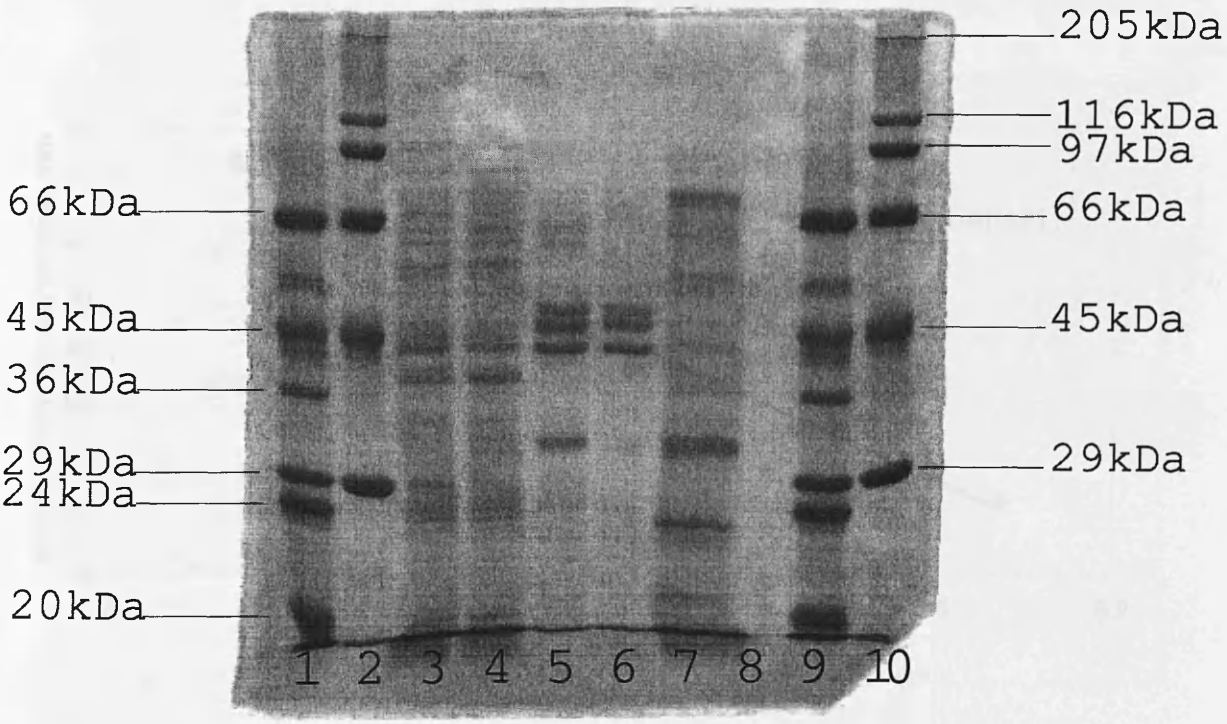


Figure 17. A coomassie blue stained SDS PAGE gel that enables an accurate estimation of the sizes of the maturing and germinating α-galactosidases to be made. Lanes 1 and 2 contain Dalton 7 markers and High Dalton markers respectively. Lanes 3 and 4 both contain 25μl of different crude maturing preparations. Lanes 5 and 6 contain 25μl of different crude germinating preparations. Lanes 9 and 10 contain Dalton 7 markers and High Dalton markers respectively.

In figure 17 the bands corresponding to the maturing and germinating α-galactosidases can clearly be seen in lanes 2 to 6. By plotting the mobility of the molecular weight markers through the gel against the logarithm of their Mr it is possible to create the graph showing the linear relationship in figure 18. It is then a simple matter of comparing the mobilities of the maturing and germinating α-galactosidases to figure 18 to obtain their Mr's (as shown in Table 4)

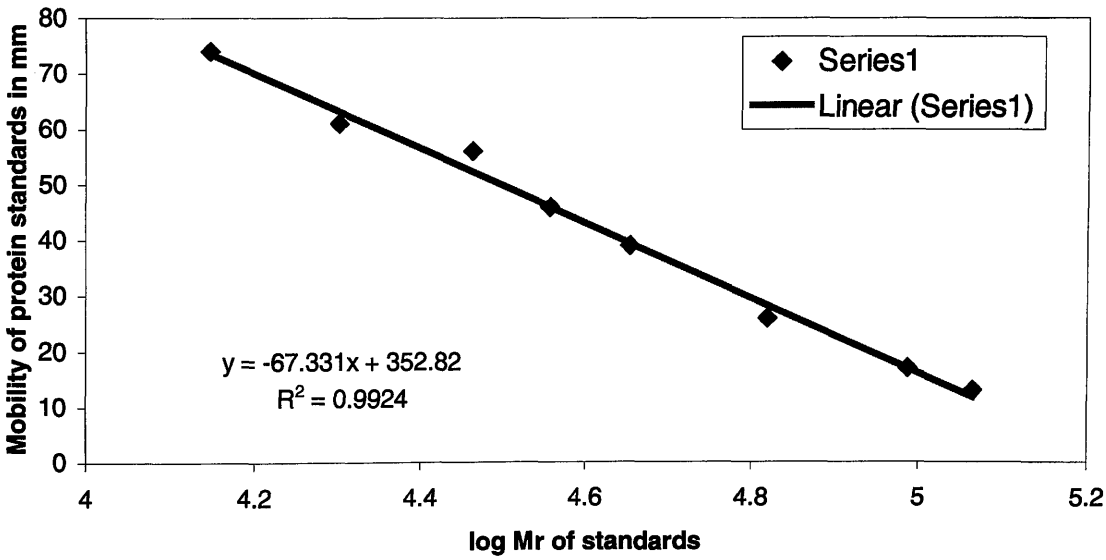


Figure 18. The Relative mobilities of Dalton 7 and High Dalton Markers plotted against the logarithm of their Mr's.

Table 4 A table showing the known molecular weights of the protein standards used to construct figure 16 and the molecular weights of the maturing and germinating α-galactosidases calculated using figure 18.

PROTEIN	RELATIVE MOBILITY (mm)	KNOWN Mr (Da)	CALCULATED MR (Da)
α-lactalbumin	74	14,000	
Trypsin	61	20,000	
Carbonic Anhydrase	58	29,000	
Glyceraldehyde-PDH	46	36,000	
Ovalbumin	39	45,000	
BSA	25	66,000	
Phosphorylase B	17	97,000	
β-Galactosidase	13	116,000	
Germinating α-gal	**41**		**43,000**
Maturing α-gal	**40**		**44,000**

The identification of the pI of *senna* α-galactosidases.

In order to identify the pI's of both the maturing and the germinating senna α-galactosidases isoelectric focusing gels were run. The isoelectric focusing gels were run as described in the methods section. The isoelectric focusing gels contained in different wells both the maturing and the germinating preparations. After running, the IEFs were cut into columns of gel containing the focused proteins of either the maturing or germinating preparations. Two of these columns (one maturing and one germinating) were cut into 2mm slices and assayed for α-galactosidase activity. A second pair of columns were measured for pH. The final gel columns were stained to show the locations of all the proteins present. Figure 19 shows the relationship produced between the pH of the slices along the gel and the localisation of α-galactosidase activity when the maturing protein preparation is run on the IEF.

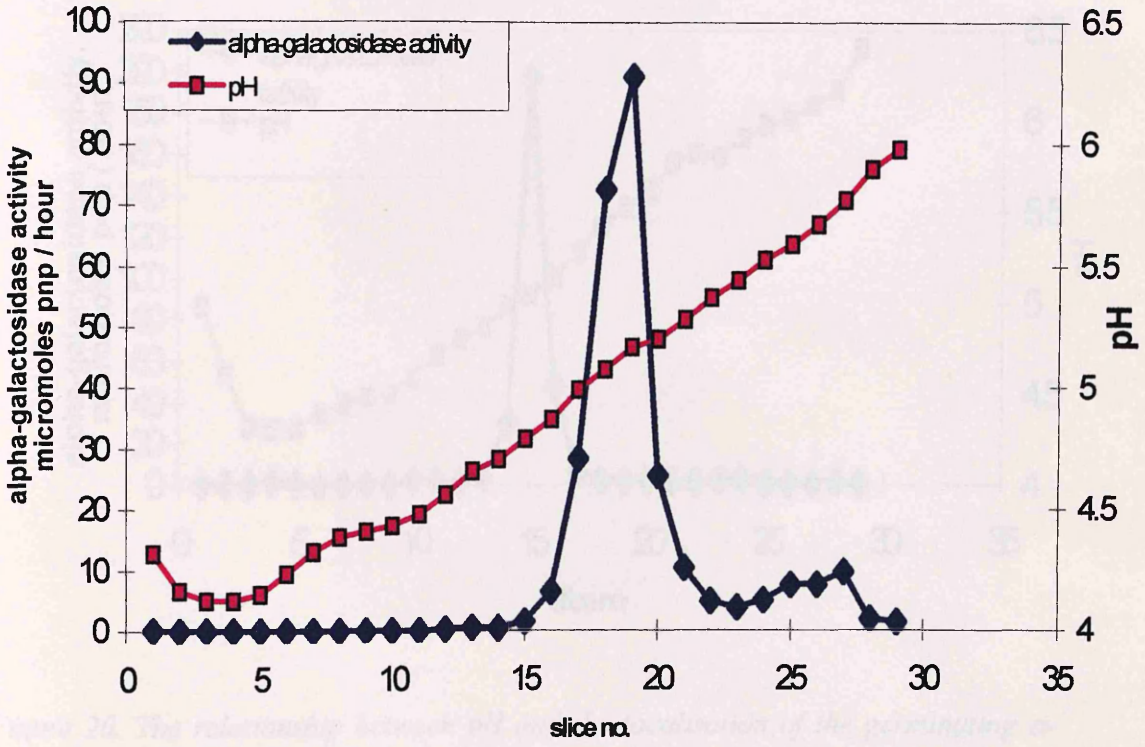


Figure 19. The relationship between pH and the localisation of the maturing α-galactosidase activity on an IEF.

Figure 20 shows the pH of gel slices and the localisation of α-galactosidase activity produced when the germinating α-galactosidase is run on the IEF gel.

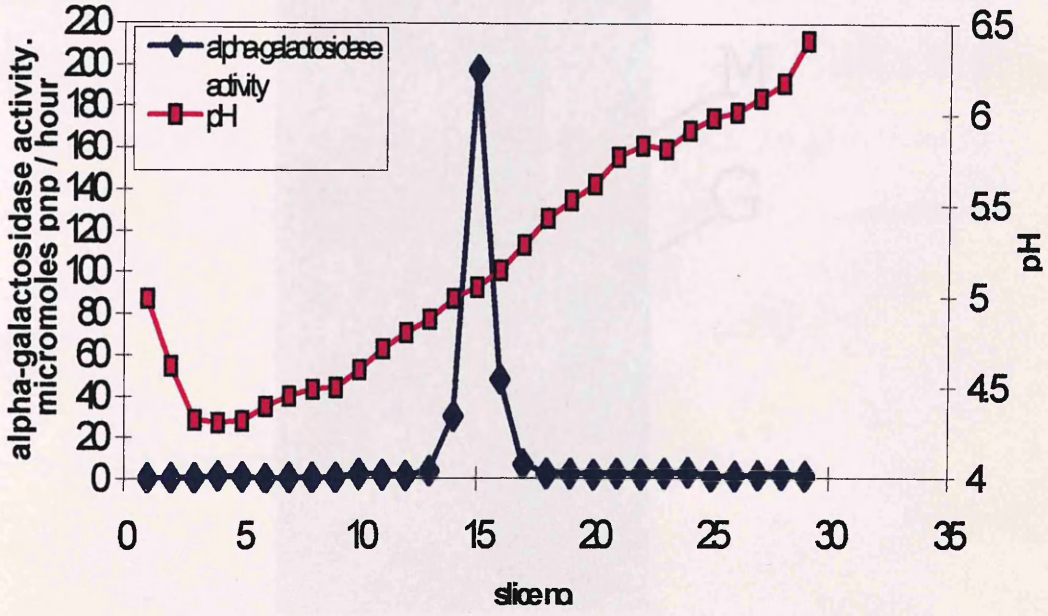


Figure 20. The relationship between pH and the localisation of the germinating α-galactosidase activity on an IEF.

Using Figures 19 and 20 it is possible to estimate the pI of both the maturing and germinating enzymes. These are pH 5.1 for the maturing α-galactosidase and pH 4.9 for the germinating α-galactosidase.

Figure 21 shows the stained sections of an IEF gel.

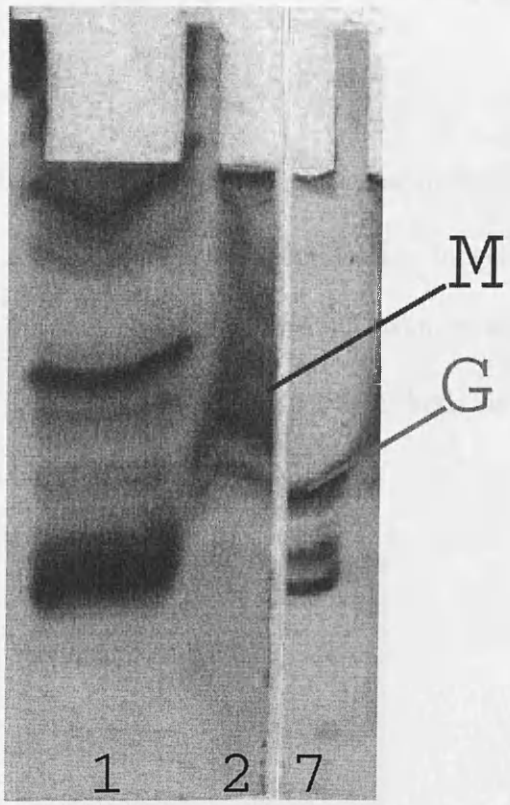


Figure 21. The localisation of α-galactosidase activity on stained IEF. Lane 1 contains sigma IEF standards. Lane 2 contains the focused proteins from a maturing senna seed preparation. Lanes 3 to 6 have been removed for pH and activity analysis. Lane 7 contains the focused proteins from a germinating senna seed preparation. Using information from the activity assays the bands corresponding to the α-galactosidases have been indicated. The maturing α-galactosidase is labelled M and the germinating α-galactosidase is labelled G.

A comparison of the two α-galactosidases primary peptide structure.

In order to compare the maturing and germinating α-galactosidases further, peptide mapping was carried out using two different peptidases. In gel digestions were performed as described in the methods section and the products were visualised by coomassie blue staining. Figure 22 shows a peptide map of both the α-galactosidases produced by the GluC peptidase.

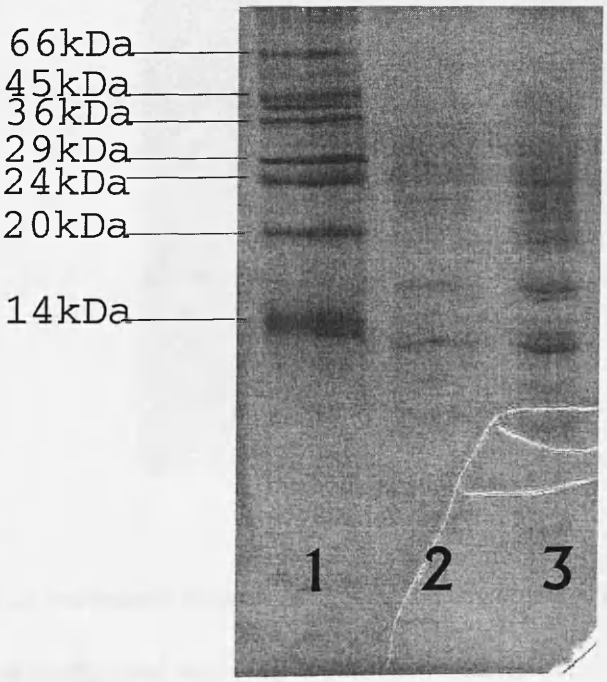


Figure 22. A coomassie stained 18% SDS PAGE showing; in lane 1 Dalton7 markers + 1 strip of undigested maturing α-galactosidase (to act as a reference), in lane 2 a GluC peptide digest of the maturing α-galactosidase and in lane 3 a GluC peptide digest of the germinating α-galactosidase.

Apart from the heavier loading of the germinating α-galactosidase both digests look identical. All the bands produced in the GluC digestion of the maturing α-galactosidase have corresponding bands produced in the GluC digestion of the germinating α-galactosidase. When a different peptidase, LysC was used the same result was found. This is shown in figure 23.

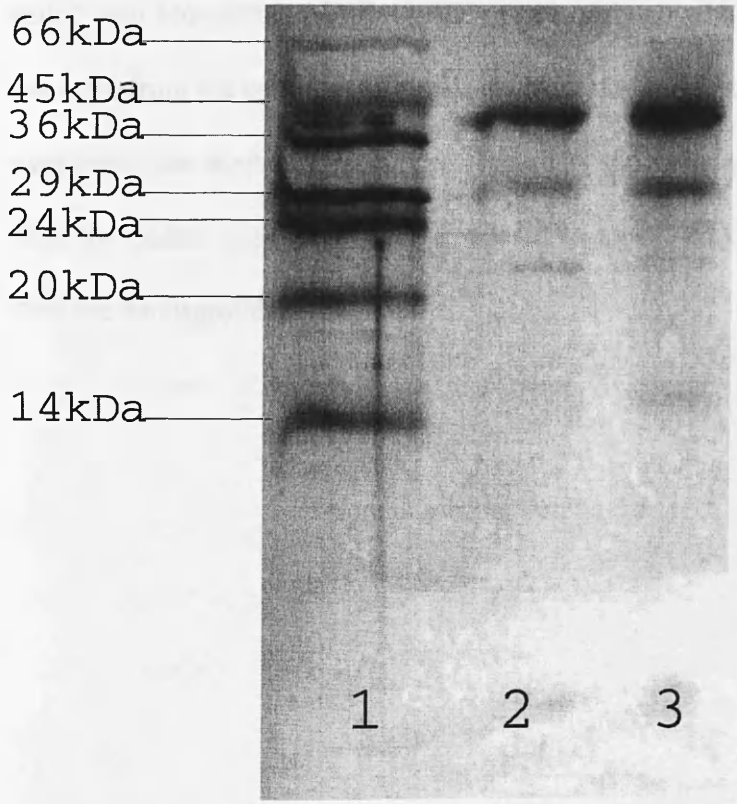


Figure 23. A coomassie stained 18% SDS PAGE showing; in lane 1 Dalton7 markers + 1 strip of undigested maturing α-galactosidase (to act as a reference), in lane 2 a LysC peptide digest of the maturing α-galactosidase and in lane 3 a LysC peptide digest of the germinating α-galactosidase.

As with the GluC digestion, the LysC digestion of the two α-galactosidases produce identical banding patterns.

To further investigate the similarities of the primary sequence of both the maturing and germinating α-galactosidases, peptide sequencing was carried out. As described in the methods, peptide digests were prepared and run on SDS PAGE gels. They were then blotted onto Problot membranes, stained and sent to Unilever for N terminal amino acid sequencing. Unfortunately it was only possible to obtain internal peptide sequence from the germinating α-galactosidase, the maturing peptides were in too low a concentration to obtain a sequence. Figure 24 shows the peptide sequences obtained from the Senna α-galactosidases compared to Guar and Coffee sequences obtained from the Swissprot database.

Maturing senna α-galactosidase N terminus= LGNGLGNTPPMG*N

Germinating senna α-galactosidase N terminus= NGLGNTPPMG*N

Coffee α-galactosidase= .LA.NGLGLTPPMGWN

Guar α-galactosidase= .LAENGLGQTPPMGWN

N-Terminal peptide sequences

Maturing senna α-galactosidase= GNGGMTTE*YR*HF

Coffee α-galactosidase= GNGGMTTTEYRSHF

Guar α-galactosidase= GNGGMTTEEYRSHF

Internal Peptide sequence A

Maturing senna α-galactosidase= VIAVNQDSLGVQGKKV*SDA

Coffee α-galactosidase= VIAVNQDKLGVQGNKVKTYG

Guar α-galactosidase= VIAVNQDKLGVQGKKVKSTN

Internal peptide sequence B

Figure 24 A comparison of peptide sequences from senna, guar and coffee α-galactosidases.

The locations of the sequenced senna peptides can be mapped against the sequences of the coffee and guar α-galactosidases to give the estimated locations of the peptides along the α-galactosidase primary structure. This map is shown in figure 25

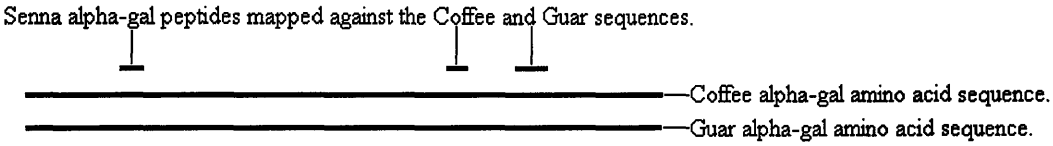


Figure 25. The locations of the regions of coffee and guar homologous to the sequenced senna peptides.

It is evident from the N terminal amino acid sequences that the maturing and germinating α-galactosidases are very similar, the only difference being an extra 2 amino acids on the N terminus of the maturing enzyme. The fact that the germinating α-galactosidase lacks a residue could be simply due to differential signal sequence excision taking place during protein translation.

The isolation of RNA from maturing *senna* endosperms and Polymerase chain reactions.

The initial method used to isolate RNA was that of Lopez-Gomez (1992). 100 34 day old maturing endosperms were used along with a positive control of 2.5g of young leaf tissue. The endosperm preparation became very viscous and had to be abandoned. However, the positive leaf control was successful producing large amounts of total RNA. as shown in figure 26.

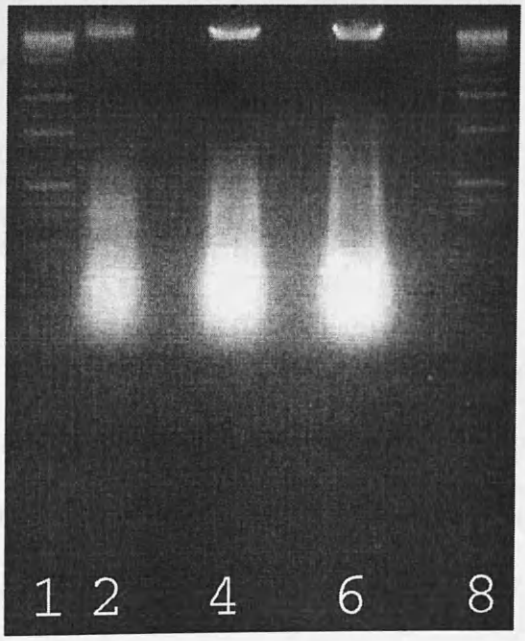


Figure 26. The RNA prepared from young senna leaves using the Lopez-Gomez method. Lanes 1 and 8 contain HindIII standards. Lanes 2, 4 and 6 contain 2µl, 5µl and 10µl respectively of young leaf RNA.

Further RNA preparations were carried out using four times the amount of lysis buffer to endosperms. These preparations, although they became viscous, were not so viscous as to prevent RNA being prepared. Figure 27 shows a comparison between preparations of 40 maturing endosperms and 1.87g of young leaf tissue.

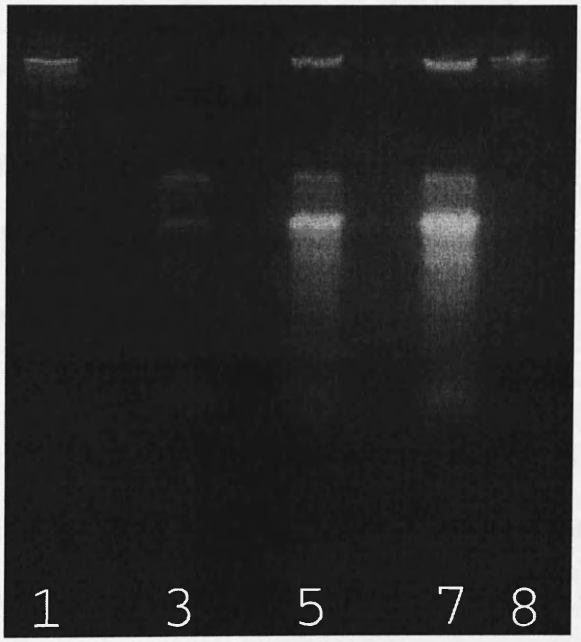


Figure 27. The result of the preparation of RNA from 40 maturing endosperms resuspended in 25µl of depc H_2O and the preparation of RNA from 1.87g of young leaves resuspended in 50µl of depc H_2O. Lanes 1and 8 contain HindIII markers. Lane 3 contains 10µl of endosperm prepared RNA lanes 5 and 7 contains 5µl and 10µl of the leaf prepared RNA

This method was repeated 3 times. However the amount of total RNA prepared was always very low, this is shown in Table 5 below.

Table 5. The amounts of total RNA prepared from maturing endosperms using the Lopez-Gomez method.

Modified prep number.	$A_{260}/_{280}$ ratio	Concentration of total RNA $\mu g\ \mu l^{-1}$	Volume of RNA prep µl	Yield of RNA in µg
1	1.6	0.038	25	0.95
2	1.9	0.055	70	3.9
3	1.7	0.013	25	0.33
4	1.5	0.013	50	0.66

Due to the fact that mRNA constitutes only 5% of the total RNA and the amount of RNA needed for RT-RACE-PCR was 6µg it was necessary to use a different method to obtain larger quantities of endosperm RNA.

The Qiagen plant RNA preparation kit was the next method used (as described in the methods). As in the previous method, difficulties were encountered with this kit because the lysed 34 day old maturing endosperm tissue became very viscous. The preparation that used the RLC lysis buffer produced a final flow-through of only 3µl rather than 30µl due to the viscosity of the sample in the spin column. The RLT lysed sample did however produce a useful amount of flow-through. The amount of RNA was measured spectroscopically and is shown in Table 6

Table 6 The volume, concentration and yield of RNA produced from the Qiagen (RLT) preparation of maturing senna endosperm.

Concentration of total RNA $\mu g\ \mu l^{-1}$	Volume of RNA prep	Yield of RNA in µg
0.06	27	1.62

The attempt to run a gel containing the Qiagen prepared RNA failed because the sample refused to remain in the well, possibly due to the presence of residual ethanol in it.

The Qiagen method was repeated with the modifications outlined in the methods section. The RNA content was again measured spectroscopically and the result of this is shown in Table 7.

Table 7. The $A_{260}/_{280}$ ratio, volume, concentration and yield of RNA produced from the modified Qiagen preparations of maturing senna endosperm.

Preparation	$A_{260}/_{280}$ ratio	Concentration of total RNA $\mu g\ \mu l^{-1}$	Volume of RNA prep	Yield of RNA in µg
Ethanol+RLC	1.86	0.025	10.5	0.26
Filtered RLC	1.47	0.045	7.5	0.34
Ethanol+RLT	1.00	0.0050	10.5	0.050
Filtered RLT	1.60	0.0075	4.5	0.034

The attempt to load a gel was again prevented because the samples floated out of the wells. Finally, to one set of samples glycerol was added, and to a second set of samples ethanol precipitation was carried out, followed by a thorough air drying and resuspension. Both methods prevented the samples from floating out of the wells and so allowed gels to be run. The gels however showed that there was **no RNA present**, the spectrometer had been measuring a small amount of contaminating DNA.

Dynabead extractions of messenger-RNA from maturing endosperms and a positive control of young senna leaves were performed.

Two preparations of mRNA were made from 0.1g of the ground tissue flour of maturing endosperms and young leaves. The preparations resulted in the production of 15µl of elution from the endosperm preparation and 17µl of elution from the leaves. 2µl of the leaf prep. was spectroscopically measured, the result of which it shown in Table 8.

Table 8. The $A_{260}/_{280}$ ratio, volume, concentration and yield of mRNA produced from the modified Dynabead preparations of young senna leaves.

$A_{260}/_{280}$ ratio	Concentration of total RNA µg $µl^{-1}$	Volume of RNA prep, µl	Yield of RNA in µg
1.083	0.015	17	0.255

The remaining 15µl of the leaf preparation was mixed with sample buffer and run on an agarose gel. The gel is shown below in figure 28.

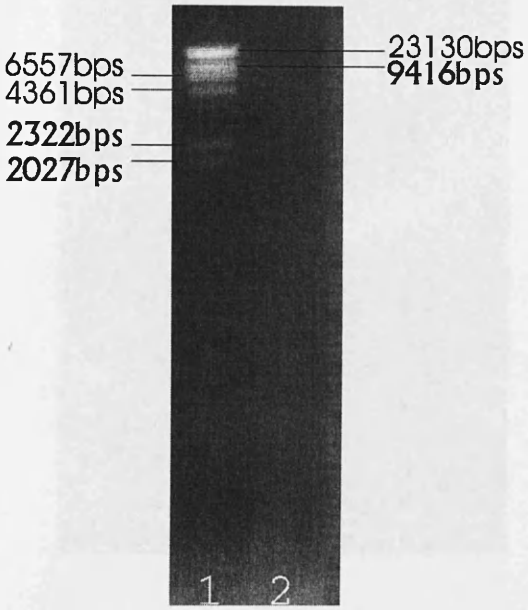


Figure 28. The product of the mRNA preparation positive leaf control run on an agarose gel. Lane 1 contains 5µl of HindIII. Lane 2 contains 15µl of leaf mRNA preparation.

Figure 28 shows that the mRNA preparation contains a light smear that could be mRNA. This being the case RT-3' RACE PCR was carried out using T_7 and primers 1, 2, 3 and a mix of 2 and 3 (produced from the amino acid sequence of the N-Terminus of the α-galactosidase) with an annealing temperature of 50°C as described in the methods section.

The PCR products were loaded onto an Agarose gel and run. The results of this PCR is shown in figure 29.

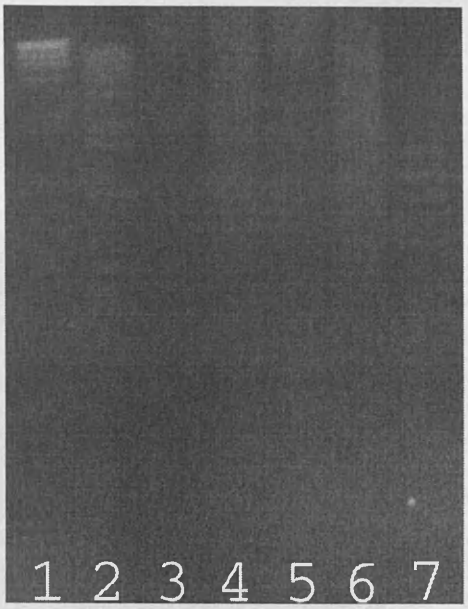


Figure 29. The results of the RT-3' RACE PCR . Lane 1 contains HindIII markers. Lane 2 contains 1kb markers. Lanes 3 to 6 contain the PCR products of the reactions using primer T7 with: in lane 3 P1, in lane 4 P2, in lane 5 P3 and in lane 6 P2 and P3.

Figure 29 shows that no definite PCR product was produced using any of the primers (no bands are visible, only a smear is evident in the PCR products). The same result was obtained using an annealing temperature of 55°C with the Dynabead produced first strand cDNA.

A further attempt to use Dynabeads to prepare mRNA was made when a solid state cDNA library was made using dilute total RNA prepared by the Lopez –Gomez method. After the production of the library, four 3' RACE PCR reactions were carried out using either primer 1 and T7 or Primer mix 2+3 and T7 on the heat incubated or non-heat incubated Dynabead libraries (see the methods section for details). The results of these PCR's can be seen in figure 30.

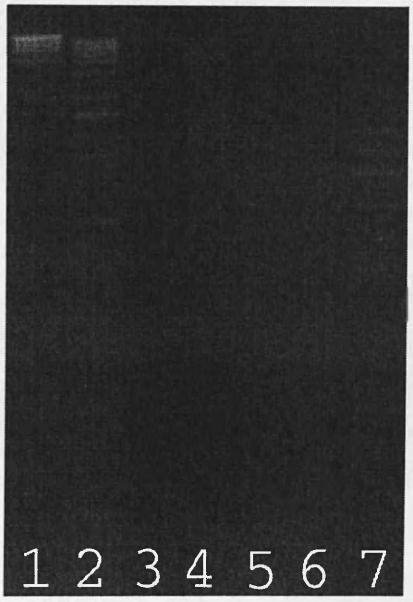


Figure 30. The PCR products of reactions 1-heated (lane 3), 2+3-heated (lane 4), 1-nonheated (lane 5) and 2+3-nonheated. The markers shown here are HindIII (lane 1), 1kb (lane 2) and ΦX174 (lane7)

Figure 30 shows that no definite PCR product bands are visible in either of the heated cDNA library PCRs, however, smears are present from very high to very low molecular weights. The lanes containing the PCR products of the non-heated cDNA library are totally empty. The PCR product from the heated cDNA library produced with primers 2+3 was "Promega Wizard" cleaned up (as per the methods section) and a further round of PCR was carried out on it in an attempt to amplify any faint bands that may be contained within the smear. Negative controls (no primers, no Taq and no template) were also set up. The result of these reactions are shown in figure 31.

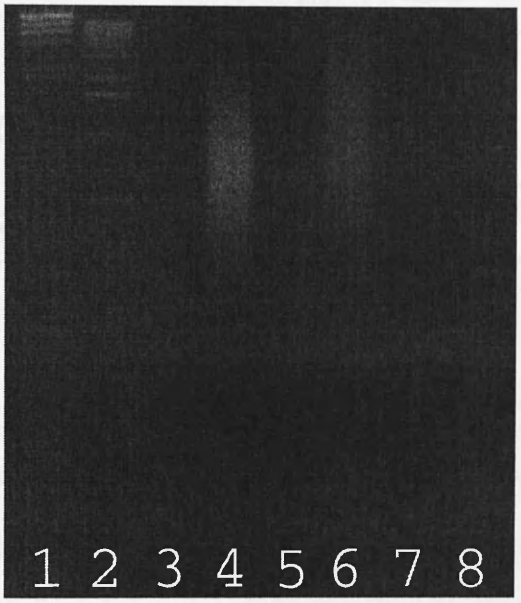


Figure 31. The PCR re-amplification of the heated cDNA-Dynabead library. Lane 1 contains HindIII markers and lane 2 contains 1kb markers. Lane 4 contains the re-amplified PCR product. Lanes 6, 7 and 8 are negative controls (no primers, no Taq and no template respectively).

Figure 31 shows that no distinct PCR product bands were produced in the re-amplification and that the smear is an artefact produced by incorrect priming. This result was repeated at higher annealing temperatures.

The isolation of RNA from germinating *senna* endosperms and Polymerase chain reactions.

Due to the fact that it had been impossible to obtain useful amounts of RNA from maturing endosperms, standard Qiagen preparations using 10 mg of endosperms from *senna* seeds that had germinated for three days were carried out. The preparations were not hampered by excessive gelling of the samples and both the RLC and RLT lysis buffers produced flow through. The amount of RNA was estimated by comparing the fluorescence of an aliquot of the sample with a serial dilution of known nucleic acid concentration. The result of the Qiagen preparation is shown in Table 9

Table 9. The volume, concentration and yield of RNA produced from the standard Qiagen preparations of germinating senna endosperms.

Preparation	Concentration of total RNA $\mu g\ \mu l^{-1}$	Volume of RNA prep µl	Yield of RNA in µg
RLC	0.01	15.5	0.155
RLT	0.01	15	0.15

Although the amount of RNA prepared was again very low, a RT-PCR was attempted as before. No PCR product was produced, as can be seen in figure 32.

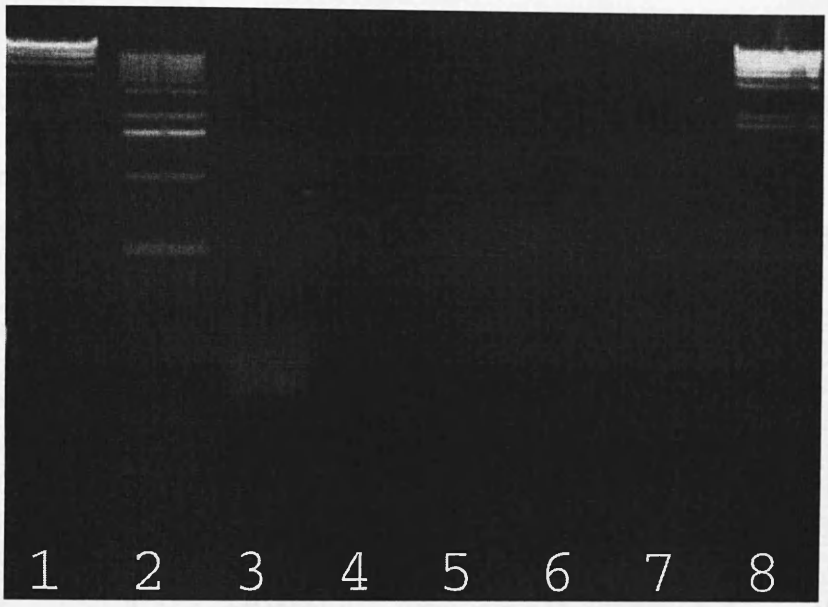


Figure 32. The products of the PCR using RNA from the Qiagen extraction of germinating senna endosperms. Lanes 1 and 8 contain HindIII. Lane 2 contains 1kb. Lane 3 contains the PCR product produced from the RLT produced RNA, lane4 the product of the RLC produced RNA. Lanes 5, 6 and 7 contains negative controls (no Taq, no primers and no template respectively).

Final Qiagen preparations of germinating endosperms were carried out in an attempt to isolate the maximum RNA possible. Two preparations were performed, A, with 15mg of ground senna endosperm flour and B, with 30mg of tissue. The standard Qiagen method was used with the exception that the final elution of the RNA solution from the spin column was repeated four times with:-1, 30µl DEPC H_2O, 2, 30µl DEPC H_2O, 3, 15 µl DEPC H_2O, 4, 15 µl DEPC H_2O. The results of the preparations are shown in Table 10.

Table 10. The, volume, concentration and yield of RNA produced from the four elutions of the Qiagen preparations A and B of maturing senna endosperm.

Preparation number	Concentration of total RNA μg μl^{-1}	Volume of RNA preparation μl	Yield of RNA in μg
A1	0.015	7	0.11
A2	0.010	27	0.27
A3	0.005	16	0.08
A4	0.0025	16	0.04
B1	0.025	7.5	0.19
B2	0.015	25	0.38
B3	0.010	22	0.22
B4	0.005	17	0.09

Although the concentration of RNA again seems very low, the total yield of RNA was 2.5μg. The RNA was therefore ethanol precipitated, dried, rehydrated and used in RT-PCR as described in the methods section. The result of this PCR is shown in figure 33.

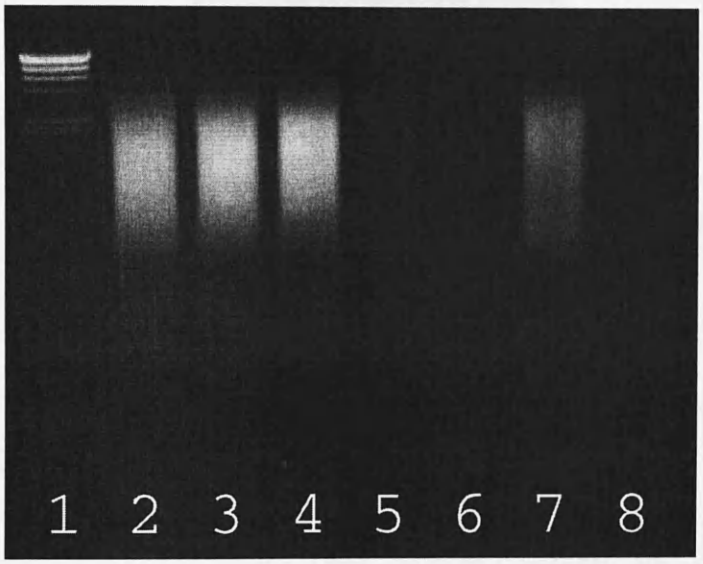


Figure 33. The products of the PCR using RNA concentrated from the Qiagen extraction of germinating senna endosperms. Lane 1 contains HindIII markers. Lane 2 contains the PCR product produced using primers P1 and T7, lane 3 contains the product of P2+3 and T7, lane 4, GS1 and T7 and lane 5 the product of primers P2+3 and GA1. Lanes 6, 7 and 8 contains negative controls (no Taq, no primers and no template respectively).

Again large smears are present in all the experimental lanes and the no primers negative control. Something in the RT-PCR must be acting as a non-specific primer. Therefore the Qiagen preparation and concentration of RNA was repeated. The reverse transcription was repeated and half of this was "Promega Wizard" cleaned up and the other half was run on a gel with the intention of cutting out and cleaning up first strand cDNA above 200 bps. The gel run is shown in figure 34.

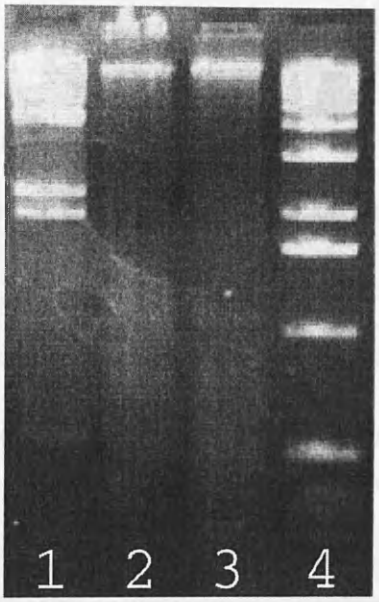


Figure 34. The RNA and cDNA produced by a Quiagen preparation using germinating endosperms. Lane 1 contains HindIII and lane 4 contains 1kb markers. Lane 2 should contain the Qiagen prepared RNA and lane 3 should contain cDNA produced from it.

Figure 34 clearly shows that the nucleic acid that had been prepared (and measured) by the Qiagen plant RNA preparation kit was not in fact RNA but high molecular weight genomic DNA. The Qiagen method was therefore abandoned.

Due to the failure of the Qiagen method a Guanidine HCl RNA preparation method was used with 0.25g of liquid nitrogen ground germinating endosperm tissue flour. The procedure was carried out and the resulting pellet was dissolved and the nucleic acids present were measured. The result of the Guanidine HCl preparation was that **no RNA was isolated.**

The Lopez-Gomez method of RNA isolation was returned to, with 150 3 day old germinating senna endosperms being ground in liquid Nitrogen and used in the modified method. The nucleic acids produced were measured and shown in Table 11.

Table 11. The volume, concentration and yield of RNA produced from the modified Lopez-Gomez preparations of 150 germinating senna endosperms.

Concentration of total RNA $\mu g\ \mu l^{-1}$	Volume of RNA preparation µl	Yield of RNA in µg
0.05	20.0	1.00

Table 11 shows that this method has isolated the most RNA so far. The RNA was run out on a gel and this is shown in figure 35.

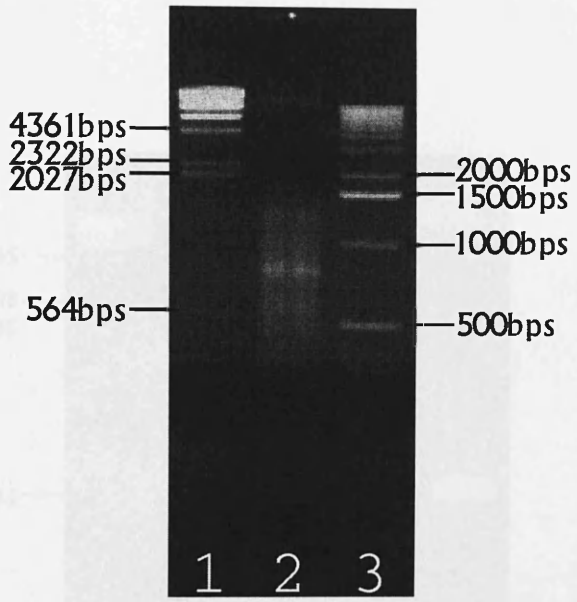


Figure 35. The RNA prepared from germinating senna endosperms using the modified Lopez-Gomez method. Lane 1 contains HindIII markers, lane 2 contains 3µl of the prepared RNA and lane 3 contains 1kb markers.

Figure 35 shows that the RNA preparation did indeed isolate RNA, ribosomal RNA bands are visible and so is a smear that may be mRNA. Within the RNA preparation is also high molecular weight DNA and polysaccharide can be seen remaining in the well of the gel.

First strand cDNA was made using this RNA (as previously described) which was then "Promega Wizard" cleaned up. PCR reactions were then set up using primers P1+T7, GS1+T7, P1+GA1, no enzyme control (with primers GS1+T7 present), no primer control and a no template control (with primers GS1+T7 present). The PCR was performed with the annealing temperature set to 50°C. The products of the PCR were run out on an agarose gel, this is shown in figure 36.

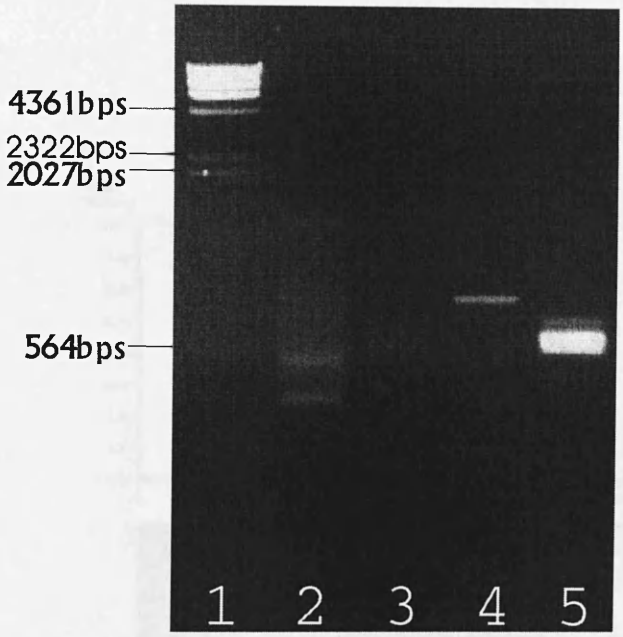


Figure 36. The products of the PCR using RNA from the Lopez-Gomez extraction of germinating senna endosperms. Lane 1 contains HindIII. Lane 2 contains the product of primer P1+T7, lane 3 the product of GS1+T7 and lane 4 the product of P1+GA1. Lane 5 contains a repeat of GS1+T7.

Figure 36 shows that PCR products were produced in all the experimental reactions (and none were produced in the controls). The lanes were analysed to give the sizes of the bands present. This analysis is shown in figure 37.

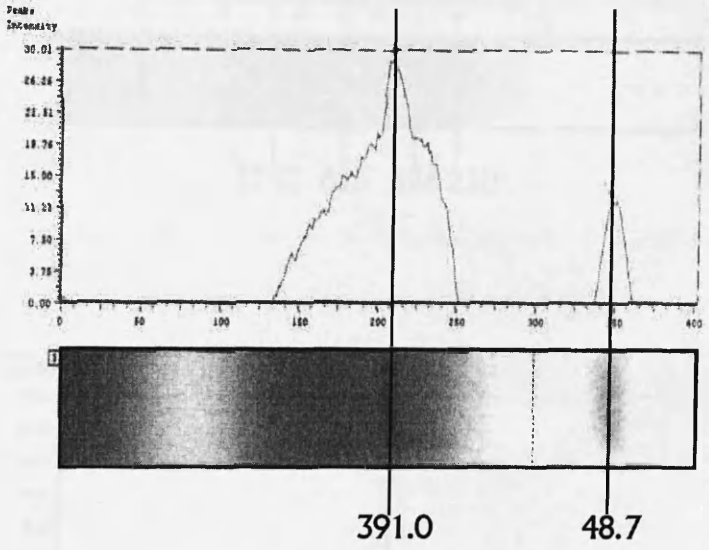


Figure 37A.

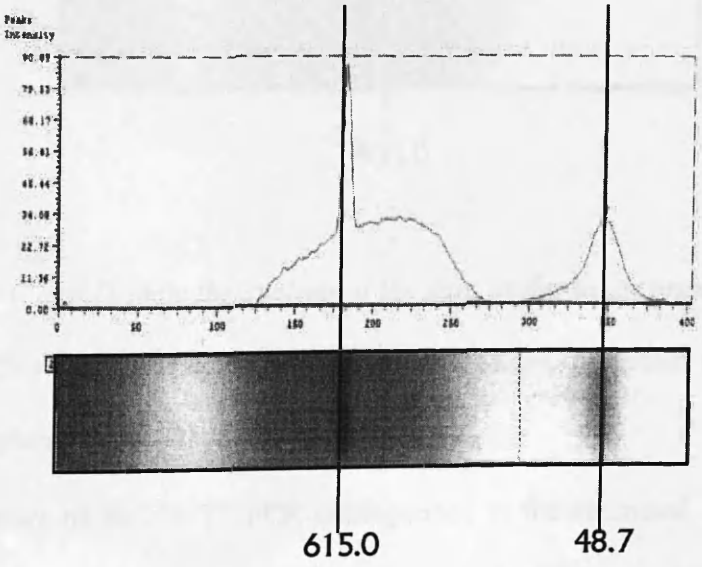


Figure 37B

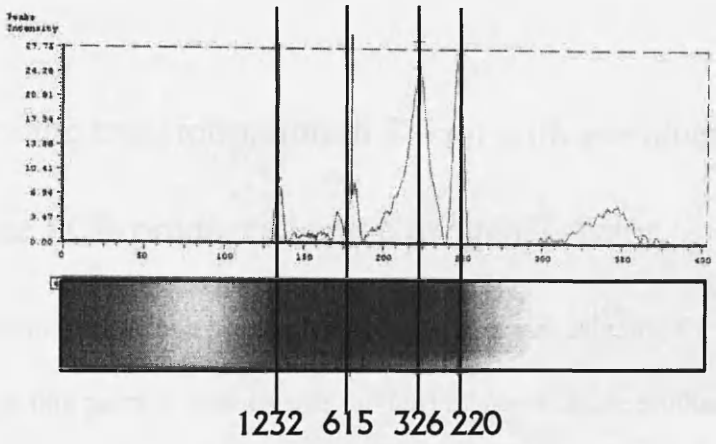


Figure 37C

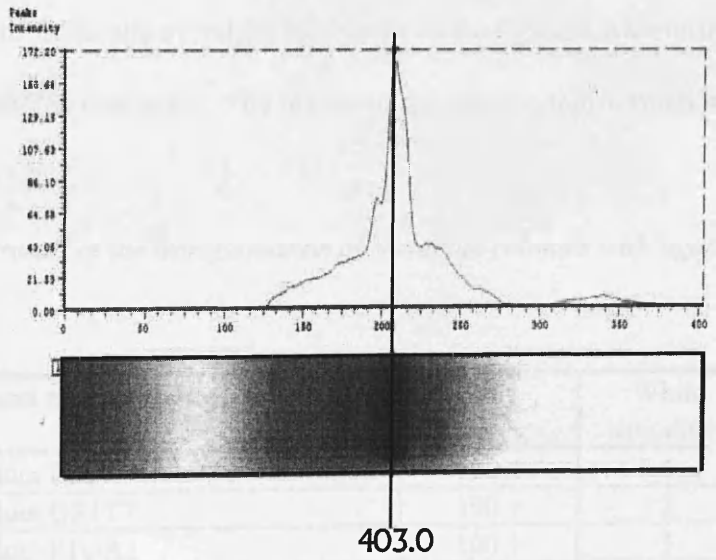


Figure 37D

Figures 37A, B, C and D show the analysis of the sizes of the bands present in the 4 PCR products. The analysis is based on the mobility of the PCR products through the agarose gel compared to the HindIII standards.

The largest product of the P1+T7 PCR corresponded to the estimated size of the cDNA that would produce the 43kDa α-galactosidase. The products of the GS1 and the GA1 primers were also approximately the size expected when a comparison of the Senna, Guar and Coffee amino acid sequences was made (figure 25).

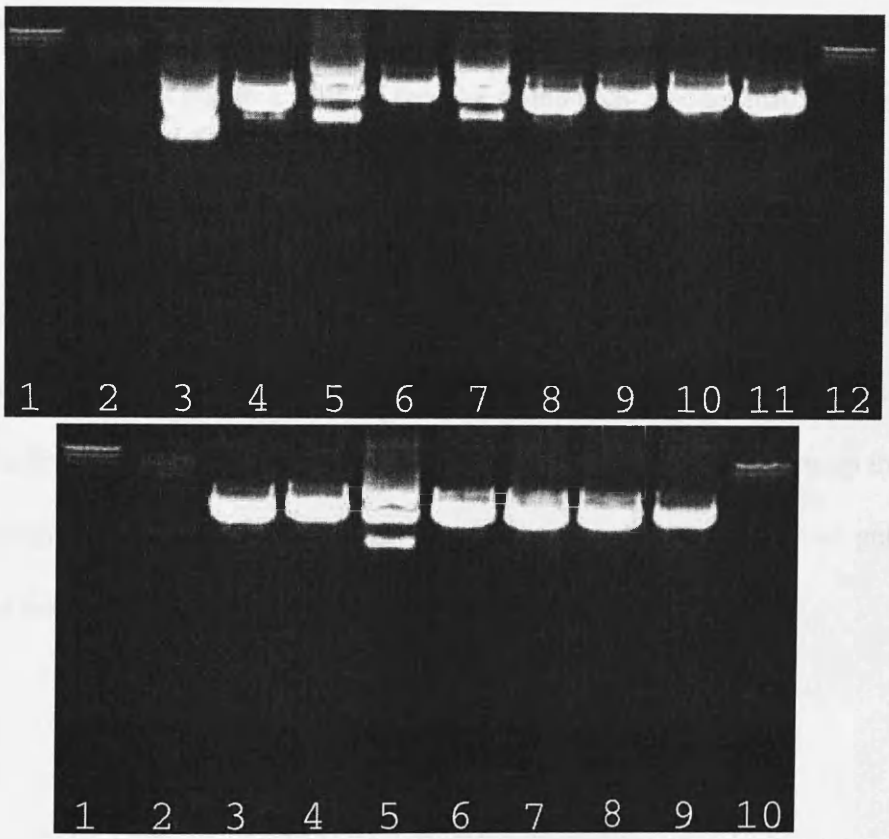
The ligation and transformation of *E. coli* with germinating α-galactosidase PCR products using Novogen Tvector.

The PCR products were heavily loaded onto an agarose gel, run and the PCR bands were excised (at this point it seemed that the GS1T7-repeat PCR product seemed to consist of two very similarly sized bands which can be seen in figure 36). The gel slices were then cleaned up and ligated into the Novogen T vector as described in the methods section. After the overnight incubation of the ligation, transformations were set-up and incubated overnight. The results of the ligation/transformation are shown below in Table 12.

Table 12. The result of the transformation of Novablue colonies with ligated Novogen T vectors.

Ligation and transformation.	Blue colonies.	White colonies.
PCR product P1T7	100 +	4
PCR product GS1T7	100 +	2
PCR product P1GA1	100 +	3
PCR product GS1T7 repeat upper band	100 +	3
PCR product GS1T7 repeat lower band	100 +	4
Novogen +ve control.	100 +	20
Novogen –ve control.	100 +	1
No vector control.	0	0
Ligated vector control.	0	100 +

The white colonies of the ligated PCR products were picked and incubated in L. Broth at 37°C overnight. The cultures were then minipreped (subjected to plasmid isolation) using the "Promega Wizard" kit and the resulting plasmids were digested with EcoR1 and HindIII restriction enzymes (as described in the methods). The resulting fragments were run on an agarose gel which is shown in figure 38.



Gel A (top) and Gel B (bottom).

Figure 38. The digestion of the ligated PCR products with EcoR1 and HindIII. Gel A shows in lanes 1 and 12 HindIII markers and in lane 2 1kb markers. Lanes 3-6 contains the 4 clones produced from the P1T7 PCR product. Lanes 7 and 8 contain the products of GS1T7. Lanes 9-11 contains the products of P1GA1. Gel B shows in lane 1 and 14 HindIII markers and in lane 2 1kb markers. Lanes 3-5 contain the product of the clones produced from the GS1T7(repeat)-lower band and lanes 6-9 contain the product of the clones produced from the GS1T7(repeat)-upper band. Lane 10 contains a positive restriction digest control. Lane 11 contains a no digest negative control. Lane 12 contains a EcoR1 only digestion (a linearisation). Lane 13 contains a HindIII only digestion (a linearisation).

Figure 38 shows that apart from gel A lane 3 (P1T7 clone A) none of the digestions of the plasmids gave rise to a band of the size of the PCR product that was ligated into them. The EcoR1 digest gave the same banding pattern as no digest at all.

In an attempt to isolate colonies of the PCR products other than P1T7 three more transformation reactions were carried out. The further transformations resulted in the production of 44 white colonies. The whites were again picked, grown up in culture, minipreped and restriction digested. On running the digests on agarose gels it was found that **no white colonies contained inserts.**

PCR confirmation of the success of the ligation and the transformation of novogen T vector with the PCR product of primers P1 and T7.

To check that P1T7 clone A did contain the PCR product resulting from the use of primers P1 and T7 a PCR was carried out using the P1T7-A clone as a template with, initially, the primers P1 and GA1. The products of this PCR were run on an agarose gel and this is shown in figure 39.

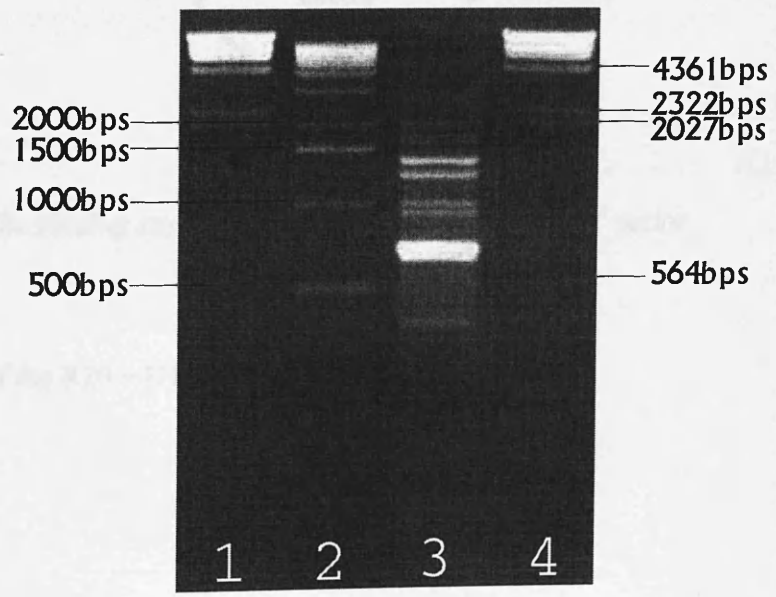


Figure 39. The result of the confirmation PCR using primers P1 and GA1. Lanes 1 and 4 contain HindIII markers. Lane 2 contains 1kb. Lane 3 shows the products of the confirmation PCR.

Figure 39 shows that the confirmation PCR produced many products (probably due to non-specific primer binding). However, the brightest PCR product was of the size that would be expected if P1T7-A did contain the correct product of the P1T7 PCR. A further confirmation was carried out by PCRing the clone using the primers R20 and U19 which bind to the Novogen T vector adjacent to the multiple cloning site as shown diagrammatically in figure 40.

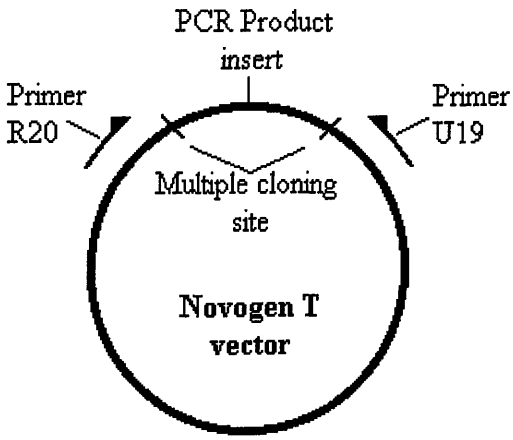


Figure 40 The binding sites of R20 and U19 on the Novogen T vector.

The result of the R20 + U19 PCR is shown in figure 41.

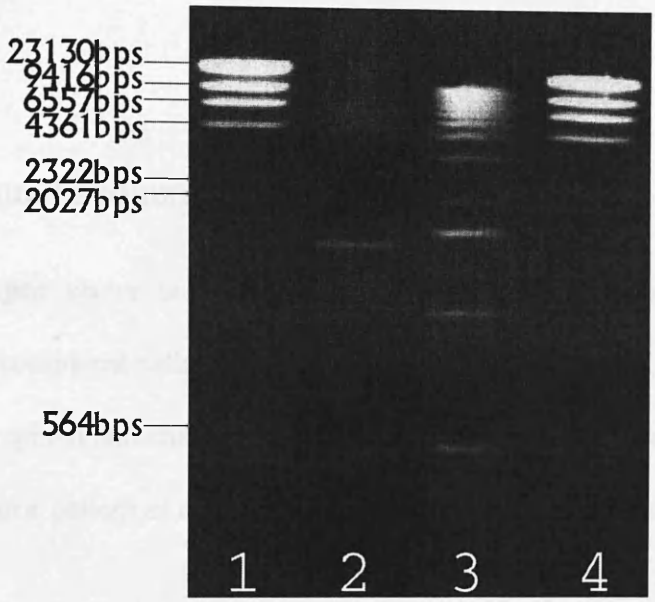


Figure 41. . The result of the confirmation PCR using primers R20 and U19. Lanes 1 and 4 contain HindIII markers. Lane 3 contains 1kb. Lane 2 shows the products of the confirmation PCR.

Figure 41 shows that the PCR has produced a range of products. However the sharpest and brightest band again corresponds to the size expected from a vector containing the P1T7 PCR product

The manual sequencing of P1T7 in the Novogen T vector.

The Novogen vector containing the PCR product was re-transformed into fresh Novablue competent cells. The resulting colonies were picked, grown up in culture, plasmid prepared and checked by restriction digest to make sure that the new vectors gave the same pattern as the original P1T7-Novogen T vector construct.

The new plasmids were prepared for sequencing (as described in the methods section) and then sequenced using the primers R20 and U19, located either side of the multiple cloning site. The sequencing reactions were run on a polyacrylamide sequencing gel. The gel was transferred to 3M paper and placed in a cassette in contact with x-ray film. After exposure the film was developed.

It was discovered that primer U19 gave no sequence data. The sequence obtained from the R20 primer was read and was compared to the novogen T vector in order to find the multiple cloning site. The comparison showed that the clone P1T7-A did not contain a PCR product insert and was missing approximately 800 bases upstream from the cloning site. This is shown in figure 42.

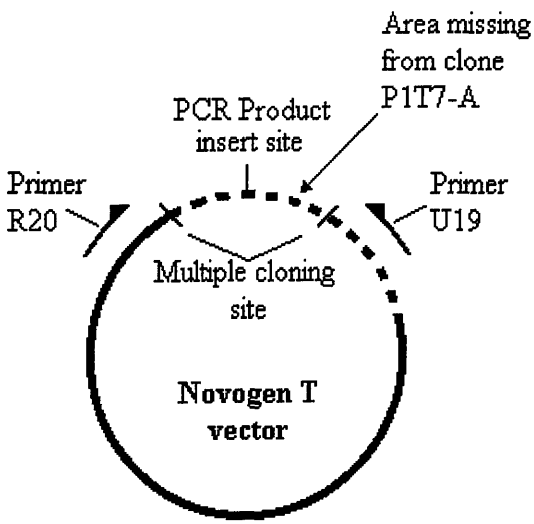


Figure 42. The resulting picture of the P1T7-clone-A after sequencing. The sequencing showed that there was no PCR insert and that 800 bases of the T vector was missing.

The ligation and transformation of *E. coli* with germinating α-galactosidase PCR products using blunt-ended ligation into pBluescript.

XL1 Blue cells were removed from the glycerol stock and used in the preparation of competent cells. After the production of the cells a test transformation was carried out with pBluescript. The result of the test transformation was that many colonies were produced confirming that the XL1 blue were competent.

pBluescript plasmids were grown up and then digested with SmaI using either IPA or React-4 buffers as described in the methods section. The linearised pBluescript was cleaned up using the Hybaid kit and then ligated with P1GA1 PCR product that had been gel purified. Control reactions were also set up. The ligations were then transformed into the competent XL1 blue cells.

The result of the ligation/transformation was that the transformation reaction worked, as shown by the fact that the uncut pBluescript control produced many blue colonies. However, the ligation reaction did not work, this was shown that no other reactions produced colonies including the positive, self ligation control.

The ligation and transformation of invitrogen *INVαF'* cells with germinating α-galactosidase PCR products using T/A ligation into PCR2 vector.

This method initially required the Adenosine overhangs of the PCR products in storage to be added back. This was carried out by the method described in the methods section. The PCR products were run out on an agarose gel and the bands excised and cleaned up using the Hybaid purification kit. The amount of DNA in the cleaned samples were measured and adjusted according to Invitrogen ligation method. The ligation reactions were set up and incubated overnight. The transformation reactions were performed and plated out. The 5 transformations produced 40 white colonies. The whites were picked, grown up, minipreped and restriction digested with EcoR1. The results of the ligation/transformations are shown in figure 43

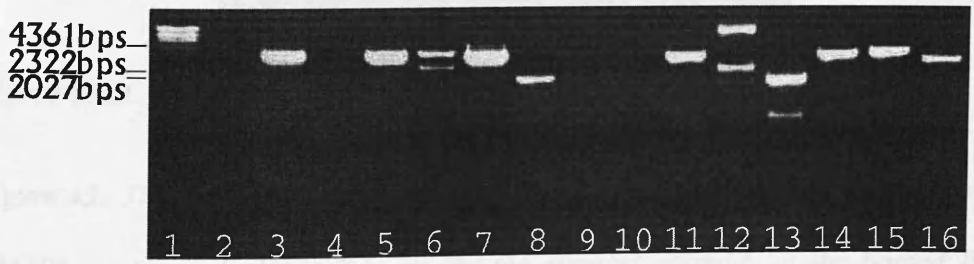


Figure 43-A

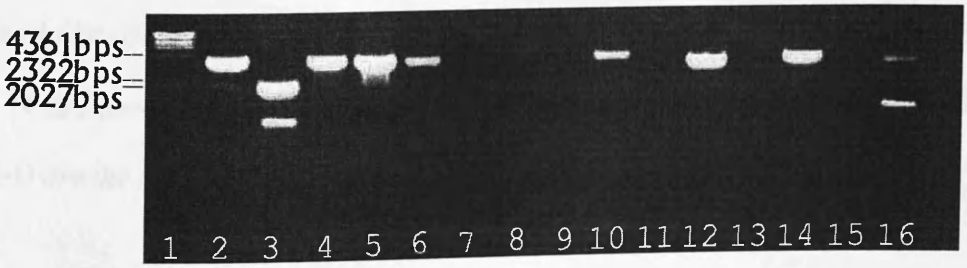


Figure 43-B

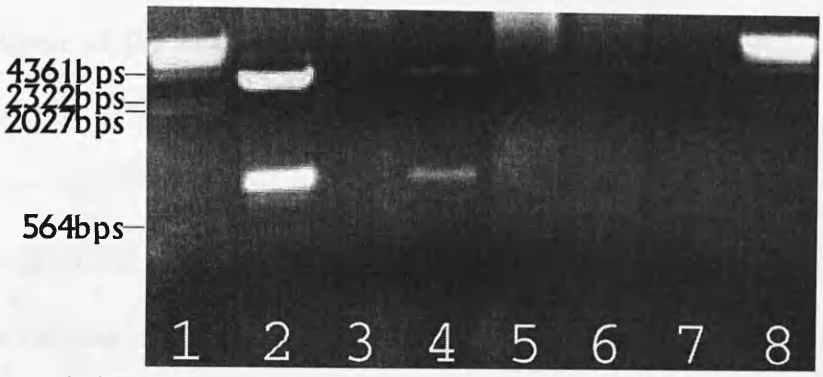


Figure 43-C

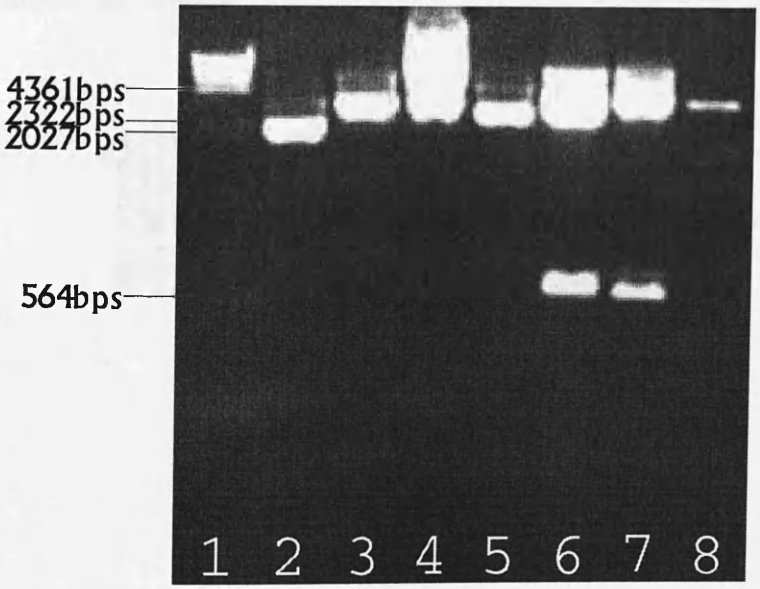


Figure 43-D

Figure 43. The result of the Invitrogen cloning method.. Figure 43-A shows in lane 1 HindIII markers. In lanes 2 to 10 are the digested plasmids of the ligated PCR product P1T7 and in lanes 11 to 16 are products of the GS1T7. Figure 43-B shows in lane 1 HindIII markers. In lanes 2 to 8 are further products of the GS1T7. In lanes 9 to 16 of figure 43-B and in lanes 2 to 5 of figure 43-C and in lanes 2 to 7 of figure 43-D are the the digested plasmids of the ligated PCR product P1GA1.

Figures 43 A to D shows that only the clones run in Figure 43-D lanes 2 and 4 contain an insert of the same size as the PCR products ligated into them (PCR product P1GA1).

Further PCRs were carried out in an attempt to clone the second half of the α-galactosidase cDNA (using GS1T7). The products of the PCR were invitrogen cloned, picked, grown up, plasmid prepared and digested as before. Figure 44 shows the result of this cloning.

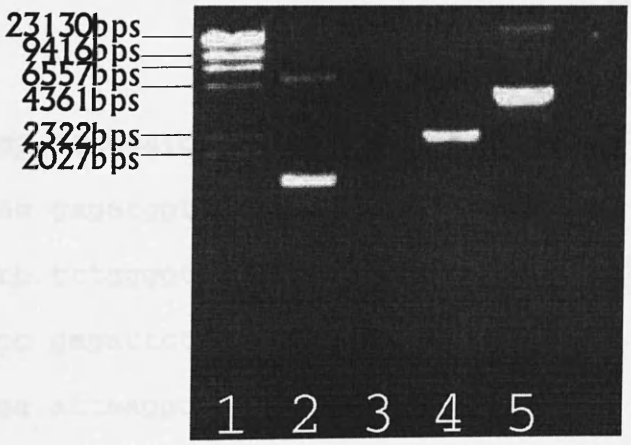


Figure 44. The result of the Invitrogen cloning of GS1T7. Lane 1 contains HindIII markers. In lanes 2 to 8 are the digested plasmids of the ligated GS1T7 PCR product.

Figure 44 shows that two clones (in lanes 6 and 7) contain inserts of the expected size.

The clones containing PCR products were used as templates in further to confirm that they did contain the correct insert. The result of this suggested that they did.

The manual sequencing of the Invitrogen prepared clones.

The plasmids were prepared for sequencing (as described in the methods section) and then sequenced using the primers T7 and P1 for the P1GA1 clone and primer GS1 for the GS1T7 clones. The sequencing reactions were run on a polyacrylamide sequencing gel. The gel was transferred to 3M paper and placed in a cassette in contact with x-ray film. After exposure the film was developed and the sequences read.

P1GA1.

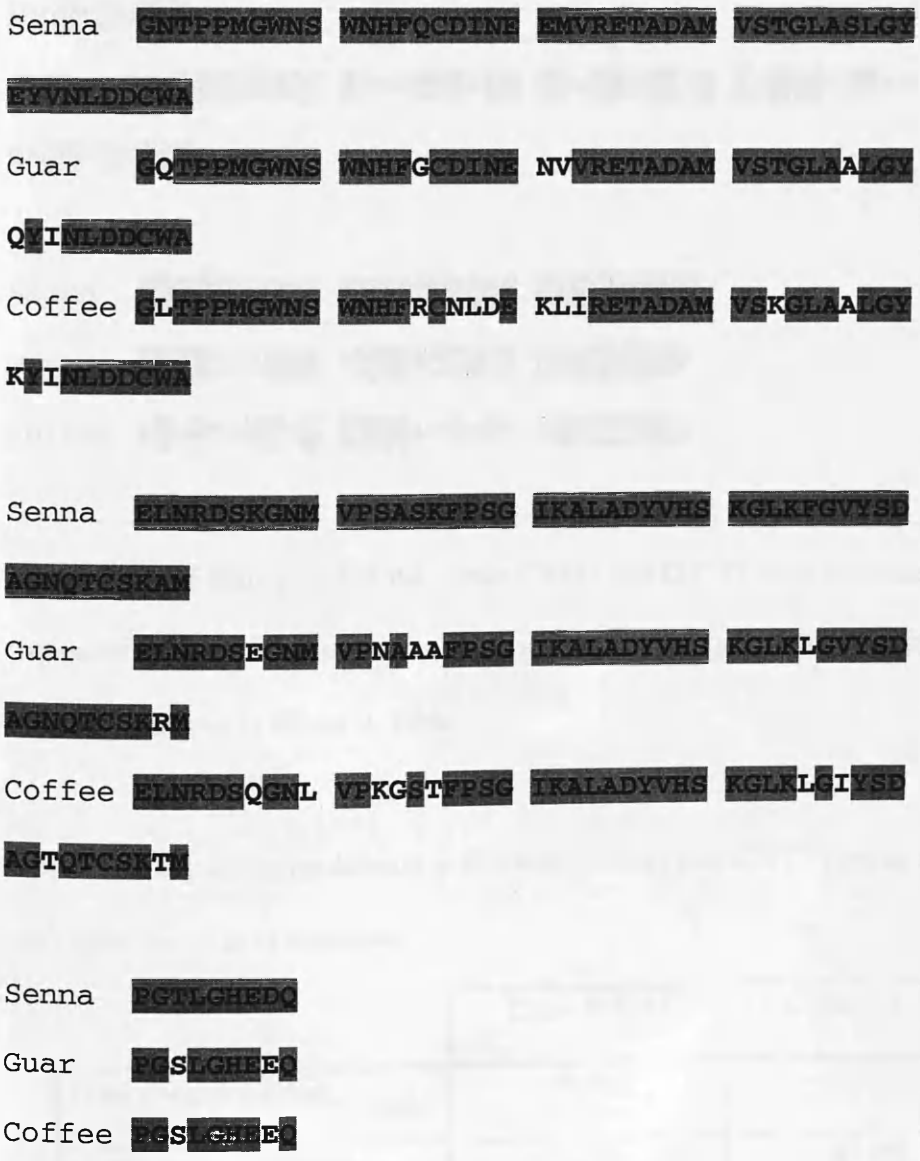
gggaacacgc cgccgatggg atggaatagc tggaaccact ttcagtgtga
cattaatgag gagatggttc gagaaacagc tgatgcaatg gtgtcaacgg
gtcttgcatc tttggggtac gaatacgtca atttagatga ttgctgggct
gaacttaacc gagactctaa gggaaatatg gttcctagtg cttcaaaatt
tccttcagga attaaggctc tggctgatta tgttcatagc aaaggattga
agtttggggt ttattctgat gctggaaacc aaacatgcag taaagctatg
cctggtacac taggacatga ggaccaa

GS1T7

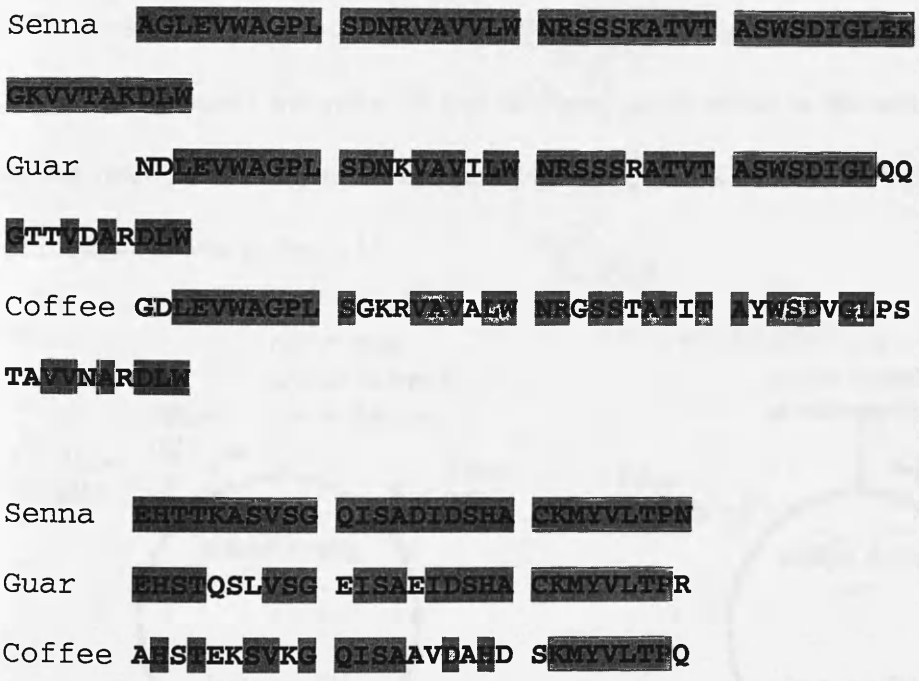
gctggtttgg aggtttgggc aggtcctctg agtgataaca gagtggcagt
ggtgttgtgg aatagaagtt catcaaaagc tactgtgact gcatcttggt
ctgacatagg gcttgaaaaa ggaaaggtgg tcactgcaaa agatttatgg
gagcacacta caaaagcatc agtttcagga caaaatttct gcagatatag
attcacatgc ttgtaagatg tatgttctga ctcccaatta aggcagacag
gaaggtgatg

The DNA sequences of the clones were translated into the corresponding amino acid sequences and then compared to the known sequences of coffee and guar.

The P1GA1 clones amino acid sequence compared to those of Guar and Coffee.



The GS1T7 clones amino acid sequence compared to those of Guar and Coffee.



The amino acid translations of the clones P1GA1 and GS1T7 were compared with the sequences of guar and coffee α-galactosidases. The percentage identities were calculated and this is shown in Table 13.

Table 13. The percentage identity of the clones P1GA1 and GS1T7 compared to guar and coffee amino acid sequences.

	Clone P1GA1	Clone GA1T7
Guar α-galactosidase.	86.5%	77.2%
Coffee α-galactosidase.	79.5%	65.8%

In order to ascertain the whether clone GA1T7 contained the full 3' untranslated region in was initially necessary to work out the direction of insertion of the GS1T7 PCR product. The was carried out by PCRing the clone using primers GS1, that anneals to the insert and either M13 or M13(rev) which anneal to the vector adjacent to the multiple cloning site. The possible arrangements of the GS1T7 insert and primers are shown in figure 45.

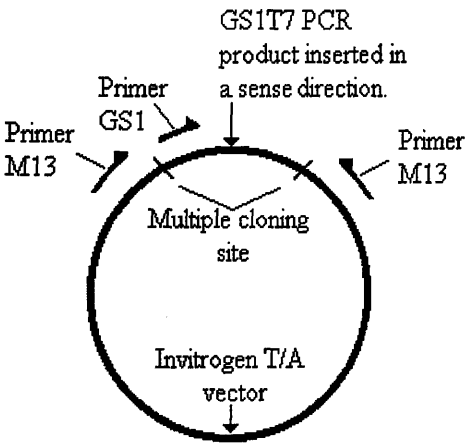


Figure 45A

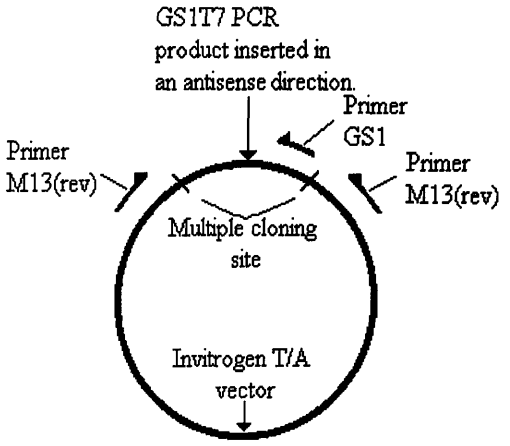


Figure 45B

Figure 45. The arrangements of vector, PCR insert and primers that would produce a product after a PCR reaction. Figure 42A shows that GS1T7 inserted in a sense direction will produce a PCR product with primers GS1 and M13. Figure 42B shows that if the GS1T7 was inserted in an antisense direction, primers GS1T7 and M13(rev) would produce a PCR product.

The PCR reaction was carried out and the results are shown in figure 46.

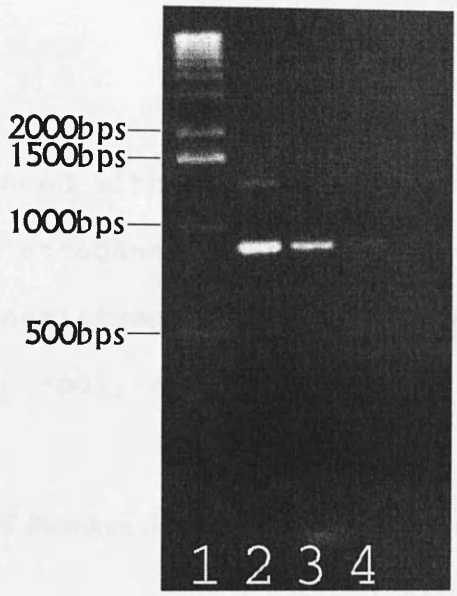


Figure 46. The orientation of clone GS1T7. Lane 1 contains 1kb markers. Lane 2 is a positive control using GS1T7. Lane 3 shows the results of the PCR using primers GS1 and M13. Lane 4 shows the result of GS1 and M13 (rev).

Figure 46 shows that a PCR product is not produced with GS1 and M13(rev) but is with GS1 and M13 therefore by referring to figure 45 it is possible to deduce that the GS1T7 clone was inserted in the sense direction. The clones were grown up in shaken culture and midipreped using the Invitrogen kit. The success of the midipreps was confirmed using restriction digests and running the digested fragments on an agarose gel as before. The success of the midipreps in producing more of the GS1T7 clone allowed further sequencing reactions (using M13 primer) to ascertain whether clone GS1T7 contained the full 3' untranslated region. The sequencing of the GS1T7 clone produced the sequence shown below.

GS1T7 sequenced with M13 primer:

taattaaataa gatagaaatt gcctgagatt tctatcatta tttgtattat

tggagcattg agatttgagac cttttgattt caattcaata attataccgc

aaatgtttct ct-poly-a-tail.....

The clone GS1T7 therefore does contain the entire 3' untranslated region and poly-A-tail.

Southern Blotting.

Southern blotting was carried out to discover whether the Senna genome contains only the germinating α-galactosidase gene already isolated or whether a number of α-galactosidase like genes are present. Two probes were used in the Southern blotting, the GS1T7 probe which is the least conserved and should indicate only the germinating α-galactosidase and the P1GS1 probe which is the more conserved and should pick out any α-galactosidase like gene.

Genomic Senna DNA was isolated according to methods section and the resulting preparation was measured spectrophotometrically. The amount of DNA isolated is shown in Table 14.

Table 14. The volume, concentration and yield of DNA produced from the genomic DNA preparations of 5.34g of young senna leaves.

Concentration of DNA μg μl^{-1}	Volume of DNA prep μl	Yield of DNA in μg
0.43	1000	430

The DNA was digested and run on two 0.8% agarose gels. These gels are shown in figure 47.

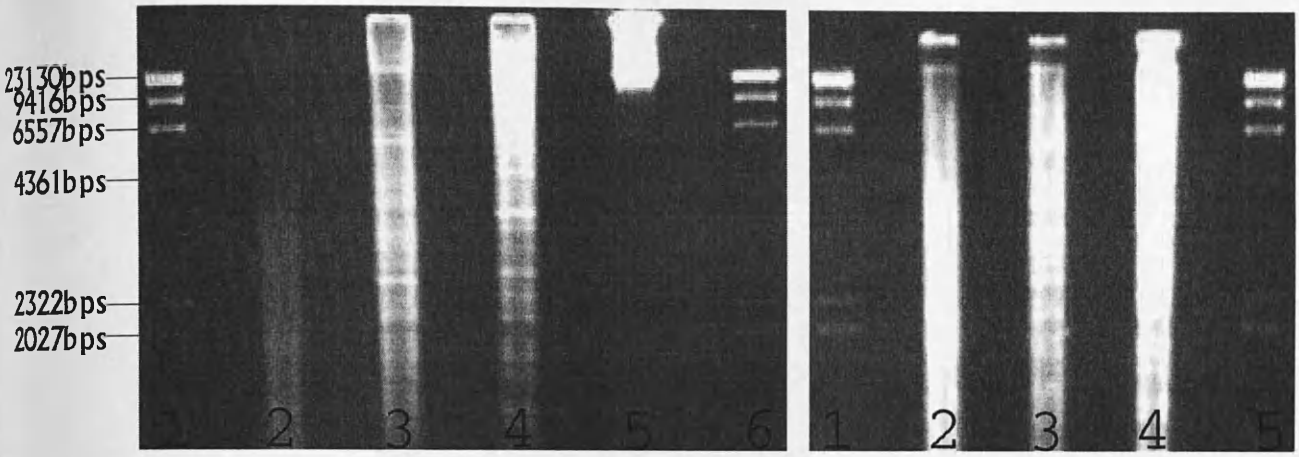


Figure 47. The digested genomic DNA. Gel 1 shows in the first and last lanes HindIII markers, in lane 2 the EcoRI digest, in lane 3 the HindIII digest and in lane 4 the BamHI digest. Lane 5 contains a undigested genomic DNA standard. Gel 2 is identical apart from containing no standard undigested genomic DNA.

The agarose gels were blotted onto Hybaid N+ membranes. After the blotting the agarose gels were re-imaged and it was shown that all the DNA had transferred.

The probes for blotting (the PCR products of P1GA1 and GS1T7) were prepared and labelled with ^{32}P as described in the methods section. Both of the blots were probed (at the same stringency level) and used to expose film in a x-ray cassette equipped with Phosphor intensifying screens. The films were developed and the blot probed with the less conserved GS1T7 probe is shown in figure 48. The blot probed with the more conserved P1GA1 probe is shown in figure 49.

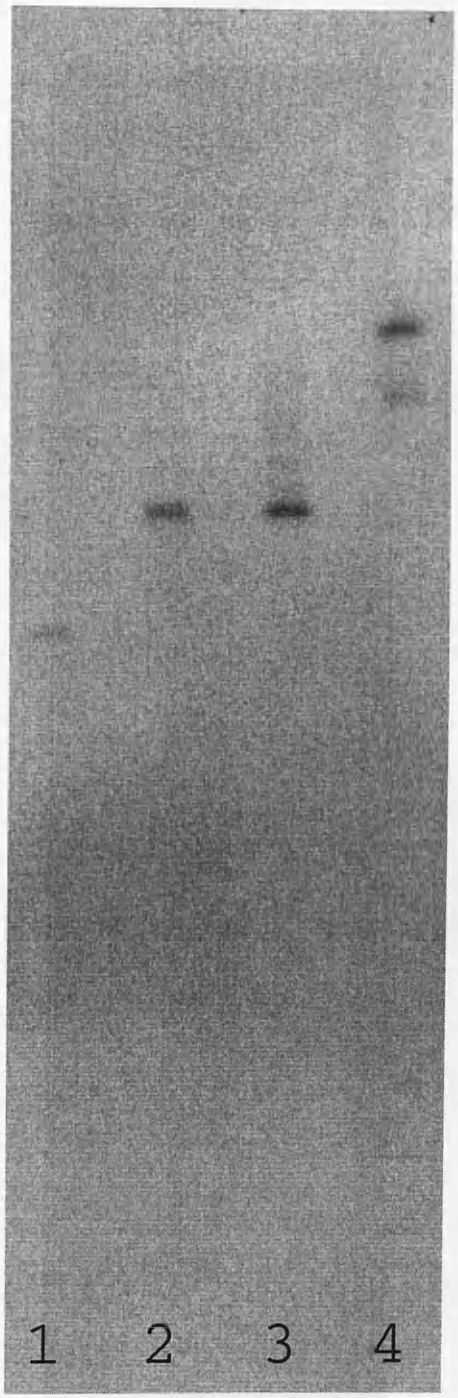


Figure 48. The blot of the senna genome probed with the GS1T7 probe. Lane 1, senna genomic DNA cut with EcoR1; lane 2, senna DNA cut with HindIII; lane 3, senna DNA cut with BamHI, lane 4, senna DNA not cut with restriction enzymes. The blot shows only one strongly labelled band in each lane, corresponding presumably to the germinating α-galactosidase.

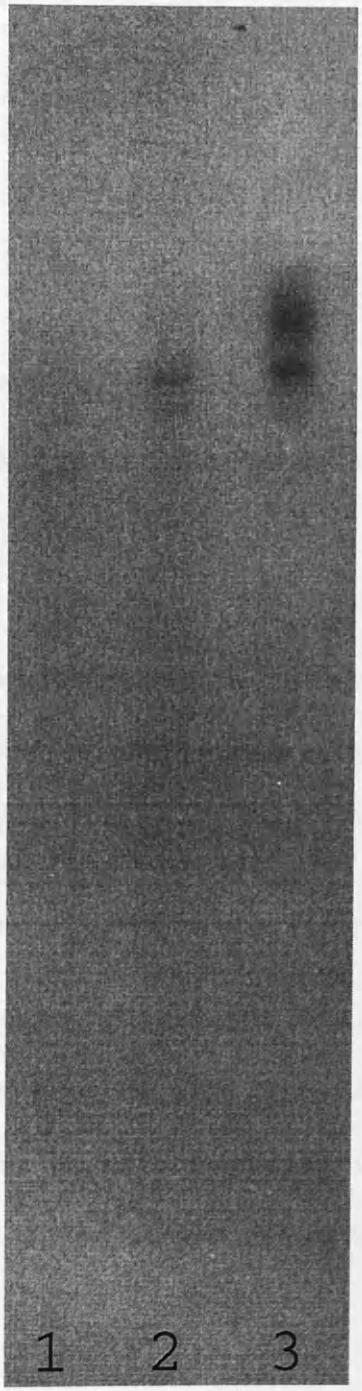


Figure 49. The blot of the senna genome probed with the P1GA1 probe. Lane 1, senna genomic DNA cut with EcoRI; lane 2, senna DNA cut with HindIII; lane 3, senna DNA cut with BamHI. A number of bands in each lane have been labelled.

The results of the blotting therefore shows that there is probably more than one α-galactosidase gene in the Senna genome.

Further preparation of RNA from maturing *senna* endosperms and PCR reactions.

RNA was prepared from 140 endosperms using the modified Lopez-Gomez method as before. The resulting RNA preparation was measured spectroscopically. The amount of RNA produced is shown in Table 15.

Table 15. The volume, concentration and yield of RNA produced from the modified Lopez-Gomez preparations of 140 maturing Senna endosperms

Concentration of total RNA $\mu g\ \mu l^{-1}$	Volume of RNA prep µl	Yield of RNA in µg
0.02	25	0.48

The 1st strand of cDNA was produced as before. The RNA remaining in the cDNA was then entirely removed using the alkali degradation method described in the methods section. PCRs were then set up as before and the products of these reactions were run on an agarose gel. The gel is shown in figure 50.

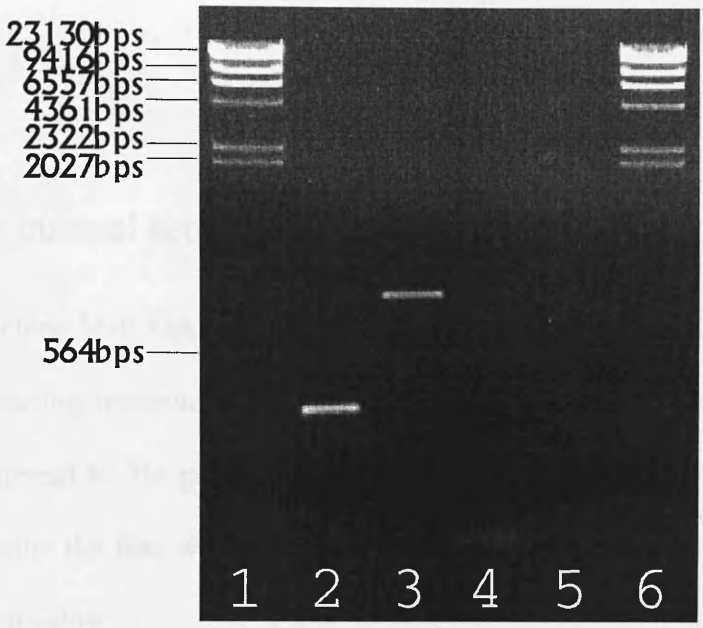


Figure 50. The result of the PCR of Alkali cleaned maturing 1st strand cDNA. Lanes 1 and 6 contain HindIII markers. Lane 2 contains the product of primers P1 and T7. Lane 3 the products of P1 + GA1. Lane 4 the products of GS1 + T7. Lane 5 contains a no primer negative control.

Figure 50 shows that PCR products were produced with primers P1+T7 and P1+GA1. However, the P1T7 product has a size that is too low to correspond with the known size of the α-galactosidase. The P1GA1 product was the size expected and was therefore cloned using the Invitrogen T/A kit.

The manual sequencing of M(aturing)-P1GA1.

The clone M-P1GA1 was prepared and sequenced using the M13(rev) primer. The sequencing reactions were run on a polyacrylamide sequencing gel. The gel was transferred to 3M paper and placed in a cassette in contact with x-ray film. After exposure the film was developed and the sequences read. The sequence obtained is shown below.

Sequence of clone M(aturing)-P1GA1

gggaacacgc cgcccatggg atggaatagc tggaaccact ttcagtgtga
cattaatgag gagatggttc gagaaacagc tgatgcaatg gtgtcaacgg
gtcttgcatc tttggggtac gaatacgtca atttagatga ttgctgggct
gaacttaacc gagactctaa gggaaatatg gttcctagtg cttcaaaatt
tccttca

The sequence of the M(aturing)-P1GA1 clone was compared to the sequence of P1GA1 and the percentage identity was calculated. This is shown in Table 16

Table 16. The percentage identity of the clone M(aturing)-P1GA1 from maturing senna endosperms compared to the clone P1GA1 from germinating senna endosperms.

	Clone M-P1GA1
Clone P1GA1.	100%

To isolate more of the maturing α-galactosidase cDNA further RNA preparations were carried out. Figure 51 shows the RNA prepared from 160 maturing endosperms complete with testa (it was intended that the testa RNA would co-precipitate the endosperm RNA).

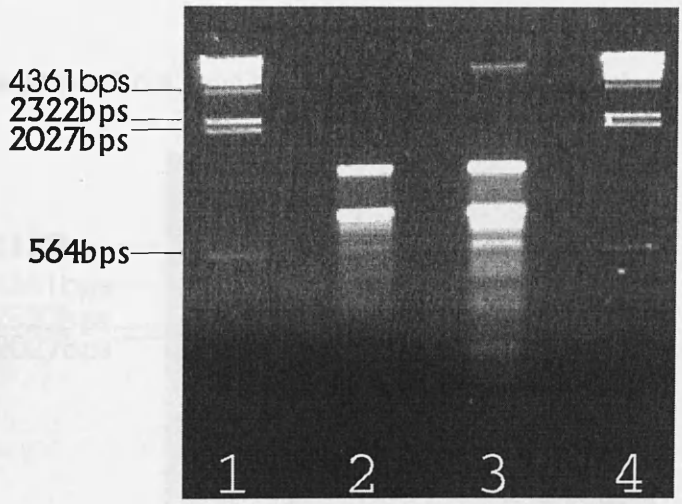


Figure 51. The RNA isolated from maturing endosperms (complete with testas). Lanes 1 and 4 contain HindIII markers. Lanes 2 and 3 contain the isolated maturing RNA; lane 2 shows the RNA post DNase treatment and lane 3 shows the RNA pre DNase treatment.

The concentration of the RNA preparation was measured spectrophotometrically as before. the concentration and yield is shown in Table 17.

Table 17. The volume, concentration and yield of RNA produced from the modified Lopez-Gomez preparations of 160 maturing senna endosperms with testas.

Concentration of total RNA µg μl^{-1}	Volume of RNA prep µl	Yield of RNA in µg
0.68	30	20.4

Table 17 shows that by including the testas with the endosperms the greatest yield of RNA was produced

Also isolated was RNA from the testas only. This RNA is shown in figure 52.

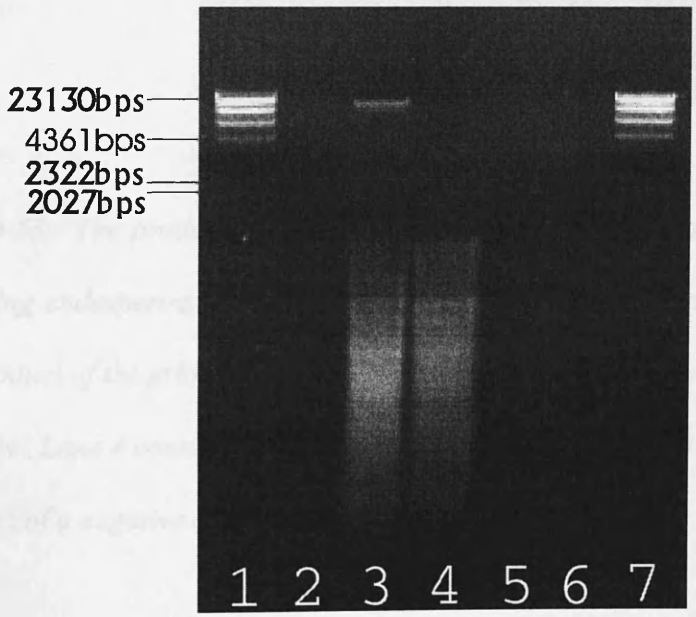


Figure 52. The RNA isolated from maturing testas only. Lanes 1 and 7 contain HindIII markers. Lanes 3 and 4 contain the isolated testa RNA; lane 3 shows the RNA pre DNase treatment and lane 4 shows the RNA post DNase treatment.

The maturing endosperm+testa and testa only cDNAs were split into 2 aliquots. Aliquots were then either alkali treated or not before being used as templates in PCRs. The results of the PCRs are shown in figures 53, 54 and 55.

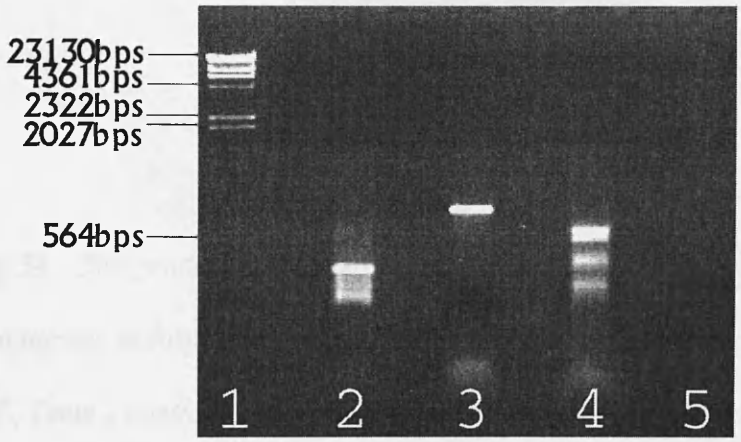


Figure 53. The products produced by PCR using alkali treated cDNA isolated from maturing endosperms + testas. Lane 1 contains HindIII markers. Lane 2 contains the product of the primers P1 and T7. Lane 3 contains the product of the P1 and GA1 primers. Lane 4 contains the product of the primers GS1 and T7. Lane 5 contains the product of a negative control that used no primers.

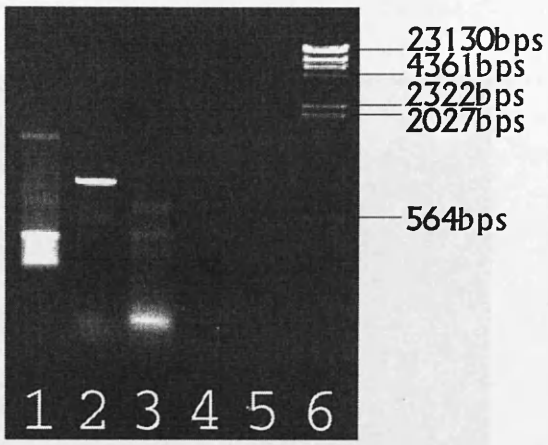


Figure 54. The products produced by PCR using non-alkali treated cDNA isolated from maturing endosperms + testas. Lane 1 contains the product of the primers P1 and T7. Lane 2 contains the product of the P1 and GA1 primers. Lane 3 contains the product of the primers GS1 and T7. Lane 4 contains the product of a negative control that used no primers. Lane 6 contains HindIII markers

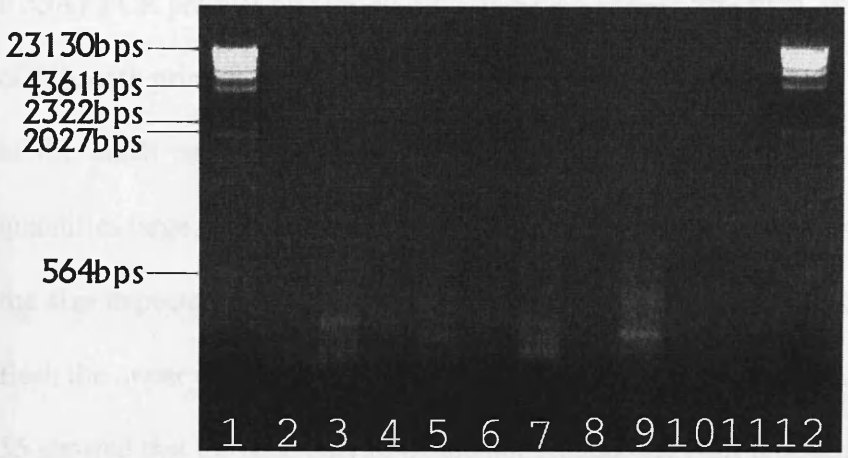


Figure 55. The products produced by PCR using alkali treated and non-treated cDNA isolated from maturing testas only. Lanes 1 and 12 contain HindIII markers. Lanes 3 to 6 contain the PCR products of the alkali treated cDNA template. Lane 3 contains the product of the primers P1 and T7. Lane 4 contains the product of the P1 and GA1 primers. Lane 5 contains the product of the primers GS1 and T7. Lane 6, a no primers negative control. Lanes 7 to 10 contain the PCR products of the non-alkali treated cDNA template. Lane 7 contains the product of the primers P1 and T7. Lane 8 contains the product of the P1 and GA1 primers. Lane 9 contains the product of the primers GS1 and T7. Lane 10 contains the product of a negative control that used no primers.

Figures 53 and 54 both show a P1T7 band that was too faint for cloning. Again the P1GA1 PCR product was produced in large quantities. The PCR of non-alkali treated cDNA with primers GS1 and T7 produced only very faint products. However the use of the alkali removal of RNA techniques have produced GS1T7 PCR products in quantities large enough to be cloned. Of note is that the product of the GS1T7 PCR of the size expected to be the 3' end of the α-galactosidase seems to be a double band. Both the upper and lower GS1T7 products were therefore excised for cloning. Figure 55 showed that the testa only RNA did not produce the PCR products produced by the endosperm and testa prep.

Cloning and sequencing of the M(aturing)-GS1T7 T(op) and B(ottom) PCR products.

The M-GS1T7-T and M-GS1T7-B PCR products were Invitrogen T/A cloned in accordance with the methods section. Colonies of the INVαF' were picked and cultured for plasmid preparations and restriction digests. Two clones (one each of M-GS1T7-T and M-GS1T7-B) were chosen and sequenced in both directions using the M13 and M13(rev) primers. The sequencing reactions were carried out by Unilever using an automatic sequencing machine. The sequences obtained are shown below.

M-GS1T7-T (5' sequence)

ttgcggtgaa ccaagatagt ctaggagtcc aaggaaagaa ggtgaaaagt
gatgctggtt tggaggtttg ggcaggtcct ctgagtgata acagagtggc
agtggtgttg tggaatanaa gttcatcaaa agctactgtg actgcatctt
ggtctgacat agggcttgaa aaaggaaagg tggtcactgc aaaagattta
tgggagcaca ctacaaaagc atcagtttca ggacaaattt ctgcagatat
agattcacat gcttgtaaga tgtatgttct gactcccaat taaggcagac
aggaaggtga tgaaagccc

M-GS1T7-B (5' sequence)

ttgcggtgaa ccaagatagt ctaggagtcc aaggaaagaa ggtgaaaagt
gatgctggtt tggaggtttg ggcaggtcct ctgagtgata acagagtggc
agtggtgttg tggaatanaa gttcatcaaa agctactgtg actgcatctt
ggtctgacat agggcttgaa aaaggaaagg tggtcactgc aaaagattta
tgggagcaca ctacaaaagc atcagtttca ggacaaattt ctgcagatat
agattcacat gcttgtaaga tgtatgttct gactcccaat taaggcagac
aggaaggtga tgaaagccc

M-GS1T7-B (3' sequence)

tcttgtttga catagggctg gaaaaaggaa agtggttcac tNcaaaagat

ttatgggagc acactacaaa agcatcagtt tcaggacaaa tttctgcaga

tatagattca catgcttgta agatgtatgt tctgactccc aattaaggca

gacaggaagg tgatgaaagc caaggtttta aggaagagaa atacaatcca

aggattcaaa gaaggatgga gaaaataaac atggagttta tttttcaata

agaaatatat agaaataatt aaataagata gaaattgcct gagatttcta

tcattatttg tattattgga gcattgagat ttgagacctt ttgatttcaa

ttcaataatt ataccgcaaa tgtttctct-poly-a-tail.

A comparison of the maturing GS1T7 top and maturing GS1T7 bottom 5' sequences shows 100% similarity. A comparison of the 3' ends of the clones could not be made because the M-GS1T7-T 3' sequence was disrupted due to the fact that the 1st strand cDNA synthesis had included a very long stretch of poly A tail the sequencing of which used up all the Stop-A nucleotide.

The sequence of the clone M-GS1T7-B shows that it contains a full 3' untranslated region. The sequence of the M-GS1T7-B clone was compared to the sequence of (germinating) GS1T7 and the percentage identity was calculated. This is shown in Table 18.

Table 18. The percentage identity of the clone M(aturing)-P1GA1 from maturing senna endosperms compared to the clone P1GA1 from germinating senna endosperms.

	Clone M-GS1T7
Clone GS1T7.	100%

The sequences of the clones obtained from the maturing and germinating α-galactosidase cDNAs were therefore identical.

In an attempt to discover the reason for the difference in size of M-GS1T7-T and M-GS1T7-B further PCRs were carried out and a further maturing clone (number 14) was produced using the primers P1 and T7. This clone was sequenced from both ends and the data obtained are shown below:-

14-forward-seq

gggaacacgc cgccgatggg atggaatagc tggaaccact ttcagtgtga
cattaatgag gagatggttc gagaaacagc tgatgcaatg gtgtcaacgg
gtcttgcatc tttggggtac ggatacgtca atttagatga ttgctgggct
gaacttaacc gagactctaa gggaaatatg gttcctagtg cttcaaaatt
tccttcagga attaaggctc tggctgatta tgttcatagc aaaggattga
agtttggggt ttattctgat gctggaaacc aaacatgcag taaagctatg
cctggatcac taggacatga ggaccaagat gcaaaaacat ttgcttcctg
ggggggttgat ttcttgaagt atgacaattg taaccacnat gatntaagcc
caagaaatag gtatncaaaa atgtctgaag ctctancaaa tctggaaggc
aatctnttct ctatgtgtaa tgggatcaaa aacctgcctn tggccaaant
gtggant

14-reverse-seq

catcaagcaa atttaacatt tattataatc nccataatta nggaaaatat

agagaancat tngcggnata attatngaat ngaaatcaaa nggnctcaaa

tctcaangct ccaataatnc aaataangat agaaatctca ggcaattnct

atcttattta attatttcta tatattnctt atngaaaaat aancnccang

ttnatttnct ccatccttct ttgaatcctn ggatgggatt tctcttcctt

aaaacctngg ctttcatcac cttccngtcn gccttaattg ggagtcagaa

catncatctt acaagcatgn gaatctatat ctgcagaaat ttggcctgaa

actgatgctt ttgnaggggg cncccataaa tcttttgcag ggaccacctt

tccttttca agccctatgn cagaccagat gcagnccngn agctttgggg

aactttnttc ncaacaccnt ggccttggat cctaaggacn gccaaccct

By aligning the sequence of the 3' ends of M-GA1T7-B and P1T7-clone-14 it is possible to see that P1T7-14 is identical to M-GA1T7-B except that it contains an extra 48 bases. This is the reason that a difference in the size of the relatively small GS1T7 clones could be observed.

Although the sequences corresponding to the region of the Senna α-galactosidases between the N terminal amino acid and the end of the poly-A tail had been isolated for both germinating and maturing isolated cDNAs, the 5' untranslated region of cDNAs from either maturing or germinating endosperms had not been isolated. To obtain these sequences a number of techniques were used.

Inverse polymerase chain reaction of senna genomic DNA.

Genomic Senna DNA was isolated and restriction digested (as described in the methods). The digested DNA was then aliquoted into 0.5, 0.05 and 0.005µg amounts. The DNA aliquots were then mixed with a ligation mix and incubated at 14°C overnight. Inverse PCR was then carried out using primers (GA2 and GS3) designed using the sequences previously obtained from the P1GA1 and M-P1GA1 clones. The result of the PCR is shown in figure 56.

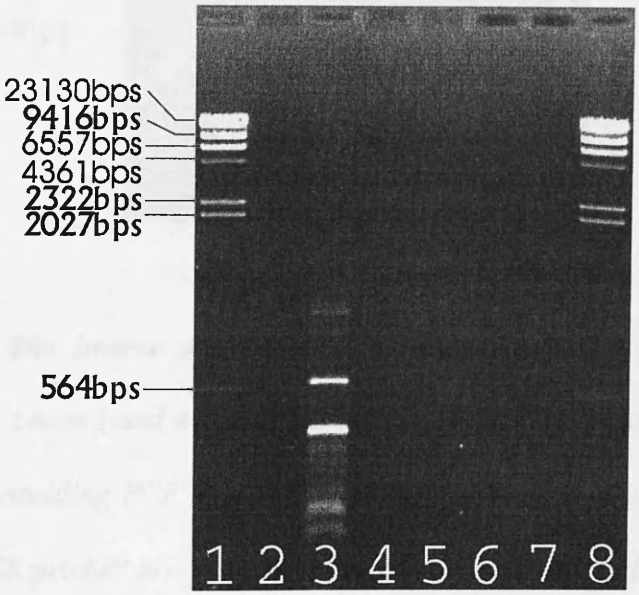


Figure 56. The result of the inverse PCR. Lanes 1 and 8 contain HindIII markers. Lane 3, 5 and 7 contain the inverse PCR products (using high, medium and low DNA concentrations respectively).

Figure 56 shows that the high-DNA inverse PCR produced a number of possible α-galactosidase upstream cDNAs. In order to discover which PCR product corresponds to the primers most closely further PCRs were carried out with increased stringency.

The inverse PCRs were performed with annealing temperatures of 55°C and 60°C. The results of these PCRs are shown in figure 57.

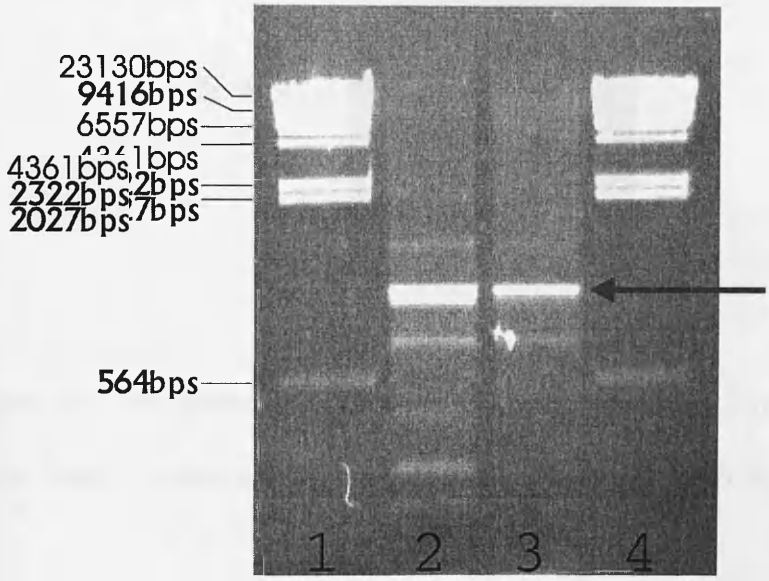


Figure 57. The inverse PCR of high concentration DNA at two annealing temperatures. Lanes 1 and 4 contain HindIII markers. Lane 2 contains the products of the 55°C annealing PCR. Lane 3 contains the products of the 60°C annealing PCR. The PCR product produced in the greatest amount at the higher stringency is indicated.

The higher stringency inverse PCR picked out one particular PCR product, to confirm that this was the 5' end of the α-galactosidase DNA a further PCR was carried out using the nested primers GA3 and GS2. If the previous PCR product was the 5' end of the α-galactosidase then the result of this PCR should be a slightly smaller PCR band. Figure 58 shows the result of the PCR

Figure 58. The confirmation inverse PCR using the nested primers GA3 and GS2. Lanes 1 and 3 contain HindIII markers. Lane 2 contains the PCR products.

Figure 58 shows that a faint band was produced using primers GA3 and GS2. However this band represents a molecule of too low a weight to be the product of the same gene that the band in figure 55 is a product of. Due to the fact that the primers GA2 + GS3 and GA3 + GS2 are amplifying different genes it is impossible to state if either is the α-galactosidase.

Final inverse PCRs were attempted using other restriction enzymes to cut the genomic DNA (these were Taq1, Mbo1 and EcoR1). The reactions produced no PCR products.

Single sided polymerase chain reaction of senna genomic DNA.

Genomic senna DNA was again isolated and restriction digested with EcoR1, HindIII and BamHI. The digested DNA was then added to EcoR1, HindIII and BamHI linearised pBluescript-KS plasmid. The 3 mixes were split into two aliquots, the first aliquot had DNA-ligase added and the second aliquot was used as a no ligation control (and therefore had no DNA-ligase added). The ligation reactions were carried out at 14°C overnight. Single sided PCR was then carried out using primers T7 and GA3. The result of the PCR is shown in figure 59.

Figure 59. The result of the single sided PCR. Lanes 1 and 8 contain HindIII markers. Lane 2 contains the PCR product of the ligated EcoRI digest. Lane 3 contains the PCR product of the non- ligated EcoRI digest control. . Lane 4 contains the PCR product of the ligated HindIII digest. Lane 5 contains the PCR product of the non- ligated HindIII digest control. Lane 6 contains the PCR product of the ligated BamHI digest. Lane 7 contains the PCR product of the non- ligated BamHI digest control.

Figure 59 shows that three faint and one bright PCR products were produced using the EcoRI ligated template. This PCR was repeated using a greater amount of template and the bright PCR product was isolated and cloned using the Invitrogen T/A cloning kit.

The clone was sequenced in both directions providing the nucleic acid sequences shown overleaf:-

Gupj-for-cot.seq

aaggatccgt cgacatccgc agggaccatc gcaatgcttt gttttaatta

aacagtcgga ttccccttgt ccgtaccagt tctgagtcga ctgttcgacg

cccggggaag ggaccccgaa gggcccgttc ccagtccgtc ccccggccgg

cacgcggcga cccgctctcg ccgcgggagc agctcgagca gtccaccgac

agccgacggg ttcggggctg ggacccccgt gcccagtcct cagagccaat

ccttttcccg aggttacgga tccattttgc cgacttccct tgcctacatt

gttccatcga ccagaggctg ttcaccttgg agacctgatg cggttatgag

taccaccggg ccttgggaag cacttcggtc cttccggatt tnaagggccg

ccggggcgca ccggaaacac gcgactgcgg tgctntttca gccgntggac

cttactncgg ctgaaccgtt cnaggtggca agcttttaaa cagaaaanaa

acttttccaa ggccccgcca nttttcggat cctaantttc cgtaaccncn

cttccgttaa gaattaacca

Gupj-rev-cut.seq

agggtccatg ngaacngnnc ttgcncatgg gttantcgat cctaaaaaac
gggggaancc cntntganan cgngcancac ncntncttnn aaagggaanc
gggttaaaat tcctgaaccg ggacgtggcg gctgacngga acgttaggga
gtccngaaac ctnggcgggg gccnngggaa aanttctntt ttntggttaa
naacctgccc nccctggaaa cnggttaacc ngnggtaggg nccnacnggt
ngaaaaacnc cccacgttnc gtggngnccg gtgccccccc ggnggccctt
taaaatnccg aggaccnant gccttncacc cccggnngtn ctnanaaccn
cataaggttt ncaangggaa nnaccctntg gncnanggaa caanttnncc
cngggaaatt ccnccaaang gatccntacc ntgggaaaag attngctttt
aaantggccc cggggntccn cccnnaaccc ctgttncgng gnctttcaac
tttnccng

The sequence was compared to the known sequence of the region of the α-galactosidase adjacent to the primer sites. **No** similarities were found.

Further PCRs were carried out using the ligated EcoRI DNA digest with nested primers. The first reaction used primers T7 and GA1. The results of this PCR along with the T7 only, GA1 only and non-ligated template controls are shown in figure 60.

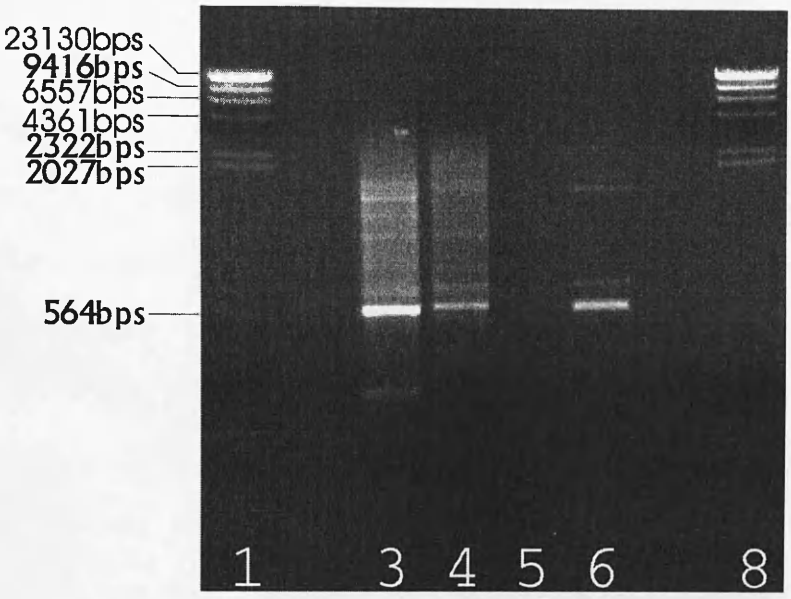


Figure 60. The initial PCR. Lanes 1 and 8 contain HindIII markers. Lane 3 contains the PCR products of T7 and GA1. Lane 4 contains the T7 only control. Lane 5 contains the non-ligated control. Lane 6 contains the GA1 only primer control.

Figure 60 shows that the T7 + GA1 reaction has produced many products however many of these are also produced with the single primer controls. The T7 + GA1 PCR product was then used as a template in a second PCR using the primers KS and GA3 (found adjacent to the T7 and GA1 sites on the α-galactosidase and pBluescript). The result of this PCR along with the KS only, GA3 only and non-ligated template controls are shown in figure 61

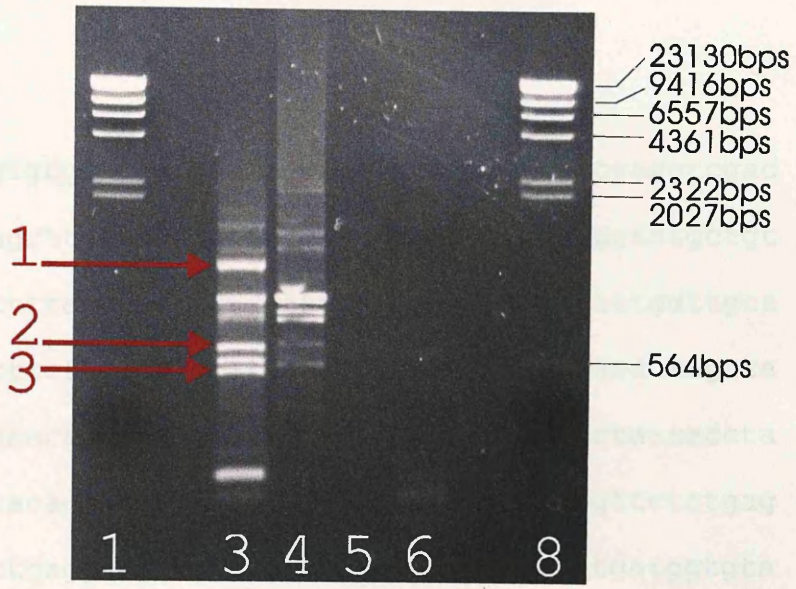


Figure 61. The second PCR. Lanes 1 and 8 contain HindIII markers. Lane 3 contains the PCR products of KS and GA3. Lane 4 contains the KS only control. Lane 5 contains the GA3 only primer control. Lane 6 contains the non-ligated control. The bands labelled 1, 2 and 3 are present in the KS + GA3 PCR product and not in the single primer controls.

Figure 61 shows that the sequential PCRs of the EcoRI digested and ligated DNA has produced three PCR bands that are not produced using single primer controls. These three bands were isolated and cloned using the Invitrogen T/A kit.

The clones were sequenced in both directions providing the nucleic sequences shown below:-

ssup1for

gcggccgcca gtgtgatgga tatctgcaga attcggcttt cgaggtcgac

ggtatcgata agcttgatat cgaattcaga taaacgtgta ggtatgctgt

ttaaatttcg ttttagcaat agtggttctg cgttcaaaaa tttgattgct

gaaaattaaa cttcttaatt gcataatagg ttcatgggtt acataagtta

aaaaaaatat aaacaaccaa aagcaagcaa tcctacaaag ctaaaaccta

aaaggctccc tacacctcac tattgaccgt cctggtcctc gttctctgag

taatcttcct ctgaggaaga cggatcctgc tgctgccccg tgatggtgta

ggccctgtcc ttgaaagtgt anacacatct gctcatgcgg gcctgtctcc

gttctaacct anattggngt cgcacgatct cccctgtcct ggaaaaatct

tgtatccngg aagattcttg gtgtcctggg cagttctgat gggcccantc

ttcctaactc cnatctacgg agtatgtgcg gatgaggtcc caaagncggt

atttttcctc c

ssup1rev

gaattcggct tagcaatcat ctaaattgac gtccatggta aatcgtcatc

acccattccc caagcaatac atcaaggaga aaaaagaccc gttcttccat

caaaacatcg caactgcagt cctgaccctc ctgacacacc gacatggcaa

gagaaacgag accctggtca gagagtacta catcgccgcc aagatggcag

gatccctcga gatgactatt agaggggtga gggtnccagc cctagatgca

aactatgtca atgaacactt ccagctcccc gttctacctg tatgcgacta

tgagaggtnc atcctagaca gatcaganga gtnccctagc gacnagatca

tancagctct gggaaccgct agggcggctg ggaaggtctg gaaatcccca

ggaaaatttg aacaacag

ssup2for

gcggccgcca gtgtgatgga tatctgcaga attcggctta gcaatcatct

aaattgacgt actttgtttc cactttaatg gttcgcagct cttggtagaa

ccaatgtagt ctttaggctt tctggtatga agctgtgttg tgcacttata

gatgaatgcg tgttgatgaa gaatgatgga atggaatgga agacaggtga

cttgggcagt acattgtaat atgaaaatga gtgcagatca aaatggaaaa

tgaccgtgtt taaggggaaa atatanaatt atactatttg atctgcagtt

ctgacttgga tcaagtatac ataacagata gtgtatttac tatttggata

atatatattt gtgtttctan gaagtaggaa tagcaataat aaaggaaaca

gaacatggta gttgtccacc gggctgtctt tgaaccctga atccttctgt

ntccagaagc aaaaccccct atgaaggtac cctactatga accnttgggg

agttccntta tgagaatntc tnatcantca cttcancccc accccaaacc

aatccn

ssup2rev

tcgaggtcga tggtatcgat aagcttgata tcgaattcat gtatacttca

tcactaaact gtaaacttcc ccaaattgct tcactagtaa gggtaacctt

caataaagag gtttctgcct cctggataca gaaacgactc aagcgttcaa

agacagcccg gtggacaact accatgttct gtttccttta ttattgctat

tcctacttcc tagaaacaca aatatatatt atccaaatag taaatacact

atctgttatg tatacttgat ccaagtcaga actgcagatc aaatagtata

attctatatt ttccccttaa acacggtcat tttccatttt gatctgcact

cattttcata ttacaatgta ctgcccaagt cacctgtctt ccattccatc

attcttcatc cacgcattct ctataaaa

ssup3rev

cgccaagctt ggtaccgagc tcggatccac tagtaacggc cgccagtgtg

ctggaattcg gctttcgagg tcgacggtat cgataagctt gatatcgaat

tcatgtatac ttcatcacta aactgtaaac ttccccaaat tgcttcacta

gtaagggtaa ccttcaataa agaggtttct gcgtcctgga tacagaaacg

actcaagcat tcaaagacag cccggtggac aactaccatg ttctgtttcc

tttattattg ctattcctac ttcctagaaa cacaaatata tattatccaa

atagtaaata cactatttgt tatgtatact tgatccaagt cagaactgca

gatcaaatag tataattcta tattttcccc tcaaacacgg tcattttcca

ttttgatctg cactcatttt catattacaa tgtactgccc aagtcacctg

tcttccattc cattccatca ttcttcatca acacgcattc atctataagt

gcacaacaca gcttcatacc agaaagccta aaaatacttg gtctncagaa

cnggaacntt aagtggaaca atnctcatta aagatgnnag ccaatcngcn

g

ssup3for

Cattacgact cactatcggg gcgaattggg ccctctagat gcatgctcga

gcggccgcca gtgtgatgga tatctgcaga attcggctta gcaatcatct

aaattgacgt actttgtttc cactttaatg gttcgcagct cttggtagaa

ccaatgtagt ctttaggctt tctggtatga agctgtgttg tgcacttata

gatgaatgcg tgttgatgaa gaatgatgga atggaatgga agacaggtga

cttgggcagt acattgtaat atgaaaatga gtgcagatca aaatggaaaa

tgaccgtgtt tgaggggaaa atatagaatt atactatttg atctgcagtt

ctgacttgga tcaagtatac ataacaaata gtgtatttac tatttggata

atatatattt gtgtttctag gaagtaggaa tagcaataat aaaggaaaca

gaac

The single sided sequences (above) were compared to the previously isolated sequence of the Senna α-galactosidase gene and also to each other. The UP2 and UP3 sequences can be seen to be very similar however neither UP1, 2 or 3 is related to senna α-galactosidase. These results can be seen in Table 19.

Table 19. The percentage similarity of the clones UP1, UP2 and UP3 compared to each other and the previously sequenced α-galactosidase.

	UP1	UP2	UP3	α-galactosidase
UP1		23.6	19.1	20.8
UP2	23.6		91.9	19.1
UP3	19.1	91.9		20.1
α-galactosidase.	20.8	19.1	20.1	

5' Rapid Amplification of cDNA Ends - Polymerase Chain Reactions.

5' RACE-PCR was carried out in accordance with the methods of sections. RNA was isolated from maturing endosperms as before. The RNA was split into two aliquots, one was DNase treated and the other was not. The two RNA samples were used to produce 1st strand cDNA using either random hexamer primers or an α-galactosidase gene specific primer (GA1). The 1st strands were cleaned up using the Hybaid clean-up kit and then poly-A tailed using the TdT enzyme. A 2nd strand of cDNA was produced using the oligo-dT-RiRoT7 primer. PCRs were then carried out using primers Ro and GA1 for the random primed cDNA and primers Ro and GA3 for the gene specific primed cDNA the result of this PCR is shown in figure 62.

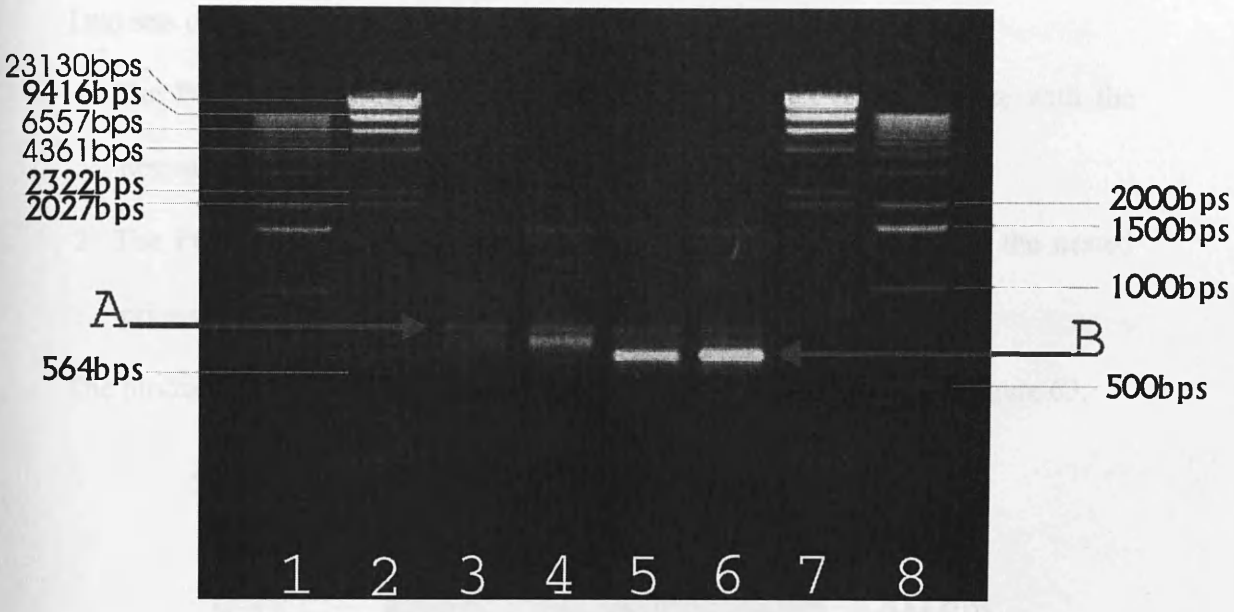


Figure 62. The PCR products of the 5' RACE-PCR. Lanes 1 and 8 contain 1kb markers. Lanes 2 and 7 contain HindIII markers. Lane 3 contains the products of the DNased random primed 1st strand. . Lane 4 contains the products of the non-DNased random primed 1st strand. Lane 5 contains the products of the DNased gene specific primed 1st strand. . Lane64 contains the products of the non-DNased gene specific primed 1st strand.

Figure 62 shows that the PCRs on random hexamer cDNA and on gene specific cDNA have produced a number of products of the same size. However, two bright bands, labelled A and B in figure 62 are of slightly different sizes. This is what would be expected, as the primers Ro and GA1 used on the hexamer produced 2nd strand PCR are located further apart on the α-galactosidase gene than the Ro and GA3 used on the gene specific produced 2nd strand α-galactosidase-oligo-dT-RiRoT7.

Two sets of confirmation PCRs were carried out using:-

1. The PCR products of the gene-specific primed cDNA as a template with the nested primers Ri + GA2 and P1 + GA2.
2. The PCR products of the random primed cDNA as a template with the nested primers Ri + GA3 and a negative control deionised H_2O + GA2.

The products of these PCRs were run on an agarose gel and are shown in figure 63.

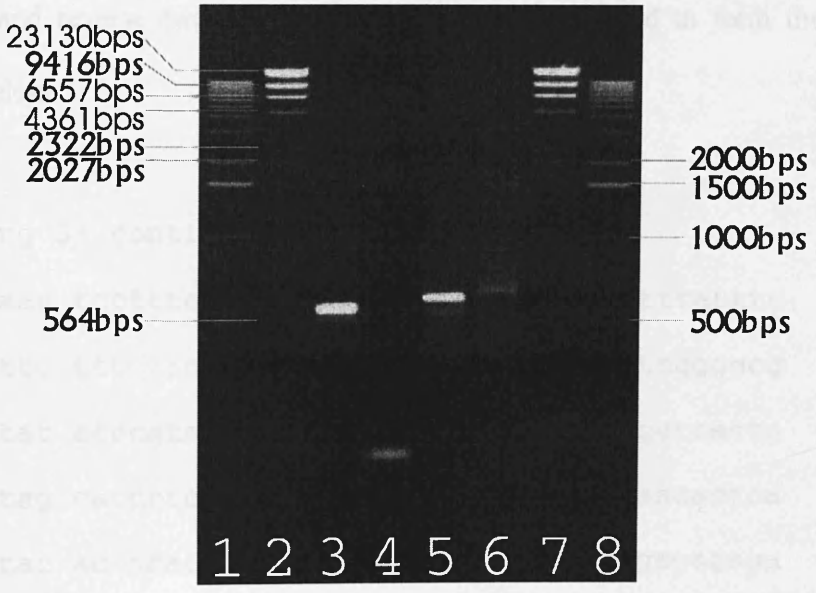


Figure 63. The confirmation PCRs of the 5' RACE PCRs. Lanes 1 and 8 contain 1kb markers. Lanes 2 and 7 contain HindIII markers. Lane 3 contains the products of Ri + GA2. Lane 4 contains the products of P1 + GA2. Lane 5 contains the products of Ri + GA3. Lane 6 contains the products of the deionised H_2O + GA2 negative control.

Figure 63 shows that the confirmation PCRs have all produced products of the size calculated using the known locations of the primer sites along the sequenced α-galactosidase. The negative control has produced a faint PCR band, probably due to the transfer of primers from the previous PCR that was used as the template.

The band B from the PCR products of the gene specific primed cDNA was isolated and cloned using the invitrogen T/A kit. The PCR product was sequenced in a forward and reverse direction and the two sequences joined to form the sequence shown below.

Maturing 5′ contig.

ngccctaaa tcctttaaa anccnccgc ccaagctac ttttatttc

tcttctttc tttcttngg ttcttaacc tattatcag cttggggcg

ttggcctat atccatata tatccatat ncatacaaa agatcatta

tataagtag catcttcca cttcatcta tatatcata taacactca

cacgtatac actacacac acactcata tatatatat agagagaga

gagagagag agagattta tatagaaag aaatggaga aaatgatga

tgtgggcaa aggttgtgt tgngcttgt tttgggtct tgaangctt

ctaattgtt caggncgct tgttgaaca caattggca atgatcaca

acaacatcc atggaagac tacttcttg gaaatggac ttggaaaca

ctcctccca tgggatgga atagctgga accactttc agtgtgaca

ttaatgagg agatggttc gagaaacag ctgatgcaa tggtgtcaa

cgggtcttg catctttgg ggtacgaat acgtcaatt tagatgatt

gct

The 5' RACE PCR procedure was repeated using RNA isolated from germinating endosperms. The PCRs again produced products which were cloned and sequenced. The obtained sequence corresponding to the α-galactosidase germinating 5' untranslated region is shown below.

germinating 5' contig

accctaaanc ctttaaaagc cnccgcccaa gctactttta tttctcttct
ttctttcttt ggttcttaac tattatcagc ttggggcgtt ggcctatata
catatatatc catatccata ccaaagatca ttatataagg ngcatcttcc
acttcatcta tatatcatat aacactcaca cgtatacact acacacacac
tcatatatat atatatagag agagagagag agagaggaga tttatataga
aagaaatggg gaaaatgatg atgtgggcaa aggttgtgtt gngcttgttt
tgggtcttga angcttctaa ttgttcaggn cgcttgttga acacaattgg
caatgatcac aacaacatcc atggaagact acttcttgga aatggacttg
gaaacactcc tcccatggga tggaatagct ggaaccactt tcagtgtgac
attaatgagg agatggttcg agaaacagct gatgcaatgg tgttcaacgg
gtctgcatct ttggggtacg aatacgtcaa tttagatgat tgct

The 5' untranslated region clones of the maturing and germinating α-galactosidases were compared to each other. The region where the 5' clones overlapped the P1T7 clones were also compared. The resulting percentage identities are shown in Table 20.

Table 20. The percentage identity of the maturing and germinating 5' compared to each other and the previously sequenced P1T7 α-galactosidase clone.

	Maturing 5' clone	Germinating 5' clone	P1T7 clone
Maturing 5' clone		99.2	97.3
Germinating 5' clone	99.2		97.3
P1T7 clone	97.3	97.3	

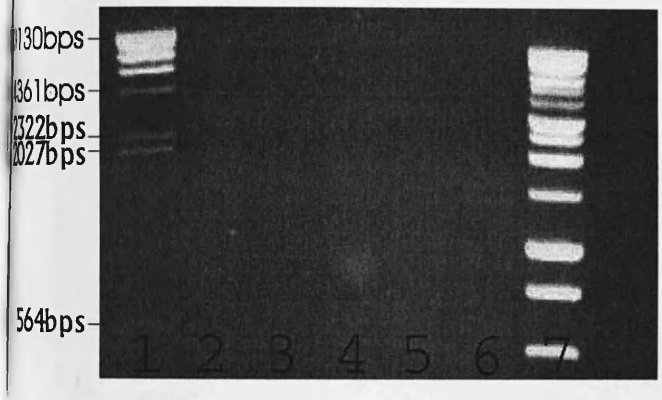
As can be seen from table 20 the maturing and germinating 5' UTR clones are essentially identical. The overlap between the 5' clones and P1T7 is identical apart from missmatches incorporated into the P1T7 clone through the use of the P1 primers that was of a degenerate nature.

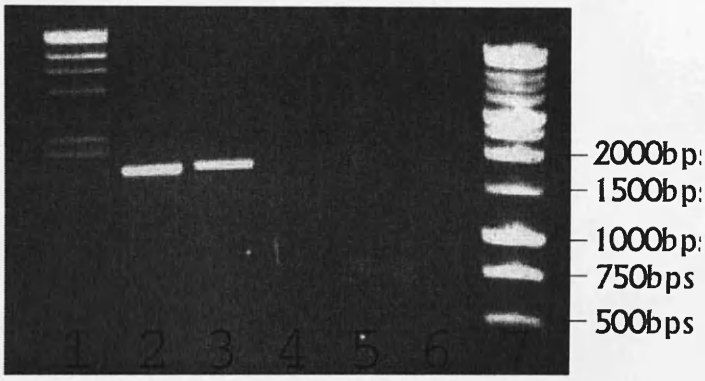
The isolation of Pfu proof-read long and short clones of maturing and germinating α-galactosidases.

Due to the fact that the extreme 5' and 3' ends of the long and short α-galactosidases from both maturing and germinating α-galactosidases were now known it was decided that a final proof reading PCR would be carried out to isolate full 5' to 3' clones for full sequencing.

Primers were designed to anneal to the extreme 5' end of the α-galactosidase (S1) and also to the 3' ends of the previously sequenced short and long versions of the cDNA (A1 and A3).

Maturing endosperm RNA was isolated as before and Dnase treated. 1st strand cDNA was produced using the oligo-dT-RiRo-T7 primer. PCRs were then set up using S1+A1 primers, S1+A3 primers or S1, A1 and A3 single primer controls. The PCRs were carried out in duplicate using Pfu and Taq. The results of the reactions can be seen in figure 64.





Gel A *Gel B*

Figure 64. Full length α-galactosidase PCRs. Gel A shows the results of the Pfu reactions and gel B shows the results of the Taq controls. Both gels contain in lane 1 HindIII markers and in lane 7 1kb markers. Lane 2 contains the results of the S1+A1 PCR. Lane 3 contains the results of the S1+A3 PCR. Lane 4 contains the S1 only single primer control, lane 5 contains the A1 only control and lane 6 the A3 only control.

Figure 64 shows that the Taq polymerase produced PCR products of the size expected of the short and long versions of the maturing α-galactosidase. However, the Pfu has produced no PCR produsts at all.

The Pfu PCRs were repeated using highly extended chain elongation times as shown below

Standard Taq PCR		Pfu PCR	
1 cycle of:-		1 cycle of:-	
50°C (annealing)	2 minutes	50°C (annealing)	10 minutes
72°C (cDNA extension)	10 minutes	72°C (cDNA extension)	20 minutes
then 30 cycles of:-		then 30 cycles of:-	
94°C (melting of strands)	1 minute	94°C (melting of strands)	1 minutes
50°C	1 minute	50°C	1 minutes
72°C	$1^1/_2$ minutes	72°C	5 minutes
then 1 cycle of:-		then 1 cycle of:-	
72°C	15 minutes	72°C	15 minutes

The results of the extended PCR are shown in figure 65.

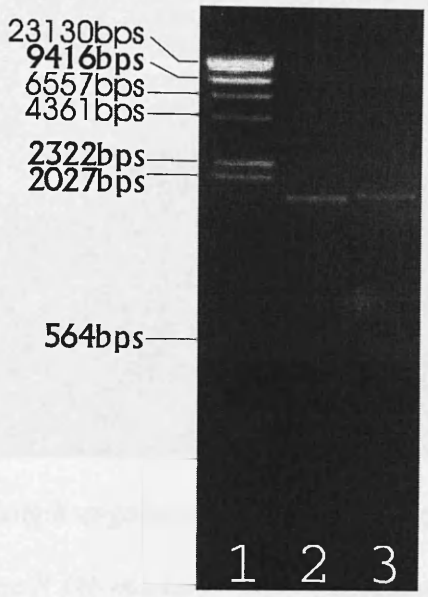


Figure 65. Full length α-galactosidase Pfu PCRs. Lane 1 contains HindIII markers. Lane 2 contains the results of the S1+A1 PCR and lane 3 contains the results of the S1+A3 PCR.

Figure 65 shows that the extended Pfu PCR has produced products of the size expected of the long and short α-galactosidases. The bands were excised and used in invitrogen T/A cloning. The resultant white colonies were checked using PCR screening. Midi-sized plasmid preparations were carried out using the positive colonies and these preparations were used for sequencing.

The Pfu PCRs were repeated to isolate full length clones from germinating endosperms. Germinating endosperm RNA was isolated as before and DNased. 1st strand cDNA was produced using the oligo-dT-RiRo-T7 primer. PCRs were then set up using S1+A1 primers, S1+A3 primers or S1, A1 and A3 single primer controls and the PCRs were carried out using Pfu with extended elongation times. The results of the reactions can be seen in figure 66

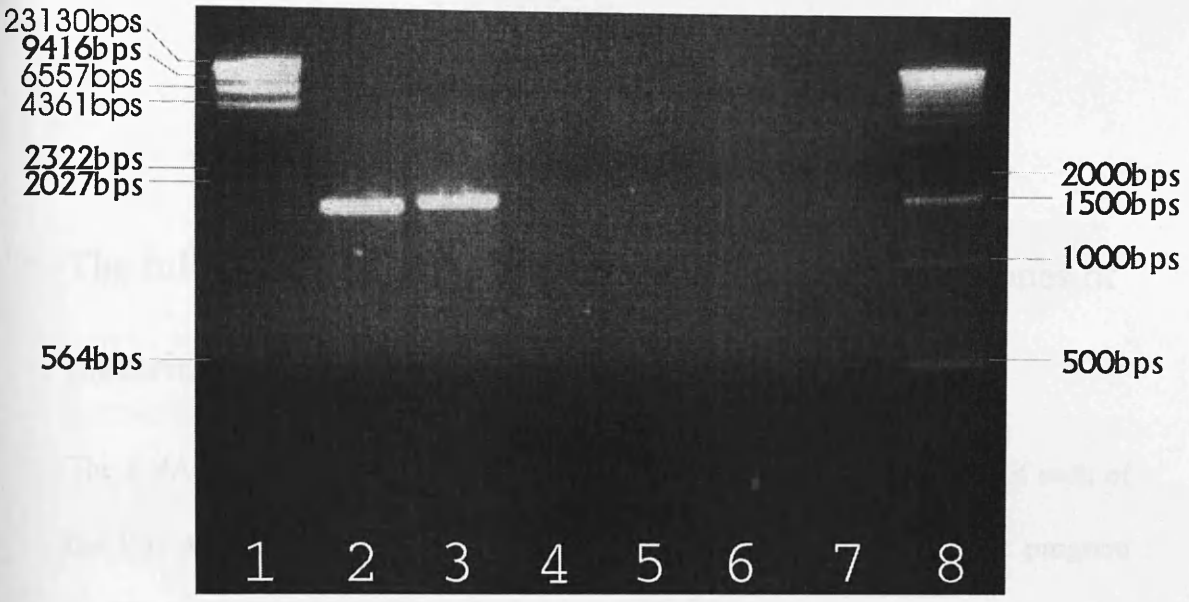


Figure 66. Full length α-galactosidase PCRs. The gel contains in lane 1 HindIII markers and in lane 8 1kb markers. Lane 2 contains the results of the S1+A1 PCR. Lane 3 contains the results of the S1+A3 PCR. Lane 4 contains the S1 only single primer control, lane 5 contains the A1 only control and lane 6 the A3 only control.

Figure 64 shows that the Pfu PCR has produced products of the size expected of the long and short α-galactosidases. The bands were excised and used in invitrogen T/A cloning. The resultant white colonies were checked using PCR screening and positives were used in plasmid preparations which were then feed into automatic sequencing reactions.

The full sequences of the Pfu proof-red long and short clones of maturing and germinating α-galactosidases.

The DNA sequences obtained by Unilever from the automatic sequencing of each of the Pfu produced α-galactosidase clones were analysed by the computer program SeqMan (DNAStar) and organised into single whole sequences (contigs). The sequences are shown below:-

Maturing alpha-galactosidase contig.

GCCCTAAATCCTTTAAAAGCCACCGCCCAAGCTACTTTTATTTCTCTTCTTTCTTTC

TTTGTTTCTTAACTATTATCAGCTTGGGGCGTTGGCCTATATACATATATATACATA

TACATACAAAAGATCATTATATAAGTAGCATCTTCCACTTCATCTATATATCATATA

ACACTCACACGTATACACTACACACACACTCATATATATATATATAGAGAGAGAGAG

AGAGAGAGAGATTTATATAGAAAGAAATGGAGAAAATGATGATGTGGGCAAAGGTTG

TGTTGTGCTTGTTTTGGGTCTTGAATGCTTCTAATTGTTCAGGTCGCTTGTTGAACA

CAATTGGCAATGATCACAACAACATCCATGGAAGACTACTTCTTGGAAATGGACTTG

GAAACACTCCTCCCATGGGATGGAATAGCTGGAACCACTTTCAGTGTGACATTAATG

AGGAGATGGTTCGAGAAACAGCTGATGCAATGGTGTCAACGGGTCTTGCATCTTTGG

GGTACGAATACGTCAATTTAGATGATTGCTGGGCTGAACTTAACCGAGACTCTAAGG

GAAATATGGTTCCTAGTGCTTCAAAATTTCCTTCAGGAATTAAGGCTCTGGCTGATT

ATGTTCATAGCAAAGGATTGAAGTTTGGGGTTTATTCTGATGCTGGAAACCAAACAT

GCAGTAAAGCTATGCCTGGATCACTAGGACATGAGGACCAAGATGCAAAAACATTTG

CTTCCTGGGGGGTTGATTTCTTGAAGTATGACAATTGTAACACCAATGATATAAGCC

CAAGAAATAGGTATCCAAAAATGTCTGAAGCTCTAGCAAATTCTGGAAGGGCAATCT

TCTTCTCTATGTGTGAATGGGGATCAGAAGACCCTGCACTATGGGCCAAAAGTGTGG

GAAATAGTTGGAGAACAACAGGAGATATTGAAGATAAGTGGGAAAGTATGGCATCTA

TTGCTGACCAAAATGACAAATGGGCATCTTATGCTGGACCTGGAGGGTGGAATGATC

CTGATATGCTTGAAGTTGGAAATGGAGGCATGACAACAGAAGAATATCGTTCTCATT

TTAGCATATGGGCATTAGCTAAGGCTCCTTTGTTGATTGGTTGTGATGTTCGATCAA

TGGATGGCGCAACATACGGACTGCTAAGCAACAAGGAAGTTATTGCAGTAAACCAAG

ACAGTCTAGGAGTCCAAGGAAAGAAGGTGAAAAGTGATGCTGGTTTGGAGGTTTGGC

CAGGTCCTCTGAGTGATAACAGAGTGGCAGTGGTGTTGTGGAATAGAAGTTCATCAA

AAGCTACTGTGACTGCATCTTGGTCTGACATAGGGCTTGAAAAAGGAAAGGTGGTCA

CTGCAAAAGATTTATGGGAGCACACTACAAAAGCATCAGTTTCAGGACAAATTTCTG

CAGATATAGATTCACATGCTTGTAAGATGTATGTTCTGACTCCCAATTAAGGCAGAC

AGGAAGGTGATGAAAGCCAAGGTTTTAAGGAAGAGAAATACAATCCAAGGATTCAAA

GAAGGATGGAGAAAATAAACATGGAGTTTATTTTTCAATAAGAAATATATAGAAATA

ATTAAATAAGATAGAAATTGCCTGAGATTTCTATCATTATTTGTATTATTGGAGCAT

TGAGATTTGAGACCTTTTGATTTCAATTCAATAATTATACCGCAAATGTTTCTCTAT

ATTTTCCATAATTATGGTGATTATAATAAATGTTAAATTTGCTTGATG

Germinating alpha-galactosidase contig.

TGCCCTAAATCCTTTAAAAGCCACCGCCCAAGCTACTTTTATTTCTCTTCTTTCTTT

CTTTGTTTCTTAACTATTATCAGCTTGGGGCGTTGGCCTATATACATATATATACAT

ATACATACAAAAGATCATTATATAAGTAGCATCTTCCACTTCATCTATATATCATAT

AACACTCACACGTATACACTACACACACACTCATATATATATATATATAGAGAGAGAGA

GAGAGAGAGAGATTTATATAGAAAGAAATGGAGAAAATGATGATGTGGGCAAAGGTT

GTGTTGTGCTTGTTTTGGGTCTTGAATGCTTCTAATTGTTCAGGTCGCTTGTTGAAC

ACAATTGGCAATGATCACAACAACATCCATGGAAGACTACTTCTTGGAAATGGACTT

GGAAACACTCCTCCCATGGGATGGAATAGCTGGAACCACTTTCAGTGTGACATTAAT

GAGGAGATGGTTCGAGAAACAGCTGATGCAATGGTGTCAACGGGTCTTGCATCTTTG

GGGTACGAATACGTCAATTTAGATGATTGCTGGGCTGAACTTAACCGAGACTCTAAG

GGAAATATGGTTCCTAGTGCTTCAAAATTTCCTTCAGGAATTAAGGCTCTGGCTGAT

TATGTTCATAGCAAAGGATTGAAGTTTGGGGTTTATTCTGATGCTGGAAACCAAACA

TGCAGTAAAGCTATGCCTGGATCACTAGGACATGAGGACCAAGATGCAAAAACATTT

GCTTCCTGGGGGGTTGATTTCTTGAAGTATGACAATTGTAACACCAATGATATAAGC

CCAAGAAATAGGTATCCAAAAATGTCTGAAGCTCTAGCAAATTCTGGAAGGGCAATC

TTCTTCTCTATGTGTGAATGGGGATCAGAAGACCCTGCACTATGGGCCAAAAGTGTG

GGAAATAGTTGGAGAACAACAGGAGATATTGAAGATAAGTGGGAAAGTATGGCATCT

ATTGCTGACCAAAATGACAAATGGGCATCTTATGCTGGACCTGGAGGGTGGAATGAT

CCTGATATGCTTGAAGTTGGAAATGGAGGCATGACAACAGAAGAATATCGTTCTCAT

TTTAGCATATGGGCATTAGCTAAGGCTCCTTTGTTGATTGGTTGTGATGTTCGATCA

ATGGATGGCGCAACATACGGACTGCTAAGCAACAAGGAAGTTATTGCAGTAAACCAA

GACAGTCTAGGAGTCCAAGGAAAGAAGGTGAAAAGTGATGCTGGTTTGGAGGTTTGG

CCAGGTCCTCTGAGTGATAACAGAGTGGCAGTGGTGTTGTGGAATAGAAGTTCATCA

AAAGCTACTGTGACTGCATCTTGGTCTGACATAGGGCTTGAAAAAGGAAAGGTGGTC

ACTGCAAAAGATTTATGGGAGCACACTACAAAAGCATCAGTTTCAGGACAAATTTCT

GCAGATATAGATTCACATGCTTGTAAGATGTATGTTCTGACTCCCAATTAAGGCAGA

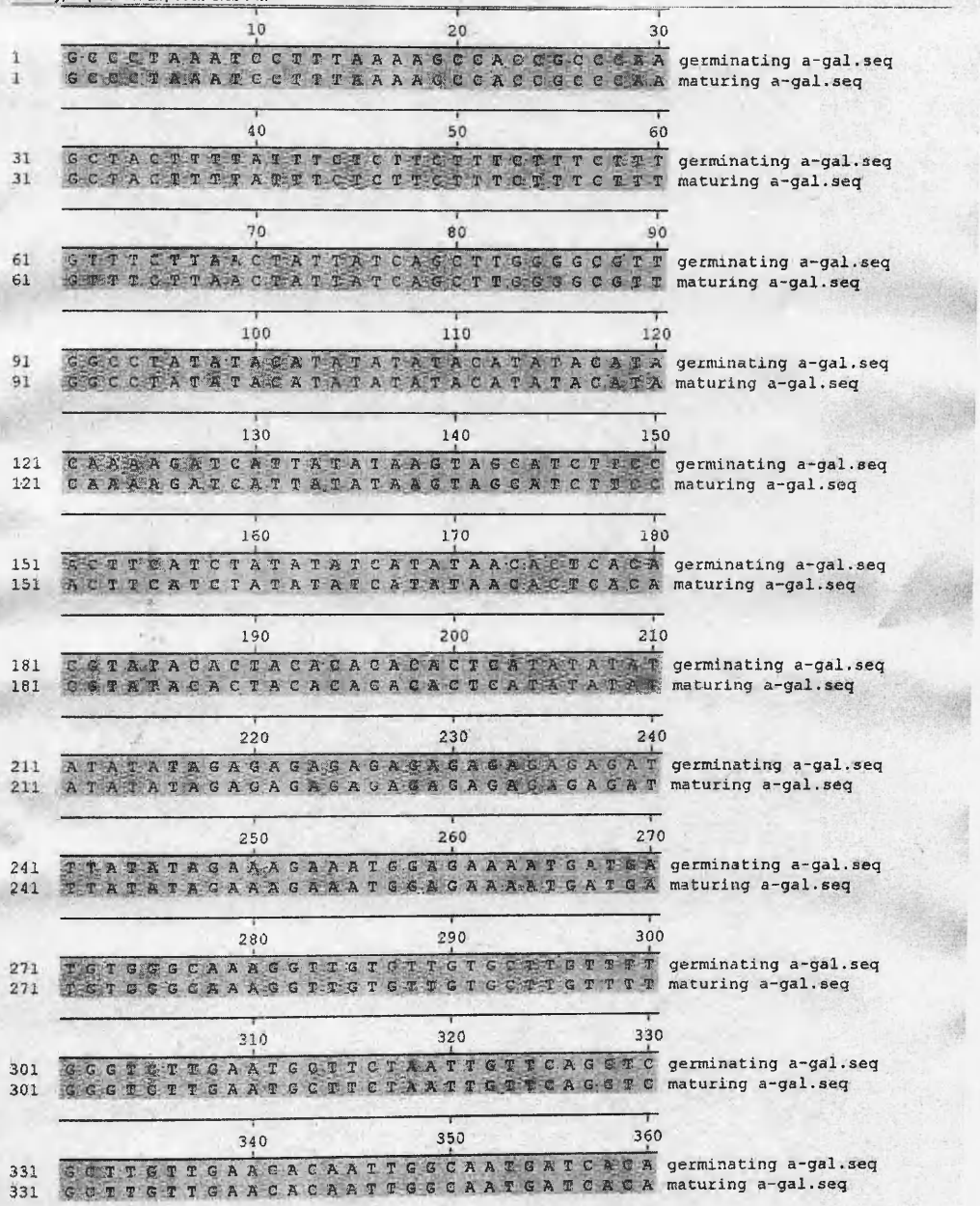
CAGGAAGGTGATGAAAGCCAAGGTTTTAAGGAAGAGAAATACAATCCAAGGATTCAA

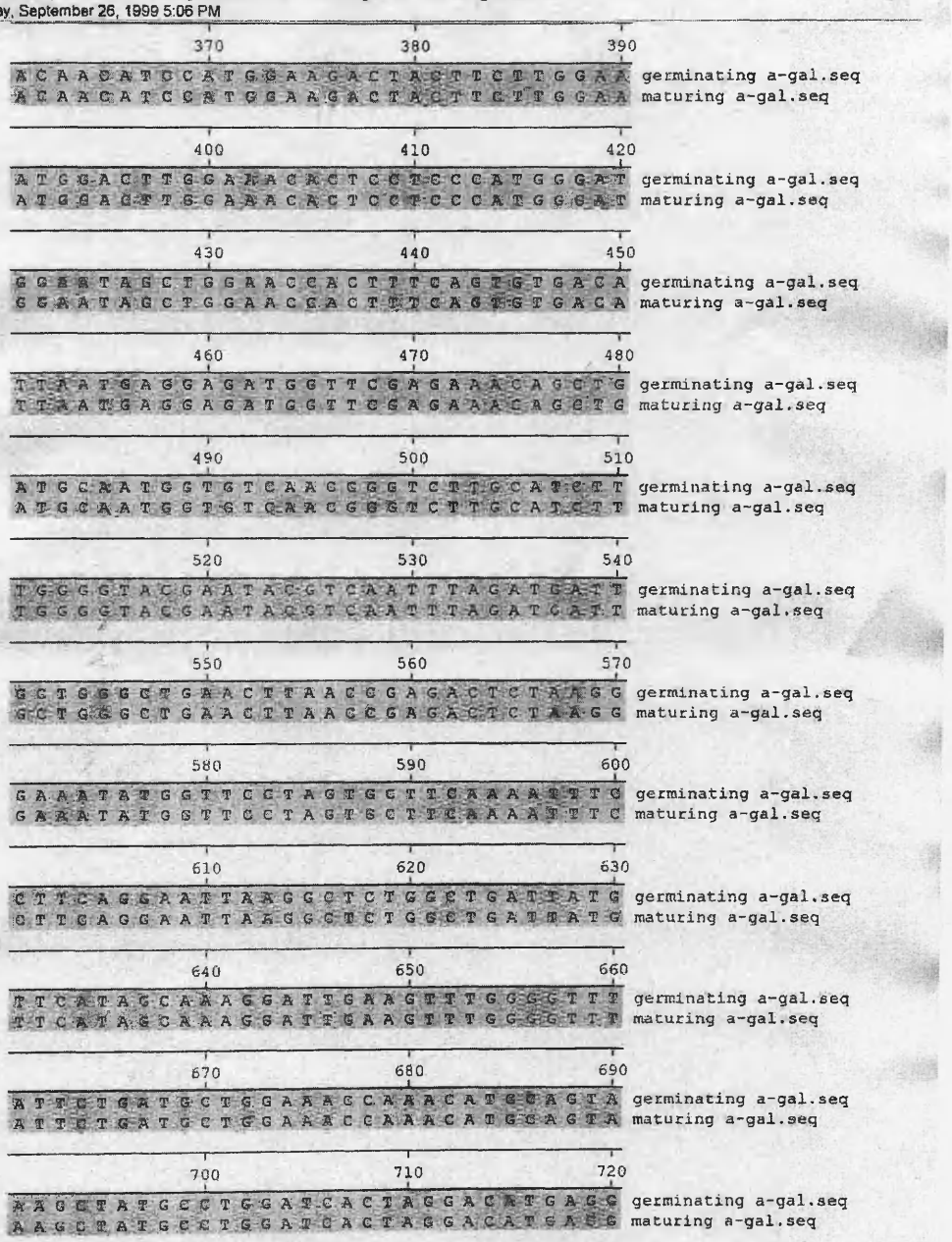
AGAAGGATGGAGAAAATAAACATGGAGTTTATTTTTCAATAAGAAATATATAGAAAT

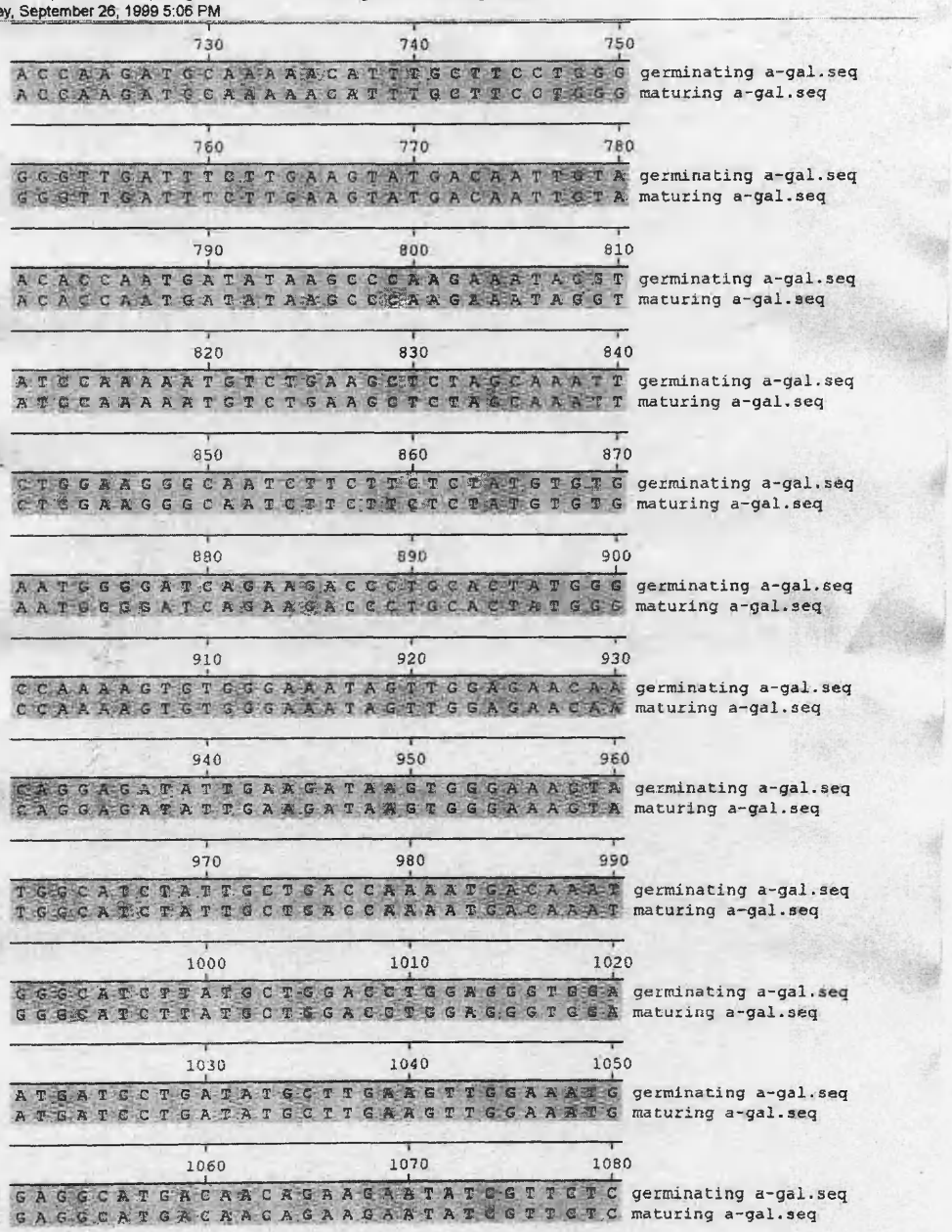
AATTAAATAAGATAGAAATTGCCTGAGATTTCTATCATTATTTGTATTATTGGAGCA

TTGAGATTTGAGACCTTTTGATTTCAATTCAATAATTATACCGCAAATGTTTCTCT

The two Senna α-galactosidase nucleotide sequences were then compared using the program MegAlign (Dnastar) which utilised the Clustal sequence alignment algorithm (Higgins and Sharp 1989). The alignments are shown overleaf.


10 20 30
1 germinating a-gal.seq
1 maturing a-gal.seq
40 50 60
31 germinating a-gal.seq
31 maturing a-gal.seq
70 80 90
61 germinating a-gal.seq
61 maturing a-gal.seq
100 110 120
91 germinating a-gal.seq
91 maturing a-gal.seq
130 140 150
121 germinating a-gal.seq
121 maturing a-gal.seq
160 170 180
151 germinating a-gal.seq
151 maturing a-gal.seq
190 200 210
181 germinating a-gal.seq
181 maturing a-gal.seq
220 230 240
211 germinating a-gal.seq
211 maturing a-gal.seq
250 260 270
241 germinating a-gal.seq
241 maturing a-gal.seq
280 290 300
271 germinating a-gal.seq
271 maturing a-gal.seq
310 320 330
301 germinating a-gal.seq
301 maturing a-gal.seq
340 350 360
331 germinating a-gal.seq
331 maturing a-gal.seq


370 380 390
361 ACAACATCCATGGAAGACTACTTCTTGGAA germinating a-gal.seq
361 ACAACATCCATGGAAGACTACTTCTTGGAA maturing a-gal.seq
400 410 420
391 ATGGACTTGGAAACACTCCTCCCATGGGAT germinating a-gal.seq
391 ATGGACTTGGAAACACTCCTCCCATGGGAT maturing a-gal.seq
430 440 450
421 GGAGTAGCTGGAACCACTTTCAGTGTGACA germinating a-gal.seq
421 GGAATAGCTGGAACCACTTTCAGTGTGACA maturing a-gal.seq
460 470 480
451 TTAATGAGGAGATGGTTCGAGAAACAGCTG germinating a-gal.seq
451 TTAATGAGGAGATGGTTCGAGAAACAGCTG maturing a-gal.seq
490 500 510
481 ATGCAATGGTGTCAACGGGTCTTGCATCTT germinating a-gal.seq
481 ATGCAATGGTGTCAACGGGTCTTGCATCTT maturing a-gal.seq
520 530 540
511 TGGGGTACGAATACGTCAATTTAGATGATT germinating a-gal.seq
511 TGGGGTACGAATACGTCAATTTAGATGATT maturing a-gal.seq
550 560 570
541 GCTGGGCTGAACTTAACGGAGACTCTAAGG germinating a-gal.seq
541 GCTGGGCTGAACTTAACGGAGACTCTAAGG maturing a-gal.seq
580 590 600
571 GAAATATGGTTCCTAGTGCTTCAAAATTTC germinating a-gal.seq
571 GAAATATGGTTCCTAGTGCTTCAAAATTTC maturing a-gal.seq
610 620 630
601 CTTCAGGAATTAAGGCTCTGGCTGATTATG germinating a-gal.seq
601 CTTCAGGAATTAAGGCTCTGGCTGATTATG maturing a-gal.seq
640 650 660
631 TTCATAGCAAAGGATTGAAGTTTGGGGTTT germinating a-gal.seq
631 TTCATAGCAAAGGATTGAAGTTTGGGGTTT maturing a-gal.seq
670 680 690
661 ATTCTGATGCTGGAAACCAAAACATGCAGTA germinating a-gal.seq
661 ATTCTGATGCTGGAAACCAAAACATGCAGTA maturing a-gal.seq
700 710 720
691 AAGCTATGCCTGGATCACTAGGACATGAGC germinating a-gal.seq
691 AAGCTATGCCTGGATCACTAGGACATGAGG maturing a-gal.seq


730 740 750
721 ACCAAGATGCAAAAACATTTGCTTCCTGGG germinating a-gal.seq
721 ACCAAGATGCAAAAACATTTGCTTCCTGGG maturing a-gal.seq
760 770 780
751 GGGTTGATTTCTTGAAGTATGACAATTGTA germinating a-gal.seq
751 GGGTTGATTTCTTGAAGTATGACAATTGTA maturing a-gal.seq
790 800 810
781 ACACCAATGATATAAGCCCAAGAAATAGGT germinating a-gal.seq
781 ACACCAATGATATAAGCCCAAGAAATAGGT maturing a-gal.seq
820 830 840
811 ATCCAAAAAATGTCTGAAGCTCTAGCAAATT germinating a-gal.seq
811 ATCCAAAAAATGTCTGAAGCTCTAGCAAATT maturing a-gal.seq
850 860 870
841 CTGGAAGGGCAATCTTCTTCTCTATGTGTG germinating a-gal.seq
841 CTGGAAGGGCAATCTTCTTCTCTATGTGTG maturing a-gal.seq
880 890 900
871 AATGGGGATCAGAAGACCCTGCACTATGGG germinating a-gal.seq
871 AATGGGGATCAGAAGACCCTGCACTATGGG maturing a-gal.seq
910 920 930
901 CCAAAAAGTGTGGGAAATAGTTGGAGAACAA germinating a-gal.seq
901 CCAAAAAGTGTGGGAAATAGTTGGAGAACAA maturing a-gal.seq
940 950 960
931 CAGGAGATATTGAAGATAAGTGGGAAAGTA germinating a-gal.seq
931 CAGGAGATATTGAAGATAAGTGGGAAAGTA maturing a-gal.seq
970 980 990
961 TGGCATCTATTGCTGACCAAAATGACAAAT germinating a-gal.seq
961 TGGCATCTATTGCTGACCAAAATGACAAAT maturing a-gal.seq
1000 1010 1020
991 GGGCATCTTATGCTGGACCTGGAGGGTGGA germinating a-gal.seq
991 GGGCATCTTATGCTGGACCTGGAGGGTGGA maturing a-gal.seq
1030 1040 1050
1021 ATGATCCTGATATGCTTGAAGTTGGAAAATG germinating a-gal.seq
1021 ATGATCCTGATATGCTTGAAGTTGGAAAATG maturing a-gal.seq
1060 1070 1080
1051 GAGGCATGACAACAGAAGAATATCGTTCTC germinating a-gal.seq
1051 GAGGCATGACAACAGAAGAATATCGTTCTC maturing a-gal.seq

```
                    1090                1100                1110
1081 ATTTTAGCATATGGGCATTAGCTAAGGCTC germinating a-gal.seq
1081 ATTTTAGCATATGGGCATTAGCTAAGGCTC maturing a-gal.seq

                    1120                1130                1140
1111 CTTTGTTGATTGGTTGTGATGTTCGATCAA germinating a-gal.seq
1111 CTTTGTTGATTGGTTGTGATGTTCGATCAA maturing a-gal.seq

                    1150                1160                1170
1141 TGGATGGCGCAACATACGGACTGCTAAGCA germinating a-gal.seq
1141 TGCATGGCGCAACATACGGACTGCTAAGCA maturing a-gal.seq

                    1180                1190                1200
1171 ACAAGGAAGTTATTGCAGTAAACCAAGACA germinating a-gal.seq
1171 ACAAGGAAGTTATTGCAGTAAACCAAGACA maturing a-gal.seq

                    1210                1220                1230
1201 GTCTAGGAGTCCAAGGAAAGAAGGTGAAAA germinating a-gal.seq
1201 GTCTAGGAGTCCAAGGAAAGAAGGTGAAAA maturing a-gal.seq

                    1240                1250                1260
1231 GTGATGCTGGTTTGGAGGTTTGGCCAGGTC germinating a-gal.seq
1231 GTGATGCTGGTTTGGAGGTTTGGCCAGGTC maturing a-gal.seq

                    1270                1280                1290
1261 CTCTGAGTGATAACAGAGTGGCAGTGGTGT germinating a-gal.seq
1261 CTCTGAGTGATAACAGAGTGGCAGTGGTGT maturing a-gal.seq

                    1300                1310                1320
1291 TGTGGAATAGAAGTTCATCAAAAGCTACTG germinating a-gal.seq
1291 TGTGGAATAGAAGTTCATCAAAAGCTACTG maturing a-gal.seq

                    1330                1340                1350
1321 TGACTGCATCTTGGTCTGACATAGGGCTTG germinating a-gal.seq
1321 TGACTGCATCTTGGTCTGACATAGGGCTTG maturing a-gal.seq

                    1360                1370                1380
1351 AAAAAGGAAAGGTGGTCACTGCAAAAGATT germinating a-gal.seq
1351 AAAAAGGAAAGGTGGTCACTGCAAAAGATT maturing a-gal.seq

                    1390                1400                1410
1381 TATGGGAGCACACTACAAAAGCATCAGTTT germinating a-gal.seq
1381 TATGGGAGCACACTACAAAAGCATCAGTTT maturing a-gal.seq

                    1420                1430                1440
1411 CAGGACAAATTTCTGCAGATATAGATTCAG germinating a-gal.seq
1411 CAGGACAAATTTCTGCAGATATAGATTCAC maturing a-gal.seq
```

Decoration 'Decoration #1': Shade (with black at 25% fill) residues that match germinating a-gal.seq exactly.

The DNA sequences of the maturing and germinating α-galactosidases can be seen to be 100% identical. The two forms of the α-galactosidase must therefore be transcribed from the same gene within the senna genome. Using the SWISS-PROT Internet resource both the Senna α-galactosidase DNA sequence was translated into amino acid sequence. The full amino acid sequence (including the signal sequences) are shown below.

Senna Occidentalis alpha-galactosidase amino acid sequence.

M E K M M M W A K V V L C L F W V L N A S N C S G R L L N

T I G N D H N N I H G R L L L G N G L G N T P P M G W N S

W N H F Q C D I N E E M V R E T A D A M V S T G L A S L G

Y E Y V N L D D C W A E L N R D S K G N M V P S A S K F P

S G I K A L A D Y V H S K G L K F G V Y S D A G N Q T C S

K A M P G S L G H E D Q D A K T F A S W G V D F L K Y D N

C N T N D I S P R N R Y P K M S E A L A N S G R A I F F S

M C E W G S E D P A L W A K S V G N S W R T T G D I E D K

W E S M A S I A D Q N D K W A S Y A G P G G W N D P D M L

E V G N G G M T T E E Y R S H F S I W A L A K A P L L I G

C D V R S M D G A T Y G L L S N K E V I A V N Q D S L G V

Q G K K V K S D A G L E V W P G P L S D N R V A V V L W N

R S S S K A T V T A S W S D I G L E K G K V V T A K D L W

E H T T K A S V S G Q I S A D I D S H A C K M Y V L T P N

Stop

The Senna α-galactosidase amino acid sequence along with the amino acid sequences of the α-galactosidases of Coffee and Guar (the only published plant α-galactosidases on the SWISS-PROT database) were used in a MegAlign clustal multi-sequence alignment. The results of the alignment are shown overleaf:-

```
                    10                  20
1   MEKM--MMWAKVVLCLFWVL  senna-germ-agal.PRO
1   MEKM--MMWAKVVLCLFWVL  senna-maturing-agal.PRO
1   MATHYSIIGGMIIVVLLMII  guar-a-gal.PRO
1   MVK-------------------  coffee-a-gal.PRO

                    30                  40
19  NASNCSGRLLN-----TIGN  senna-germ-agal.PRO
19  NASNCSGRLLN-----TIGN  senna-maturing-agal.PRO
21  GSE--GGRLLEKKNRTSAEA  guar-a-gal.PRO
4   ----------------SPGT  coffee-a-gal.PRO

                    50                  60
34  DHNNIHGRLLLGNGLGNTPP  senna-germ-agal.PRO
34  DHNNIHGRLLLGNGLGNTPP  senna-maturing-agal.PRO
39  EHYNVR-RYLAENGLGQTPP  guar-a-gal.PRO
8   EDYT-R-RSLLANGLGLTPP  coffee-a-gal.PRO

                    70                  80
54  MGWNSWNHFQCDINEEMVRE  senna-germ-agal.PRO
54  MGWNSWNHFQCDINEEMVRE  senna-maturing-agal.PRO
58  MGWNSWNHFGCDINENVVRE  guar-a-gal.PRO
26  MGWNSWNHFRCNLDEKLIRE  coffee-a-gal.PRO

                    90                  100
74  TADAMVSTGLASLGYEYVNL  senna-germ-agal.PRO
74  TADAMVSTGLASLGYEYVNL  senna-maturing-agal.PRO
78  TADAMVSTGLAALGYQYINL  guar-a-gal.PRO
46  TADAMVSKGLAALGYKYINL  coffee-a-gal.PRO

                    110                 120
94  DDCWAELNRDSKGNMVPSAS  senna-germ-agal.PRO
94  DDCWAELNRDSKGNMVPSAS  senna-maturing-agal.PRO
98  DDCWAELNRDSEGNMVPNAA  guar-a-gal.PRO
66  DDCWAELNRDSQGNLVPKGS  coffee-a-gal.PRO
```

```
                    130                 140
114 KFPSGIKALADYVHSKGLKF senna-germ-agal.PRO
114 KFPSGIKALADYVHSKGLKF senna-maturing-agal.PRO
118 AFPSGIKALADYVHSKGLKL guar-a-gal.PRO
86  TFPSGIKALADYVHSKGLKL coffee-a-gal.PRO

                    150                 160
134 GVYSDAGNQTCSKAMPGSLG senna-germ-agal.PRO
134 GVYSDAGNQTCSKAMPGSLG senna-maturing-agal.PRO
138 GVYSDAGNQTCSKRMPGSLG guar-a-gal.PRO
106 GIYSDAGTQTCSKTMPGSLG coffee-a-gal.PRO

                    170                 180
154 HEDQDAKTFASWGVDFLKYD senna-germ-agal.PRO
154 HEDQDAKTFASWGVDFLKYD senna-maturing-agal.PRO
158 HEEQDAKTFASWGVDYLKYD guar-a-gal.PRO
126 HEEQDAKTFASWGVDYLKYD coffee-a-gal.PRO

                    190                 200
174 NCNTNDISPRNRYPKMSEAL senna-germ-agal.PRO
174 NCNTNDISPRNRYPKMSEAL senna-maturing-agal.PRO
178 NCENLGLSVKERYPPMGKAL guar-a-gal.PRO
146 NCNNNNISPKERYPIMSKAL coffee-a-gal.PRO

                    210                 220
194 ANSGRAIFFSMCEWGSEDPA senna-germ-agal.PRO
194 ANSGRAIFFSMCEWGSEDPA senna-maturing-agal.PRO
198 LSSGRPIFFSMCEWGWEDPQ guar-a-gal.PRO
166 LNSGRSIFFSLCEWGEEDPA coffee-a-gal.PRO

                    230                 240
214 LWAKSVGNSWRTTGDIEDKW senna-germ-agal.PRO
214 LWAKSVGNSWRTTGDIEDKW senna-maturing-agal.PRO
218 IWAKSIGNSWRTTGDIEDNW guar-a-gal.PRO
186 TWAKEVGNSWRTTGDIDDSW coffee-a-gal.PRO
```

```
                    250                 260
234 E S M A S I A D Q N D K W A S Y A G P G senna-germ-agal.PRO
234 E S M A S I A D Q N D K W A S Y A G P G senna-maturing-agal.PRO
238 N S M T S I A D S N D K W A S Y A G P G guar-a-gal.PRO
206 S S M T S R A D M N D K W A S Y A G P G coffee-a-gal.PRO

                    270                 280
254 G W N D P D M L E V G N G G M T T E E Y senna-germ-agal.PRO
254 G W N D P D M L E V G N G G M T T E E Y senna-maturing-agal.PRO
258 G W N D P D M L E V G N G G M T T E E Y guar-a-gal.PRO
226 G W N D P D M L E V G N G G M T T T E Y coffee-a-gal.PRO

                    290                 300
274 R S H F S I W A L A K A P L L I G C D V senna-germ-agal.PRO
274 R S H F S I W A L A K A P L L I G C D V senna-maturing-agal.PRO
278 R S H F S I W A L A K A P L L V G C D I guar-a-gal.PRO
246 R S H F S I W A L A K A P L L I G C D I coffee-a-gal.PRO

                    310                 320
294 R S M D G A T Y G L L S N K E V I A V N senna-germ-agal.PRO
294 R S M D G A T Y G L L S N K E V I A V N senna-maturing-agal.PRO
298 R A M D D T T H E L I S N A E V I A V N guar-a-gal.PRO
266 R S M D G A T F Q L L S N A E V I A V N coffee-a-gal.PRO

                    330                 340
314 Q D S L G V Q G K K V K S D A G L E V W senna-germ-agal.PRO
314 Q D S L G V Q G K K V K S D A G L E V W senna-maturing-agal.PRO
318 Q D K L G V Q G K K V K S T N D L E V W guar-a-gal.PRO
286 Q D K L G V Q G N K V K T Y G D L E V W coffee-a-gal.PRO

                    350                 360
334 P G P L S D N R V A V V L W N R S S S K senna-germ-agal.PRO
334 P G P L S D N R V A V V L W N R S S S K senna-maturing-agal.PRO
338 A G P L S D N K V A V I L W N R S S S R guar-a-gal.PRO
306 A G P L S G K R V A V A L W N R G S S T coffee-a-gal.PRO
```

```
                                370                               380
354 A T V T A S W S D I G L E K G K V V T A  senna-germ-agal.PRO
354 A T V T A S W S D I G L E K G K V V T A  senna-maturing-agal.PRO
358 A T V T A S W S D I G L Q Q G T T V D A  guar-a-gal.PRO
326 A T I T A Y W S D V G L P S T A V V N A  coffee-a-gal.PRO

                                390                               400
374 K D L W E H T T K A S V S G Q I S A D I  senna-germ-agal.PRO
374 K D L W E H T T K A S V S G Q I S A D I  senna-maturing-agal.PRO
378 R D L W E H S T Q S L V S G E I S A E I  guar-a-gal.PRO
346 R D L W A H S T E K S V K G Q I S A A V  coffee-a-gal.PRO

                                410
394 D S H A C K M Y V L T P N               senna-germ-agal.PRO
394 D S H A C K M Y V L T P N               senna-maturing-agal.PRO
398 D S H A C K M Y V L T P R S             guar-a-gal.PRO
366 D A H D S K M Y V L T P Q               coffee-a-gal.PRO
```

Decoration 'Decoration #1': Shade (with black at 25% fill)
residues that match senna-germ-agal.PRO exactly.

From the plant α-galactosidase multiple sequence alignment it was possible to draw up a table of percentage similarity, this is shown in Table 21.

Table 21. The percentage similarity of the senna, coffee and guar α-galactosidase amino acid sequences.

Senna	*Coffee*	*Guar*	
	74.1	*75.7*	*Senna*
		75.9	*Coffee*
			Guar

The multiple alignment was also used to produce a phylogenetic tree of the Senna, Coffee and Guar α-galactosidases. The tree is shown in figure 67.

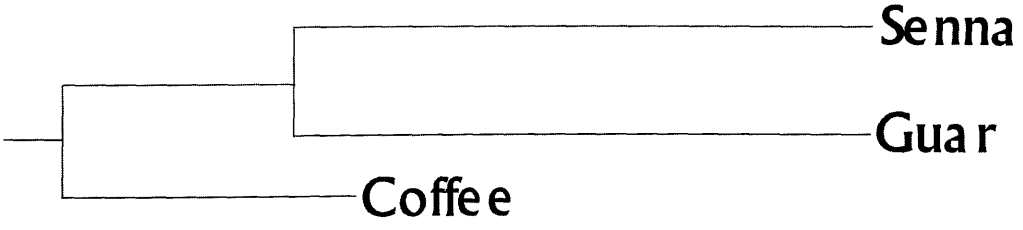


Figure 67. A phylogenetic tree showing the distances between the amino acid sequences of maturing Senna, germinating Senna, Guar and Coffee α-galactosidases.

Exhaustive database searching using the ExPASy bot revealed a further 11 published non-plant α-galactosidase sequences. These sequences were used with the plant sequences to produce a definitive α-galactosidase multiple sequence alignment. The percentage similarities of the α-galactosidases areshown in table 22.

Table 22. The percentage similarity of all published α-galactosidase amino acid sequences.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	
	25.1	32.3	29.2	25.6	26.7	23.4	23.8	23.1	23.1	23.1	24.4	28.8	24.7	A
		33.9	27.5	37.8	39.9	21.3	21.7	22.4	22.4	22.4	22.0	28.1	24.4	B
			75.9	36.0	31.5	33.1	35.2	35.7	35.7	35.7	33.9	75.7	35.2	C
				28.7	28.5	29.2	29.2	29.7	29.7	29.7	31.6	74.1	32.8	D
					78.0	23.1	23.8	22.6	22.6	22.6	22.8	30.3	26.1	E
						22.4	23.2	22.7	22.7	22.7	23.2	30.0	23.9	F
							97.5	96.0	96.0	96.0	80.3	30.8	68.0	G
								96.8	96.8	96.8	80.9	30.3	67.8	H
									100.0	100.0	81.5	31.5	67.0	I
										100.0	81.5	31.5	67.0	J
											81.5	31.5	67.0	K
												32.3	68.7	L
													31.0	M
														N

Key
A-Aspergillus niger (den Herder et al 1992)
B-Caenorhabditis elegans (Wilson et al 1994)
C-Coffea arabica (Zhu and Goldstein 1994)
D-Cyamopsis tetragonoloba (Overbeeke et al 1989)
E-Homo sapiens (Kornreich, Desnick and Bishop 1989)
F-Mus musculus (Ohshima et al 1995)
G-Saccharomyces cerevisiae MEL1 (Liljestroem 1985)
H-Saccharomyces cerevisiae MEL2 (Turakainen, Kristo and Korhola 1994)
I-Saccharomyces cerevisiae MEL5 (Turakainen, Kristo and Korhola 1994)
J-Saccharomyces cerevisiae MEL6 (Turakainen, Kristo and Korhola 1994)
K-Saccharomyces paradoxus (Naumova et al 1996)
L-Saccharomyces spp (Naumova et al 1996)
M-Senna occidentalis
N-Zygosaccharomyces cidri (Turakainen et al 1994)

The full multiple alignment was also used to produce a phylogenetic tree of all the α-galactosidase sequences. The tree is shown in figure 68.

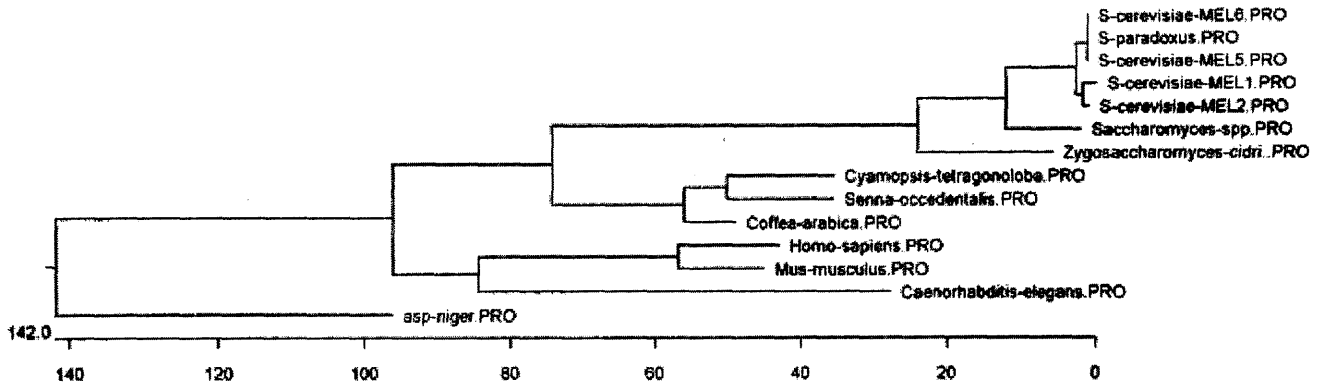
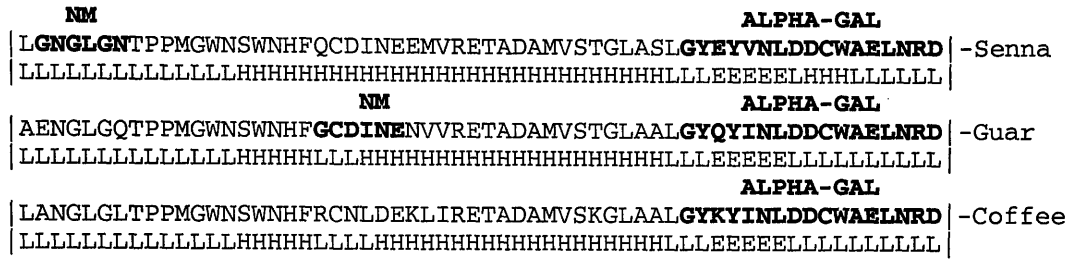
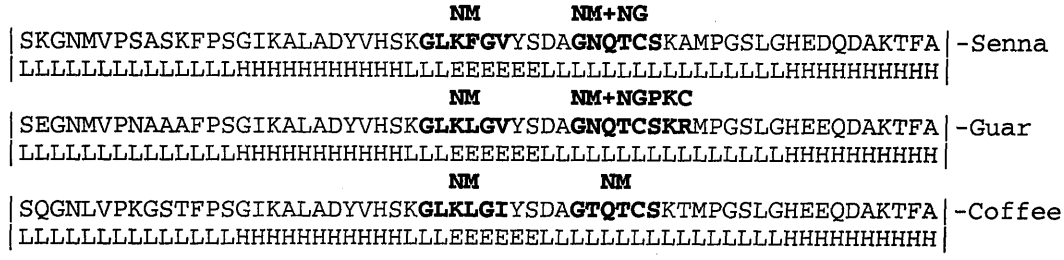
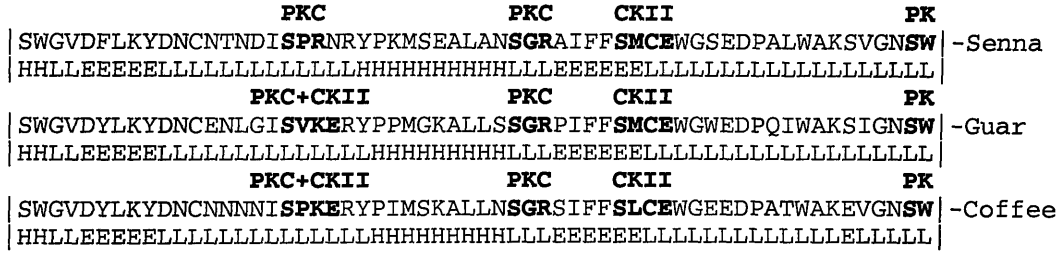


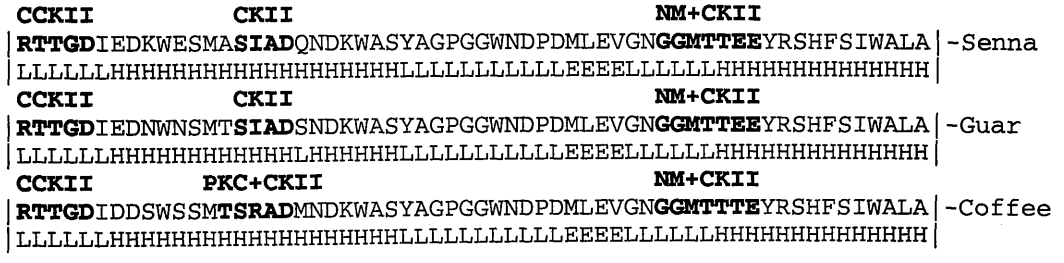
Figure 68. A phylogenetic tree with scaled braches showing the distances between the amino acid sequences of all the published α-galactosidases. The lower scale indicates the degree of divergence.

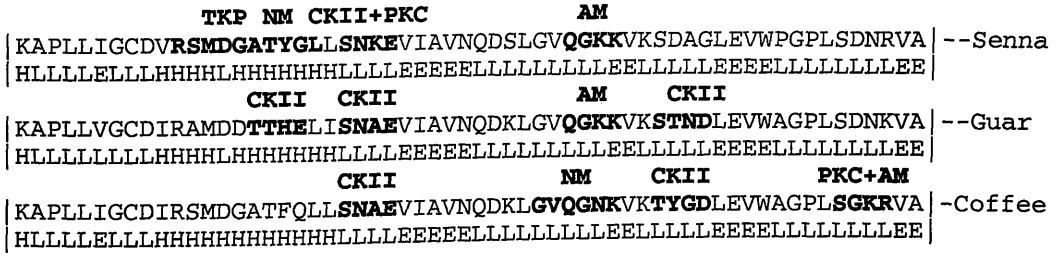
The primary amino acid structures of the three plant α-galactosidases were used with the Swiss-model secondary structure prediction utility (Guex, N. and Peitsch, 1997. Peitsch, 1996. Peitsch, 1995) to produce an alignment of secondary structures, shown in figure 69.


NM
ALPHA-GAL
LGNGLGNTPPMGWNSWNHFQCDINEEMVRETADAMVSTGLASLGYEYVNLDDCWAELNRD -Senna
NM
ALPHA-GAL
AENGLGQTPPMGWNSWNHFGCDINENVVRETADAMVSTGLAALGYQYINLDDCWAELNRD -Guar
ALPHA-GAL
LANGLGLTPPMGWNSWNHFRCNLDEKLIRETADAMVSKGLAALGYKYINLDDCWAELNRD -Coffee


NM
NM+NG
SKGNMVPSASKFPSGIKALADYVHSKGLKFGVYSDAGNQTCSKAMPGSLGHEDQDAKTFA -Senna
NM
NM+NGPKC
SEGNMVPNAAAFPSGIKALADYVHSKGLKLGVYSDAGNQTCSKRMPGSLGHEEQDAKTFA -Guar
NM
NM
SQGNLVPKGSTFPSGIKALADYVHSKGLKLGIYSDAGTQTCSKTMPGSLGHEEQDAKTFA -Coffee


PKC
PKC
CKII
PK
SWGVDFLKYDNCNTNDISPRNRYPKMSEALANSGRAIFFSMCEWGSEDPALWAKSVGNSW -Senna
PKC+CKII
PKC
CKII
PK
SWGVDYLKYDNCENLGISVKERYPPMGKALLSSGRPIFFSMCEWGWEDPQIWAKSIGNSW -Guar
PKC+CKII
PKC
CKII
PK
SWGVDYLKYDNCNNNNISPKERYPIMSKALLNSGRSIFFSLCEWGEEDPATWAKEVGNSW -Coffee


CCKII
CKII
NM+CKII
RTTGDIEDKWESMASIADQNDKWASYAGPGGWNDPDMLEVGNGGMTTEEYRSHFSIWALA -Senna
CCKII
CKII
NM+CKII
RTTGDIEDNWNSMTSIADSNDKWASYAGPGGWNDPDMLEVGNGGMTTEEYRSHFSIWALA -Guar
CCKII
PKC+CKII
NM+CKII
RTTGDIDDSWSSMTSRADMNDKWASYAGPGGWNDPDMLEVGNGGMTTTEYRSHFSIWALA -Coffee


TKP NM CKII+PKC
AM
KAPLLIGCDVRSMDGATYGLLSNKEVIAVNQDSLGVQGKKVKSDAGLEVWPGPLSDNRVA --Senna
CKII CKII
AM
CKII
KAPLLVGCDIRAMDDTTHELISNAEVIAVNQDKLGVQGKKVKSTNDLEVWAGPLSDNKVA --Guar
CKII
NM
CKII
PKC+AM
KAPLLIGCDIRSMDGATFQLLSNAEVIAVNQDKLGVQGNKVKTYGDLEVWAGPLSGKRVA -Coffee

```
     NG+PKC      CKII    NM  PKC+CKII  PKC       NM
|VVLWNRSSSKATVTASWSDIGLEKGKVVTAKDLWEHTTKASVSGQISADIDSHACKMYVL|-Senna
|EEEEELLLLLLEEEEEEEELLLLLLLLEELELHHHLLLLLLLLLLEEEEELLLEEEEE|
     NG+PKC      CKII    NM CKII
|VILWNRSSSRATVTASWSDIGLQQGTTVDARDLWEHSTQSLVSGEISAEIDSHACKMYVL|-Guar
|EEEELLLLLLLLEEEEEEELLLLLLLELLHHHHHHHLLLLLLHHHHHEELLLLLLEEEEE|
         NM              NM               PKCPKC   NM
|VALWNRGSSTATITAYWSDVGLPSTAVVNARDLWAHSTEKSVKGQISAAVDAHDSKMYVL|-Coffee
|EEEELLLLLEEEEEEEELLLLLLLLLLLEEHHHHHHLLLLLLLEEEEEELLLLEEEEE|

|TPN|-Senna
|LLL|
 PKC
|TPRS|-Guar
|LLLL|

|TPQ|-Coffee
|LLL|
```

Figure 69. The alignment of the secondary structures of the Senna, Guar and Coffee a-galactosidases. Secondary structure codes are H = helix, E = extended (sheet), L = loop (other structure). In bold type are the substitution sites, PKC = Protein kinase C phosphorylation site, CKII = Casein kinase II phosphorylation site, TKP = Tyrosine kinase phosphorylation site, NM = N-myristoylation site, AM = Amidation site, ALPHA-GAL = Alpha-galactosidase signature.

Discussion

The initial aim of this project was to compare the α-galactosidase present in the maturing endosperm of Senna seeds and responsible for the late developmental change in the Man/Gal ratio of galactomannan with the α-galactosidase produced in the germinating Senna endosperm, responsible for the mobilisation of the galactomannan reserves.

The further aim of this project was to identify potential promoter and enhancer regions responsible for controlling the expression of the Senna α-galactosidase gene. If it was discovered that the maturing and germinating versions of the α-galactosidase were produced from different genes, then it would have been the maturing version's regulatory elements that would have been the target for investigation. The regions of DNA upstream from the coding region would be sequenced and used in a database probe in an attempt to identify consensus sequences indicating where transcription factors would interact with the gene. If no matches were obtained, these upstream sequences could have provided the basis for an extensive series of reporter gene experiments (Goto et al 1989, Small et al 1992).

The reason for the interest in the (maturing) α-galactosidase promoter regions is that a potential application exists to utilise this work. Genetic manipulation of Guar, resulting in the expression of α-galactosidase in maturing guar seeds could producing the structural equivalent of the more costly Locust Bean Gum galactomannan (or indeed Senna galactomannan) in the annual guar plant.

Due to the fact that the Senna plants had to be hand pollinated and also that maturing endosperm tissue had to be isolated by hand dissection, only very small amounts of tissue were available. Because of this, it was clearly impractical to attempt large-scale protein purifications in order to compare the two α-galactosidases. Therefore it was decided to carry out small-scale extractions of the two tissues, identify the appropriate bands on gels and to compare as far as possible, the properties of the enzymes. The ability to identify the α-galactosidase on gels was provided by the availability of an antiserum raised against the α-galactosidase from the endosperm of germinated guar (provided by Unilever).

The antibody anti-α-galactosidase used on the Westerns localised only a single α-galactosidase protein in the maturing and germinated endosperms (figures 15 and 17 show this). Using the mobilities of the proteins on SDS-PAGE compared with molecular weight markers, a comparison of the apparent relative molecular masses of the α-galactosidases was made (figure 18). It was discovered that although the two α-galactosidases were very similar in size they did differ slightly. The size of the maturing α-galactosidase was calculated to be 44,000 Daltons and the germinating α-galactosidase was calculated to be 43,000 Daltons (Table 4). These two sizes fall directly between the sizes of the other plant α-galactosidases published, these are 41.3kDa for Coffee (Zhu and Goldstein 1994) and 45.1kDa for Guar (Overbeek, et al 1989).

A further characteristic of the two proteins, which may be used to distinguish between them, is isoelectric point (pI). IEFs were run in order to identify the pI's of both the maturing and germinating α-galactosidases. The activities of the α-galactosidases were localised on the IEF gels by using the PNP-α-gal assay. The activities of the α-galactosidases were then plotted against the pH of the region of the IEF where the activity was measured (figures 19 and 20). The pI of the maturing α-galactosidase was found to be pH 5.1 and the pI of the germinating α-galactosidase was pH4.9.

Further investigations into the similarity of the two Senna α-galactosidases were carried out by performing peptide mapping. Two different peptidases that cut at specific points were used to produce peptide fragments, which were run on SDS gels. It was discovered that in both cases maturing and germinating α-galactosidases produced identical digestion patterns (figures 22 and 23).

To further analyse the similarities between the maturing and germinating forms of the α-galactosidases, peptide fragments were generated by endo-peptidase digestion and then N terminally sequenced (the sequencing facilities were provided by Unilever). Due to the low levels of the α-galactosidases present in the protein preparations, only the non-peptidase treated samples produced N-terminal sequences that allowed direct comparisons between the maturing and germinating α-galactosidases. These sequences showed that the maturing enzyme contained 2 extra amino acids on the extreme N terminus when compared to the germinating enzyme.

From these studies on the α-galactosidase proteins it was possible to summarise that:

1. The two α-galactosidases are of a very similar but slightly different size.
2. The two α-galactosidases have a very similar but slightly different pI.
3. The two α-galactosidases produce identical fragmentation patterns on digestion with endo-peptidases
4. The two α-galactosidases have identical N terminus amino acid sequences apart from the extreme N terminus where the maturing enzyme has two extra residues missing from the germinating enzyme.

It was therefore possible to conclude that the two α-galactosidases were different. However, it was impossible to conclude whether the two enzymes were different due to:-

A. Being produced from similar but slightly different genes expressed at the different developmental stages.
B. Differential post-translational modifications (such as differential signal sequence cleavage, phosphorylation, glycosylation, etc.) at different developmental stages taking place on the translated products of the same gene.
C. Artefactual exo-peptidase degradation due to the protein preparation method used.

To discover which of possibilities A, B or C above are true it was necessary to isolate the cDNA of the α-galactosidases from both the maturing and germinating endosperms.

The initial RNA isolation attempts were made using the maturing endosperm tissue. A variety of RNA preparation methods were used as starting points for a number of experiments, these included, Lopez-Gomez (1992), Logemann (1987), Qiagen and Dynabead methods. In all the isolation attempts, the galactomannan present within the maturing endosperm presented problems, causing the lysed tissue to become very viscous and preventing the isolation of useful amounts of RNA. Attempted RT-PCR carried out using this limited amount of RNA resulted in the production of smears of nucleic acids (from very large to very small) on agarose gels (figure 31 demonstrates this).

Due to the problems encountered with the galactomannan present in the endosperms it was decided to temporarily abandon attempts to isolate RNA from the very galactomannan rich maturing endosperm. Instead, effort was concentrated on preparing RNA from endosperms of seeds that had been germinating for three days (and had, consequently, far less galactomannan present).

The isolation of RNA from the germinating endosperms again proved impossible using variations of the Qiagen kit. However the use of a modified version of the Lopez-Gomez method did produce a low but useable amount of RNA (shown in figure35).

The germinating RNA was subjected to a DNA removal procedure (in order to prevent, in subsequent polymerase chain reactions, the amplification of a possible maturing α-galactosidase within the contaminating genomic DNA). First strand cDNA was produced as outlined in the methods section (Frohman et al 1988). After the first strand was produced it was subjected to an alkali lysis technique to remove contaminating RNA fragments that were postulated to be acting as primers in the previous PCRs and thus causing the nucleic acid smears seen before. 3' RACE-PCR was carried out on the fully cleaned first strand germinating cDNA. The 3' PCR produced products of the sizes expected from the known size of the α-galactosidase protein and the estimated locations of the primer sites along the cDNA (figure 36).

The PCR products were cloned using the Novogen T/A cloning system. The results of the cloning were disappointing, only one clone seemed to contain the inserted PCR product when checked using restriction digests (figures 38). This clone was used as a template in further PCRs (re-PCRed) to confirm that it did contain the correct insert, which it apparently did (figures 39 and 41). On sequencing this clone was shown not to contain any insert. The confirmation PCRs that gave a PCR band of the size that suggested that the correct PCR product had been cloned must have done so because of "mis-priming" events. Further attempts to clone the PCR products using the Novogen materials proved unsuccessful.

Blunt ended cloning was attempted using a method based on the Stratagene pCR-script SK(+) cloning method. This procedure produced no clones. The reason for this was found to be the result of the blunt ended inserts not ligating to the blunt cut pBluescript vector.

T/A cloning was again attempted, using the Invitrogen kit. This method produced clones that apparently contained inserts of the correct size (figure 43). Confirmation that the Invitrogen kit had cloned the PCR products of interest was provided again by using plasmid preparations as templates in confirmation PCRs.

The clones produced by Invitrogen cloning were sequenced manually. The resulting data was compared to the published plant α-galactosidase sequences available via the ExPASy database bot (coffee and guar). It was discovered that the obtained clones were between 66 and 87% identical to the coffee and guar nucleic sequences (Table 13). It was therefore concluded that the clones obtained from the PCRs were germinating senna α-galactosidase.

In order to discover whether the germinating α-galactosidase gene was the only α-galactosidase present in the senna genome or whether a separate maturing α-galactosidase was present, southern blotting was carried out. Genomic senna DNA was restriction digested and run out on agarose gels as described in the methods (and shown in figure 47). The DNA-containing gels were then blotted to bind the DNA onto membrane supports.

To localise the germinating α-galactosidase on the blot, the 3' portion of the cloned senna germinating α–galatosidase was radio-labelled. This section of DNA was then used to probe the blot. Due to the fact that the 3' portion of the cloned germinating α-galactosidase contained the 3' untranslated region, it was therefore (and observably) less conserved across the plant α-galactosidases than the rest of the germinating α-galactosidase clone. Because this 3' probe was conserved to only a low degree it should only hybridise with only the germinating α-galactosidase on the southern blot. Figure 48 shows that the 3' probe did only pick out one gene; presumably the germinating α-galactosidase.

To probe the blot for other α-galactosidase-like genes, the 5' portion of the cloned germinating senna α-galactosidase was radio-labelled. This probe, highly conserved across the plant α-galactosidases was used (at the same stringency level as the 3' probe) to label any α-galactosidase–like genes present on the senna genomic southern blot. Figure 49 shows that the 5' probe indicated a number of bands had germinating α–galactosidase similarities.

The results of the southern blots were therefore inconclusive. It was indicated that there was only one germinating α-galactosidase gene within the senna genome. However, there was at least four α-galactosidase-like genes, any or none of which could have been a separate maturing α-galactosidase gene.

To clarify whether the maturing α-galactosidase was encoded on the same or different gene as the germinating enzymes, it was necessary to return to maturing endosperm RNA preparations. In order to isolate the endospermic RNA, preparations were carried out using maturing endosperms still attached to the testas. The rational behind this approach was that testa RNA would co-precipitate the endosperm RNA, giving a larger yield of both. The yield of RNA did increase and after testing testa only RNA for the presence of α-galactosidase encoding mRNAs (none were found) RT-PCR was carried out. The PCR amplified cDNAs of the expected size and these were (Invitrogen) cloned.

The maturing α-galactosidase clones were sequenced and aligned against the germinating α-galactosidase sequence. Both sequences proved to be identical. At this stage of the investigation it seemed that both the maturing and germinating α-galactosidases were produced from the same gene. However because of the locations of the primer sites it had not been possible to obtain DNA sequence that corresponded to the area of the extra amino acids possessed by the maturing version of the enzyme and therefore conclusive proof had yet to be obtained.

To obtain sequence information upstream from the most 5' primer possessed it was necessary to use new techniques. The techniques using genomic DNA as a starting point were inverse-PCR and single-sided PCR (as described in the methods section). These techniques offered the possibility of obtaining 5' regions of the α-galactosidase(s) and also upstream promoter elements. Although in both I-PCR and SS-PCR, promising PCR products were produced using nested sequential PCR

techniques (in order to increase the stringency of the reactions), on sequencing, none of the clones contained the expected α-galactosidase sequence designed to be incorporated into a successful α-galactosidase amplification.

The 5' regions of the two α-galactosidase were eventually isolated using 5' RACE-PCR techniques. These two sequences were compared and it was discovered that both the maturing and germinating cDNAs were almost identical and on translation, encoded identical amino acid sequences, including the two amino acids missing from the germinating version of the α-galactosidase enzyme. Therefore the difference in the N termini of the two α-galctosidase enzymes must be the result of differential signal sequence cleavage or accidental exposure to exo-peptidases.

With the knowledge of the extreme 5' end of the cDNAs primers were designed and final, full length cDNAs were produced by the proof-reading DNA polymerase, Pfu. The maturing and germinating clones produced were sequenced in sections (by gene-walking) and the sections then assimilated into a continuous whole (with the aid of the Dnastar package SeqMan).

The final full proof-red maturing and germinating DNA sequences were translated into amino acid sequences and compared. The two protein sequences were 100% identical. The two nucleic acid sequences were also aligned. These also showed a 100% similarity; even the third "wobble" position of all the codons were identical.

From the alignment of the full proof-read maturing and germinating derived sequences it was possible to conclude that the two forms of the α-galactosidase enzymes present at different stages of the senna developmental cycle were derived from the same gene. The observable differences in the mature enzymes must therefore be due to either:-

1. Accidental exo-peptidase degradation during preparation (although this unlikely because the sizes of the two forms of the α-galactosidase always differ by the same degree).
2. Differential post-translational modifications of the α-galactosidase. Differential post-translational modifications of α-galactosidase have been observed although only in transgenic mice expressing human α-galactosidase constitutively (Ishii et al 1998). It was shown that the α-galactosidase expressed in the different mouse tissues were glycosylated differently.

The senna α- galactosidase was used in a multiple sequence alignment with all the published α-galactosidases on the databases. This multiple sequence alignment permitted a sequence similarity table (table 22) and a phylogenetic tree (fig 68) to be produced.

The phylogenetic tree, with scaled branches, can be used to estimate the relationships (shown by the branches) and degree of divergence (shown by the branch lengths) between the α-galactosidases since they diverged from a common ancestor.

The tree comparing the α-galactosidases raises a particularly interesting point. The phylogenetic tree shows that the majority of α-galactosidase sequences (including the senna sequence) cluster in accordance with the genealogical ties between the organisms; the animal, plants and fungi show discrete groupings. However the sequence of the α-galactosidase from *aspergillus niger* does not fit into the phyolgeny pattern expected. Fig 68 shows that the *A.niger* α-galactosidase isolated by den Herder et al (1992) diverged from the other α-galactosidases at a very early evolutionary stage, previous to the divergence of the animals, plants and fungi from their common ancestor. This suggests that the α-galactosidases should actually be viewed as two sub-groups, Group I containing the α-galactosidases that branched off and led to the type isolated by den Herder in *A.niger* and Group II containing the other branch of the α-galactosidase family, which includes the senna α-galactosidase.

The three plant α-galactosidases display a high degree of similarity, it was found that the amino acid sequences have a similarity of between 74 and 76%. This high degree of similarity can be expected as the three plants from which α-galactosidases have been isolated all are members of the *leguminosae* family.

To further investigate the degree of similarity of the plant α-galactosidases secondary structure predictions were made. On comparing the predicted secondary structures of the senna, guar and coffee α-galactosidases it is clearly evident that the positions of α-helix and β-sheets are highly conserved. This high level of conservation is also true of the positions of the substitution and active sites of the three plant α-galactosidases.

Whilst the final proof-read maturing and germinating α-galctosidase sequences were being assembled and compared, a patent was applied for by Danisco (Jørsboe, Brunstedt and Petersen 1997) giving the full sequence of a senna endosperm α-galactosidase cDNA. Although the details contained within the patent are sketchy, the fact that the RNA preparation method used was that described in Logemann (1987), a method shown in this thesis not to isolate RNA from maturing endosperms, makes it seems very likely that the RNA amplified by Danisco must have been isolated from germinating endosperms. This senna α-galactosidase was amplified by using primers designed to match conserved regions of the guar and coffee sequences. Although Danisco did not know this at the time, my alignment clearly shows that the conserved regions of the guar and coffee are also conserved in senna. When the amino acid sequence obtained by Danisco was aligned with the maturing/germinating senna sequence obtained by me it was shown that there was a slight difference, the Danisco α-galactosidase sequence has 2 amino acid substitutions when compared to my α-galactosidase sequence. This was also the case with the nucleic acid sequence alignments; the Danisco sequence has 6 base substitutions when compared to my sequence. The reason for these slight differences is probably due to Dansico's use of a different strain of senna, than that used in my study (the Danisco patent does not state of what type or where its senna was obtained from).

The reason for Danisco's interest in the α-galactosidase was, like ours, the potential commercial importance of a transgenic guar that would express an active α-galactosidase within the maturing endosperm and so produce a galactomannan with a low galactose content.

Although Danisco had cloned their α-galactosidase from germinating senna endosperms, they did not know that in doing so they had also cloned the α-galactosidase present in the maturing senna endosperm and responsible for producing the final Man/Gal ratio of galactomannan.

After cloning their α-galactosidase, Danisco produced transgenic guar plants (Jørsboe et al 1997). These plants do express α-galactosidase activity. However, this expression is controlled by a constitutive promoter, meaning that the α-galactosidase is inefficiently expressed throughout the guar plants at all stages of development. There is no mention within the patent of any attempts to isolate upstream promoter elements that would cause the transgenic α-galactosidase to be expressed only in the guar endosperms and only during the final stages of seed maturation. The Danisco α-galactosidase genetic modification of guar resulted in only a small shift in the harvested galactomannan's Man/Gal ratio. The original Man/Gal ratio of 1.6:1 was increased to a ratio of 1.9:1, an increase of very low commercial significance.

Although Danisco have now applied for a patent to use the α-galactosidase gene along with their guar transformation techniques to produce a low galactose guar galactomannan, this may not be the only or most effective method of altering the galactose content of the galactomannan.

Recently the techniques developed during this thesis to produce the full length proof-read sequences of the α-galactosidase (such as:- a, RNA preparation methods, b, 3' RACE-PCR using a single degenerate primer designed against N-terminal amino acid sequences, c, Cloning techniques and d, 5' RACE-PCR to obtain the 5' regions of the cDNA), have been used by the Stirling University Plant Biochemistry Group to clone a further cell wall biosynthetic enzyme; the membrane-bound fenugreek α-(1→6) galactosyltransferase (Edwards et al 1999). It had previously been shown *in vitro* that by altering the activity of the galactosyltransferase relative to the mannosyltransferase activity, it was possible to produce galactomannans of widely varying Man/Gal ratios (Edwards 1992). Perhaps this may be achievable *in vivo* in guar by using a transgenic α-(1→6) galactosyltransferase to alter the overall expression or characteristics of the α-galactosyltransferase naturally present in the maturing guar endosperm. This could be a much more effective method of engineering a valuable low galactose guar galactomannan than the method described by Danisco.

Conclusions.

The initial aim of this project was to compare the α-galactosidase present in the maturing endosperm of senna seeds and responsible for the late developmental change in the Man/Gal ratio of galactomannan with the α-galactosidase produced in the germinating senna endosperm, responsible for the mobilisation of the galactomannan reserves. The conclusions obtained from this comparison are outlined below.

Investigation into the physical characteristicst of the two α-galactosidases revealed that:-

A. The two α-galactosidases are of a very similar but slightly different size, the maturing enzyme is 44 kDa and the germinating enzyme is 43 kDa
B. The two α-galactosidases have a very similar but slightly different pI, , the maturing enzyme's pI is 5.1 and the germinating enzyme's pI is 4.9
C. The two α-galactosidases produce identical fragmentation patterns on digestion with endo-peptidases
D. The two α-galactosidases have identical N terminus amino acid sequences apart from the extreme N terminus where the maturing enzyme has two extra residues missing from the germinating enzyme.

It was therefore possible to conclude that the two α-galactosidases were different. However, it was impossible to conclude whether the two enzymes were different due to:-

A. Being produced from slightly different genes.
B. Differential post-translational taking place on the translated products of the same gene.
C. Artefactual exo-peptidase degradation due to the protein preparation method used.

To discover which of possibilities A, B or C above are true it was necessary to isolate the cDNA of the α-galactosidases from both the maturing and germinating endosperms.

The initial RNA isolation attempts using the maturing endosperm tissue proved unsuccessful because the galactomannan present within the maturing endosperm caused the lysed tissue to become very viscous and prevented the isolation of useful amounts of RNA.

Effort was therefore concentrated on preparing RNA from endosperms of seeds that had been germinating for three days and had, consequently, far less galactomannan present. Again the isolation of RNA proved difficult. However, a modified Lopez-Gomez did produce a small amount of usable RNA.

The germinating RNA was used in 3'RACR-PCR and after a number of attempts the resulting product was cloned. The clone was sequenced and found to be highly similar to the sequences of guar and coffee α-galactosidases.

Southern blotting was carried out to discover whether the germinating α-galactosidase was the only α-galactosidase present in the senna genome. It was possible to conclude from the results obtained that the senna genome contained a number of germinating α-galactosidase-like genes, any of which could have been a maturing α-galactosidase gene.

Attempts to isolated the maturing endosperm RNA using co-precipitation eventually proved successful. Again the RNA was used in 3' RACE-PCR to produce a clone that was sequenced. The maturing α-galactosidase clone proved to be identical to the germinating α-galactosidase clone.

To prove conclusively that the two forms of α-galactosidase cDNA were identical it was necessary to obtain the 5' region of the cDNAs (were the codons encoding the region where the maturing α-galactosidase contained extra amino acids) and to carry out full sequencing of proof-read clones. This was performed and the two sequences proved 100% identical.

It conclusion it was therefore possible to state that the two versions of mature α-galactosidase derive from the same gene.

References

Ahmed Z, F, and Whistler R, L,, *J. Amer. Chem. Soc* **73** 4187-4188. 1951.

Anderson, E. *Ind. Eng. Chem* **41** 2887-2890. 1949.

Aspinall, G. O., Hirst, E. L., Percival, E. G. V. and Williamson, I. R. *J. Chem. Soc.* 3184 1953.

Baker C. W. and Whistler R, L,, *Carbohydr. Res.* **45** 237-243, 1975)

Bairoch, A., Bucher P. and Hofmann, K *Nucl. Acids Res.* **25**, 217-221. 1997

Bewley, J. D. and Black, M. Physiology and Biochemistry of seeds in relation to germination Vol 1 245-283 1978 *Springer-Verlag, Berlin.*

Campbell 1978 Ph.D Thesis, University of Stirling, Scotland.

Campbell J. M,. Reid J. S. G. *Planta*, **155**, No.2, pp.105-111, 1982.

Chouldhury , and Bagchi P., *Can. J. Chem.* **39** 375-381 1961.

Cleaveland, D. W., Fisher, S. G., Kirsscher, M. W. and Laemmli, U. K. *J. Biol. Chem.* **252** 1102-1106, 1977.

Courtois J. E.and P. Le Dizet *Bull. Soc. Chem. Biol.* **52** 15, 1968.

Crawshaw, L. A. and Reid, J. S G. *Planta* **160** 449-454 1984.

Dawson, R. M. C., Elliot, D. C., Elliot, W. H., Jones, K. M. Data for biochemical research. Clarendon Press, Oxford. 1982.

Dea, I. C. M., Morris, E. R., Rees, D. A., Welsh, E. J., Barnes, H. A., and Price, J. *Carbohydr. Res.* **57** 249-272 1977.

Dea, I. C. M. and Morrison, A. *Adv. Carbohydr. Chem. Biochem.* **31** 241-312. 1975.

de Vries, H. *Landwirtsch. Jahrb.* **6** 485-487 1877.

den Herder I.F., Rosell A.M.M., van Zuilen C.M., Punt P.J. *Mol. Gen. Genet.* **233** 404-410 1992.

Edwards M, Bowman Y. J. L, Dea I. C. M., Reid J. S. G., *J. Biol. Chem.* **.263,**.4333-4337 1988.

Edwards M, Bulpin P. V., Dea I. C. M, Reid J. S. G, *Planta*, **178**, 41-51 1989.

Edwards, M. E., Dickson C. A., Chengappa S., Sidebottom C., Gidley M. and Reid J. S. G. *The Plant Journal.(In press* **19(6)** 1-7*).* 1999.

Edwards, M. E., Scott C., Gidley, M. J., Reid J. S. G. *Planta* **187** 67-74 1992.

Frohman, M. A., Dush, M. K., Martin, G. R. *Proc. Nat. Acad.Sci* **85** 8998-9002, 1988.

Goldberg, R. *C. R. Acad. Sci* **264** 2899-2902. 1969.

Goto, T., MacDonald, P. and Maniatis, T. *Cell* **57** 413 1989.

Guex, N. and Peitsch, M. C. *Electrophoresis* **18**:2714-2723. 1997.

Hegi, G.: Flora von Mittel-europa. 2nd edn. Vol 2. Part3. Munchen: J. F. Lehmann. 1935.

Higgins, D. G. and Sharp, P. M. *CABIOS.* **5** 151-153. 1989.

Hui, P. A. and Neukom, H. *Tappi* **47** 39-42. 1964.

Ishii, S., Kase, R., Sakuraba, H., Taya, C., Yonekawa, H., Okumiya, T., Matsuda, Y., Mannen, K., Takeshita, M. and Suzuki, Y. *Glycoconjugate Journal* **15** 591-594 1998.

Kornreich R., Desnick R.J., Bishop D.F. *Nucleic Acids Res* **17** 3301-3302 1989.

Jørsboe. M., Brunstedt and J., Petersen, S. PCT patent application PCT/WO97/2093 1997.

Kooiman , P, *Carbohyd. Res.* **25**, 1 1972.

Laemmli, U. K., *Nature* **227,** 680-685 1970.

Leschziner C, and Cerezo A, S, *Carbohyd. Res.* **11**, 113-118, 1969.

Leung, D. M. W., Bewley, J. D. and Reid J. S. G. *Planta* **153** 95-100. 1981.

Liljestroem P.L. *Nucleic Acids Res.* **13** 7257-7268 1985.

Logemann, J., Schell, J., Willmitzer, L., *Anal. Biochem.* **163** 16-20 1987.

López-Gómez, R. and Gómez-Lim, M.A. *Hortscience* **27 (5),** 440-442. 1992.

Mackie, W and Preston, R. D. Polysaccharides from macroalgae (ed Stewart, W. P. D.) *Blackwell, Oxford.* 1974.

Maniatis, T. Fritsch, E. F. and Sambrook, J. Molecular cloning *Cold Spring Harbour Laboratory Press* 1982.

McCleary B, V, *Carbohydr. Res.* **71**, 205-230, 1979.

McCleary B, V,. Matheson, N, K., *Carbohydr. Res.* **119**, pp.191-219 1983.

Meier,H and Reid J. S. G. *Planta* **133** 234-248. 1977.

Nadelmann, H. *Jahrb. Wiss. Bot.* **21** 1-83 1890.

Naumova E.S., Turakainen H., Naumov G.I., Korhola M. *Mol. Gen. Genet.* **253** 111-117 1996.

Ochman, H., Gerber, A. S. and Hartl, D. L. *Genetics* **120** 621-623 1988.

Ochman, H., Medhora, M. M. Garza, D. and Hartl, D. L. (1990) p. 219-227. PCR protocols Academic press.

Ohshima T., Murray G.J., Nagle J.W., Quirk J.M., Kraus M.H. *Gene* **166** 277-280 1995.

Overbeek, N., Fellinger, A. J., Toonen, M. Y., Wassenaar, D., Verrips, C. T. *Plant. Molec. Biol.* **13** 541-550 1989.

Peitsch, M. C. *Bio/Technology* **13**, 658-660. 1995.

Peitsch, M. C. *Biochem Soc Trans* **24** 274-279. 1996.

Rao, C. V. N., Choudhury, D., Bagchi, P. *Can. J. Chem.* **39** 1408-1418 1961.

Reid J. S. G. *Planta* **100** 131-142. 1971.

Reid J. S. G and Bewley, J. D. *Planta* **147** 145-150. 1979.

Reid J. S. G., Edwards, M., Dea I. C. M. *Food Hydrocolloids* **1** 381-385. 1987.

Reid J. S. G., Edwards, M., Gidley, M. J. and Clark, A. H. *Planta* **195** 489-495. 1995.

Reid, J. S. G. and Meier, H. *Phytochem.* **9** 513-520. 1970.

Reid, J. S. G. and Meier, H. *Planta* **106** 44-60. 1972.

Reid, J. S. G. and Meier, H *Planta* **112** 301-308. 1973.

Rocks J. K., *Food technol.*, **25**, 476-483 1973.

Rosenfeld, J., Capdevielle, J., Guillemot, J. C., Ferrara, P. *Anal. Biochem.* **203** 173-179. 1992.

Saiki, R. K., et al. *Science* **230** 1350-1354, 1985.

Salisbury, F. B. and Ross, C. W. Plant Physiology. *Wadsworth Publishing, California.* 1992.

Small, S., Blair. And Levine, M *EMBO J* **11** 4047 1992.

Smith F, *J. Amer. Chem. Soc.* **70** 3249-3253 1948.

Seiler, A. *Planta* **134** 209-221 1977.

Turakainen H., Hankaanpaa M., Korhola M., Aho S. *Yeast* **10** 733-745 1994.

Turakainen H., Kristo P., Korhola M. *Yeast* **10** 1559-1568 1994.

Uebelmann, G. *Z. Pflanzenphysiol.* **88** 235-253 1978.

Unrau A.M. and Choy Y. M., *Can. J. Chem* **48** 1123-1128 1970.

Whistler R, L, and Durso D, F, *J. Amer. Chem. Soc* **73** 4189-4190. 1951.

Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A., Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L., Jones M., Kershaw J., Kirsten J., Laister N., Latreille P., Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M., Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R., Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.; *Nature* **368** 32-38 1994.

Zhu A., Goldstein J. *Gene* **140**. 227-231 1994.

Appendix

Primers designed using sequenced peptide fragments and synthesised by Unilever.

Primer name	Primer sequence
P1	GGIAAYACICCICCNATG
P2	CCICCIATGGGIACNAAC
P3	CCICCIATGGGIACNAAT
GS1	GTTATTGCIGTIAACCAAGAT
GA1	ATCTTGGTTIACIGCAATAAC
GS2	GAGAAACAGCTGATGCAATGG
GA2	CCATTGCATCAGCTGTTTCTC
GS3	ACGTCAATTTAGATGATTGCT
GA3	AGCAATCATCTAAATTGACGT