Molecular genetic analysis of the maize *terminal ear1* gene and *in silico* analysis of related genes

---

A Thesis Presented in Partial Fulfilment of the requirements for the Degree of Doctor of Philosophy in Plant Biology

at

Massey University
Palmerston North, New Zealand.

Daniel Charlton Jeffares

2001

This thesis is dedicated to Ben, Charlie and Alex.

THESIS ABSTRACT

Mutants of the maize *terminal ear1* (*te1*) gene have shortened internodes, abnormal phyllotaxy, leaf pattern defects and partial feminisation of tassels. The *te1* gene encodes an RNA recognition motif (RRM) protein, and is expressed in the vegetative shoot apex in semicircular rings that laterally oppose the positions of leaf primordia (Veit 1998). This project aimed to further characterise the molecular biology and function of the *te1* gene.

Molecular genetic studies aimed to further characterise the genes structure and expression. Genomic clones were sequenced revealing the intron exon structure. 5' RACE was used to predict a 5' transcription start site. Competitive RT-PCR showed that *te1* transcripts were highest in vegetative shoot meristems and embryos, lower in ears, roots and tassels, and undetectable in leaves. Two *te1* mutant alleles were cloned and the junctions sequenced, a further five alleles were characterised incompletely.

The TE1 peptide belongs to a subclass of RRM proteins which includes the *Schizosaccharomyces pombe* protein MEI2. More than 30 putative plant *Mei2*-like genes were identified in Genbank, no examples have been found in metazoans. Seven *Mei2*-like genes were predicted from the completed *Arabidopsis* genome. Exon structure and amino acid sequence supported three groupings of *Mei2*-like genes. Structural predictions of *Mei2*-like proteins indicate that the third RRM contained some novel structural features not present in canonical RRM proteins.

Attempts to study the function of the TE1 protein *in vitro* were limited by the inability of both *E. coli* and *Pichia pastoris* expression systems to express the full length protein, probably due to codon bias. Antibodies produced to a C-terminal portion of the protein did not specifically detect the TE1 protein in plant extracts without incurring non-specific activity.

The *te1* cDNA was ectopically expressed in *Arabidopsis* from a copper-inducible promoter both with and without the SV40 nuclear localisation signal (NLS). Although both *te1 and NLS:te1* transgenes were detected in transformants no phenotypes consistently correlated with transgene expression.

ACKNOWLEDGEMENTS

A lot of people have helped me with this PhD, but I would like to thank David Penny and Paula Jameson encouraging me to stay on after Honours, encouraging me to start my PhD, and for supporting my applications for the doctoral scholarship.

Thanks for all the members of my family for your continued encouragement, and for your attitude to things academic. It really makes a difference.

Thanks to Bruce Veit, for teaching the essential techniques of molecular biology lab work, how to critique experiments, all about meristems, and many other things besides. I know we haven't always got on Bruce, but your attitude to science, and why we do it, is very cool. Thanks also to Paul Reynolds for your help at the start of this project.

Thanks also to the many other people who have taught me during my time at Massey, in particular David Penny, David Fountain, Paula Jameson, Bill Jones, Mike McManus, Barry Scott, Stan Moore. Its been fun. Also to the people who helped me negotiate the trials of experimental work; Carolyn Young, Lyn Watson, Carmel Gillman, some of the above. Cheers also to the *Riborg Club*, Anthony Poole, and David Penny for the weird and wonderful world of *Riborgis eigensis*.

And to all those fellow students and friends around the Department of Plant Biology/IMBS/DEB Lab-Boffin Lounge and associates - Ive had a *darn fine* time: Abby, Alicia, Ant, Austin, Bennett, Carmel, Carolyn & David, Emily, Huai Bi, Lyn, Matt (bane of small mammals), Murray, Nena, Sang Dong Yoo, Tony (Yay for the *DEB Zeppelin* sessions!), Trish McLenachan, Trish Murray, Vern.

Katy. Thanks for everything, for even editing what to you must be meaningless jargon. ☺ Thanks for all your love and affection, you helped me to reject the hegemony of imputed needs.

I acknowledge the assistance of the Massey Doctoral Scholarship.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

2-ME	2-mercaptoethanol (β-mercaptoethanol)
Amp	ampicillin (superscript, concentration in μg/ml)
BCIP	x-phosphate/5-bromo-chloro-indolyl-phosphate
Cefo	Cefotaxime (superscript, concentration in μg/ml)
DEPC	diethylpyrocarbonate
DIG	Digoxigenin
DMPC	dimethylpyrocarbonate
DMSO	dimethylsulphoxide
dNTPs	deoxyribonucleotide triphosphates
DTT	dithiothreitol
EDTA	disodium ethylenediaminetetra-acetate
EMS	ethylmethane sulfonate
GA_3	Giberellic acid 3
Gent	gentamycin (superscript, concentration in μg/ml)
IPTG	Isopropylthio-β-D-galactaside
Kan	kanamycin (superscript, concentration in μg/ml)
LB	Luria Bertanni medium
NaOAc	sodium acetate
NBT	4 Nitroblue tetrazolium chloride
PBS	phosphate buffered saline
PCR	Polymerase chain reaction
PCR	polymerase chain reaction
pNPP	p-Nitrophenyl phosphate disodium
polyA+	Polyadenylated RNA
rpm	revolutions per minute
RRM	RNA recognition motif
RT-PCR	Reverse transcription-polymerase chain reaction
SAM	shoot apical meristems
SDS	sodium dodecyl sulfate
SDS-PAGE	SDS polyacrylaminde gel electrophoresis
Spec	spectinomycin (superscript, concentration in μg/ml)
SSC	sodium chloride, sodium citrate 20 x SSC is 3 M NaCl, 0.3M sodium citrate pH 7.0
Tris	Tris(hydroxymethyl)aminomethane
xg	multiples of gravitational force

TABLE OF GENES

A consistent nomenclature is used throughout this thesis to describe genes and proteins. Genes names are always italicised, proteins always listed in block capitals with suffixes describing the amino acids included for partial peptides. e.g. the portion of the *te1* gene was expressed to produce the peptide TE1$^{1-286}$.

Gene symbol	Gene	Species
ACE1	metallothionein regulatory protein	*Saccharomyces cerevisiae*
AML1, AML2, AML3, AML4, AML5	Arabidopsis Mei2-like	*Arabidopsis thaliana*
ANT	*AINTEGUMENTA*	*Arabidopsis thaliana*
CDC2a	cell division cycle 2a	*Arabidopsis thaliana*
CLV1, CLV2, CLV3	*CLAVATA*	*Arabidopsis thaliana*
CUC1, CUC2	*CUPSHAPED COTYLEDONS1, 2*	*Arabidopsis thaliana*
FIL	*FILAMENTOUS FLOWER*	*Arabidopsis thaliana*
gn1	*Gnarly*	*Zea mays*
GR	glucocorticoid receptor	*mammalian*
kn1	*knotted1*	*Zea mays*
MBP	maltose binding protein	*E. coli*
Mei2	Meiosis deficient 2	*Schizosaccharomyces pombe*
MGO1, MGO2	*MGOUN*	*Arabidopsis thaliana*
nptII	neomycin phoshotransferase	synthetic
OSH1	*Oryza sativa* homeobox1	*Oryza sativa*
STM	*SHOOTMERISTEMLESS*	*Arabidopsis thaliana*
T710	phage T7 protein 10	phage T7
te1	*terminal ear1*	*Zea mays*
TEL1, TEL2	*terminal ear1-like*	*Arabidopsis thaliana*
UFO	*UNUSUAL FLORAL ORGANS*	*Arabidopsis thaliana*
WUS	*WUSCHEL*	*Arabidopsis thaliana*
YAB3	*YABBY3*	*Arabidopsis thaliana*

CHAPTER 1: INTRODUCTION

1.1 ABSTRACT

The *terminal ear1* mutant phenotype is characterised by abnormal positioning of leaf primordia in the vegetative shoot meristem of maize (Veit 1998), indicating that the *te1* genes activity is a regulator of the positioning of organ primordia in the shoot apex. Analysis of the function of this gene may therefore bring a greater understanding of the mechanisms that determine phyllotaxy. A general introduction to developmental biology is presented with specific reference to plants, followed by a discussion of some important aspects plant development that relate to this thesis. Shoot meristems are described with a specific focus on the elements that allow the meristems to balance self propagation with lateral differentiation. Research into the phenomena associated with the positioning and initiation of lateral primordia is reviewed to set the scene for the specific context of the *terminal ear1* genes activity. The current understanding of the *terminal ear1* genes function in the positioning of leaf primordia is described. The *te1* gene belongs to a new family of RNA recognition motif (RRM) proteins that are present in all eukaryotes and includes the *S. pombe* gene *Mei2*. The conserved structures of RRM domains described with a brief introduction to the wide variety of cellular functions they regulate.

1.2. Developmental biology

At present, the mechanisms that determine phyllotaxy are not understood. Two principles of developmental biology that relate directly to this problem, and warrant discussion before proceeding; the positional information model (Wolpert 1998), and the competence of a tissue to respond (McDaniel 1984).

An important aspect of developmental biology is that spatial patterns of gene expression determine form. Wolpert (1969) formulated a model that described cells as gaining 'knowledge' of their position via gradients of morphogens. Cells interpret these gradients as positional information, and develop accordingly. Positional information is particularly important to plants, because plants cells appear to alter identity with respect to position to a much greater degree than do animal cells. Many aspects of shoot meristems rely on precise position information. Phyllotaxy is a clear example, in that certain populations of cells alter developmental fate depending entirely on their position.

It is known that cells will not always respond to positional information or may respond to the same information in a different manner. Plant phytohormones, for example, have a large variety of effects on plant growth and development. The continued polar transport of auxin induces gradual differentiation of cortical cells into different vascular tissues (Sachs 1984). It appears that a constant signal can give rise to further refinement of cell identity. The difference is not in the signal, but in the interpretation of the signal. On the molecular level, this is likely to be due to the expression of different receptor molecules, or different branch points in signal transduction cascades. Competence to respond to positional information is particularly important to the initiation of lateral organs in the shoot meristems. While in the central zone cells may receive positional information but they are not competent to respond. It appears that as cells become displaced from the central zone, they loose some of their indeterminacy, and become able to differentiate.

1.3. Three important aspects of plant development

1.3a. The control of cell division

Plant cells do not migrate, and cell death is a rare occurrence that usually does not affect primary formative development, so the regulation of cell division would seem to be crucial for the development of plant form. There are three parameters of cell division that could affect development significantly; the frequency of division, the plane of division, and the plane and extent of cell enlargement (elongation, widening *etc.*).

Alterations in the overall frequency of cell division do not appear to alter the tissues or organisation of tissues, as evidenced by the overexpression of a dominant negative *Arabidopsis* CDC2a gene in tobacco (Hemerly 1998). The resulting plants had fewer cells than normal, but were otherwise well formed. However, loss of control over the frequency of cell division in one particular tissue has drastic consequences for development, particularly the cell division of the shoot meristem. Mutations in either *CLAVATA1* or *CLAVATA3* genes result in a meristem that proliferates beyond normal limits, while *shootmeristemless* (*stm*) mutants fail to proliferate, with significant affects upon meristem function (Long 1996, Clark 1995). *AINTEGUMENTA* appears to promote division specifically in the cells of primordia, and ectopic overexpression of the gene results in enlarged mature organs derived from the shoot and the root that are normal in morphology (Mizukami 2000), indicating that the relative control over cell division rate is important in mature tissues also.

The plane of cell division seems to be much less important. Mutations in the maize *tangled1* gene have irregular planes of division in the longitudinal orientation, but normal divisions in the transverse plane (Smith 1996). Mutant leaves grow slowly, but are a normal shape. The mutant phenotype of *tangled1* suggests that planes of division are regulated differently in transverse *vs.* longitudinal dimension. The analysis of mutants of *ANGUSTIFOLIA* and *ROTUNDIFOLIA* in *Arabidopsis* indicates that cell expansion may also be regulated independently in different axes (Tsuge 1996). Mosaic analysis of the dominant *Gnarely-1R* class I homeobox gene also suggests that there are distinct pathways for communication between the lateral and transverse dimension (Foster 1999).

Cell enlargement appears to be more important to establishing the size of plant organs. *ANGUSTIFOLIA* mutant leaves are narrower than normal, while *ROTUNDIFOLIA* mutants are shorter than normal leaves (Tsuge 1996). While

narrower/shorter cells are the cause, the cells do not seem to compensate, by dividing more frequently to produce greater numbers of cells for example.

While the relative rates of cell division between different tissues is very important for development, neither the plane of cell division, the overall frequency of division, nor the extent of cell expansion needs to be controlled for the development of pattern. As we shall see in the next section, plants can be considered a supracellular network of cytoplasm, so it may not matter exactly where the boundaries between cells are formed.

1.3b. The role of intercellular trafficking in plant development

It is known from studies of periclinal chimeras that some genes act non-cell autonomously, that they affect clonal layers other than those in which they are expressed (e.g. Dudley 1993, Hake 1997, Foster 1999). Until recently however, the mechanisms underlying this phenomenon were not well understood. One of the most exciting findings in recent times has been that macromolecules such as mRNAs and proteins are transported. This transport is not only limited to specific macromolecules, but also to specific domains within the plant. Few examples have been studied well to date, but two will serve to illustrate some of the principles that are beginning to emerge. The KNOTTED1 (KN1) homeobox protein was found to be present in both the tunica and corpus of maize shoot apical meristems (SAMs), while *knotted1* mRNA was limited to the corpus (Jackson 1994), implying that the protein was able to move between cells. It was found that the KN1 protein is able to traffic *kn1* transcripts between cells (Lucas 1995). RNA transport also occurs over long distances through the phloem. Experiments with pumpkin-cucumber grafts indicated that *CmNACP* mRNA was transported across the graft through the phloem, and then through plasmodesmata into the shoot apex (Ruiz-Medrano 1999). Transcripts with a variety of putative functions (defence, cell cycle regulation, floral development *etc.)* were found to be present in cucumber phloem, and some of these transcripts were restricted from entry into the apex.

Transport of macromolecules in the symplasm appears to be controlled. Ultrastructural studies of plasmodesmata indicate that they contain some regular complexes, and that both cytoplasm and endoplasmic recticulum traverse junctions (Ehlers 1999). The molecules that are able to travel through plasmodesmata are usually restricted to a 'size exclusion limit' of about 1 kD, but overexpression of viral movement proteins increases the size exclusion limit to ~ 20 kD (Wolf 1989). For

KN1 to traverse plasmodesmata, it must induce their expansion, and the degree to which this occurs could well be regulated by specific domains of the protein. Experiments with injecting tracer dyes either directly into the shoot apex of birch (Rinne 1998), or loading dye into the *Arabidopsis* phloem (Gisel 1999) have shown that cells are not uniformly connected by functional plasmodesmata, specific domains are formed. These domains have been most extensively studied in the shoot apex, and they provide a new understanding of the structure of the SAM (discussed in Section 1.4a). Two aspects of plasmodesmata function may control the formation of symplastic domains. Primary and secondary plasmodesmata may have differences in transport capabilities, and primary plasmodesmata are only formed from clonally derived cells (van der Schoot 1999). Secondary plasmodesmata arise post cytokinesis, and the frequency of appearance appears to be developmentally regulated according to tissue fusion and cell elongation events (Lucas 1993). Also, plasmodesmata may be selectively closed by the activity of glucan synthase complexes that form callose plugs in the plasmodesmata lumen (Ehlers 1996, Lucas 1993).

The picture that emerges from these investigations is that a supracellular network allows movement of a variety of small molecules, and certain macromolecules. The discussion of the structure of the shoot apex will show that the connections of plasmodesmata are both spatially and temporally controlled.

1.3c. Structures that give rise to the plant body

Conceptually, plant development can be divided into three main stages; embryogenesis, the development of primary plant tissues from shoot and root meristems, and the development of secondary tissues from the vascular cambium and the cork cambium. Shoot and root meristems are usually morphologically distinguishable in the heart-shaped embryo (Steeves 1989), but probably begin to be defined at the molecular level with the onset of *WUSCHEL* expression at the 16 cell stage. This process is relatively complex, and is the subject of considerable study (see Lenhard 1999 and Bowman 2000 for reviews). This introduction focuses on vegetative development, and the formation of the SAM will not be described in detail. Two important points will suffice 1) the lineage of the meristem is established by gradual stages that begin very early, and 2) by the completion of embryogenesis a fully functional SAM is present. The remainder of the discussion of plant development will concentrate on a description of the structure and function of vegetative shoot

meristems (Section 1.4), with a particular focus on the initiation of lateral organs (Section 1.5).

1.4. Shoot meristems

1.4a. Structure of shoot meristems

Although shoot meristems differ in external dimensions and in various aspects of cellular morphology between taxa they share a common structural organisation (Steeves 1989). In fact shoot meristems are more similar between different taxa than are embryos, and are a better example of ontogeny recapitulating phylogeny (Sachs 1981).

Two features of vegetative meristems are visible at the level of cell histology: an exterior-interior organisation (tunica-corpus) organisation, and a radial zonation (central, peripheral and rib meristem zones). The functional nature of both levels of organisation have support from experimental data (see below). The tunica is composed of one or two layers of cells (depending on the taxa) that appear in regular files. Observations of periclinal chimeras indicates that each layer of the tunica (designated L1, and L2) is clonally derived, as is the corpus (L3) (Steeves 1989). Cell divisions in the tunica layers are therefore restricted to the anticlinal plane in the meristem, until the relaxation of division plane control in incipient primordia (Lyndon 1983).

Another cytologically visible organisation exists in the radial dimension between the cells in the center of the meristem, and those at the periphery. Cells of the central zone are larger, more vacuolate, and divide less frequently than the peripheral zone. The model of activity that is derived from the observations of histology and cell division rates is that the cells of the central zone act as the progenitors of the peripheral zone cells. Particularly large cells can sometimes be observed the apex of the central zone that are presumed to be apical initials, the ultimate source of the central zone. The expression patterns of genes such as *CLAVATA3*, *CLAVATA1*, *WUSCHEL*, *UNUSUAL FLORAL ORGANS*, and *SHOOTMERISTEMLESS* have delimited central zone cells of *Arabidopsis* more clearly, and indicated that the cytological zones represent functional differences (reviewed in Bowman 2000).

Investigations of the flow of low molecular weight dyes through plasmodesmata have added considerably to our understanding of the cellular structure of the SAM (Rinne 1998, van der Schoot 1999, Gisel 1999). Experiments with the injection of florescent dyes into the SAM of birch seedlings indicated that there were four symplastic zones in the shoot apex (Figure 1.1).

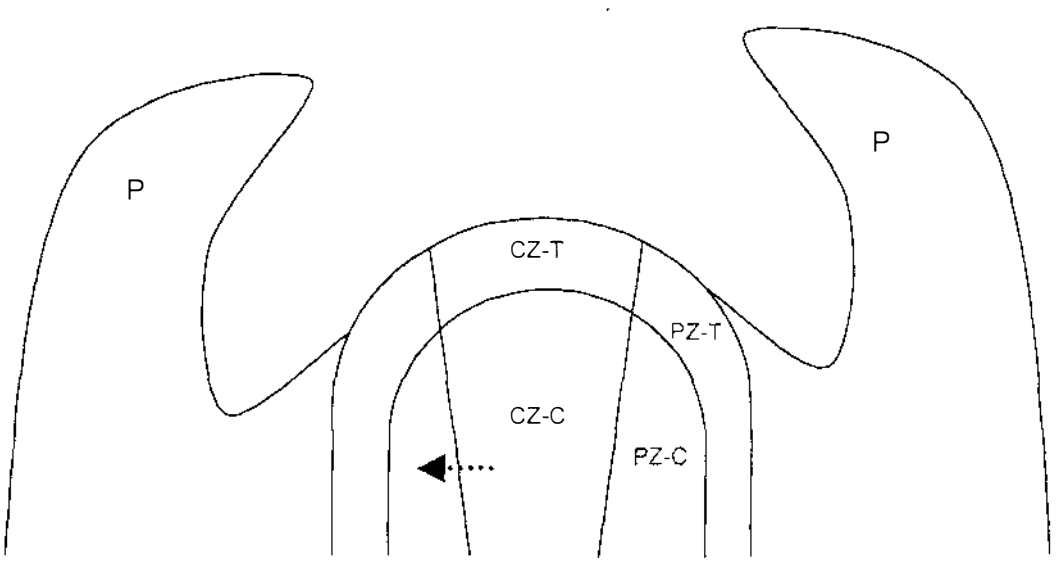


Figure 1.1. Symplastic fields in the SAM. This diagram shows an idealised meristem, with two primordia (P). Dye tracer experiments indicate that there are four fields of symplast connection within the SAM. The domains appear to correspond to central zone tunica (CZ-T), central zone corpus (CZ-C), peripheral zone tunica (PZ-T) and peripheral zone corpus (PZ-C). The peripheral zone domains are connected around the circumference of the central zone. In birch vegetative seedlings there is a transient connection between the central zone and the peripheral zone (arrow). Figure adapted from van der Schoot (1999).

While these zones were not directly related to the central and peripheral zones of birch, they appear very similar. The transient connection between the central and peripheral zones seen in these studies, which may coincide with the start of a plastochron, indicating that the limits of symplastic domains are dynamically regulated. The dynamic regulation of symplastic domains was also seen in investigations of *Arabidopsis* where the dye was loaded via the leaf phloem (Gisel 1999). In this study it was shown that the symplastic pathway from the phloem was limited to the tunica, and that the dye uptake decreased prior to the onset of flowering.

Because these experiments used low molecular weight dyes as tracers they can presumably flow though any functional plasmodesmata. These domains therefore delimit the trafficking of low molecular weight compounds (such as phytohormones) in the cytoplasm, endoplasmic recticulum (ER) lumen, or ER membrane. Intracellular macromolecules may have further restricted domains depending on the factors required to increase the size exclusion limit.

The overall structure of the shoot meristem is dynamic and changes during development. Most notably, the entire meristem becomes larger and looses some of its radial patterning with the onset of flowering, concomitant with a change in phyllotaxy (Steeves 1989). Despite the continuing displacement of central zone cells into the peripheral zone, the central zone is maintained throughout vegetative development. The dynamic maintenance of central and peripheral zones is a central theme of the next section.

1.4b. Functions of the shoot meristem

Shoot meristems continue organogenesis throughout the life of the plant, giving plant organs developmental plasticity in response to their environment. Although there are considerable complexities to meristem phenomena, their behaviour can be summarised with the following list of features:

1) *Indeterminate*; the cells can continue to divide 'indefinitely'
2) *Undifferentiated*; meristem cells are organised into other structures without prior de-differentiation
3) *Self propagating*; populations of central and peripheral zone cells are maintained and dynamically regulated
4) *Generative*; cells beyond the central zone are induced to differentiate

Each of these features will be discussed in turn, although they are interdependent.

1) Meristems are indeterminate

Clonal analysis has indicated that there may be small groups of more or less stable cells that are maintained at the summit of the meristem that give rise to the entire apical portion of the plant (Steeves 1989). While the individual cells that are in this position may not be the same throughout the life of the plant the presence of some cells in this state is maintained until the center of the apical dome differentiates at flowering. Such apical initials give rise to a large number of cells, and do not appear to have a limitation on the number of cell divisions. Differentiated cells, such as those of lateral organs only undergo a limited number of divisions. The indeterminate state is a particular feature of the central zone, and appears to be promoted by *kn1*-like homeobox genes. Misplaced expression of *kn1* in maize leaves results in 'knots' of newly dividing cells that to some extent are dedifferentiated (Jackson 1994). The overexpression of an *Arabidopsis kn1*-like gene *KNAT1* in *Arabidopsis* produced lobed leaves and ectopic meristems (Chuck 1996).

2) Meristem cells are not differentiated

There are two aspects to this feature of meristems: 1) central zone cells are able to differentiate into a variety of organs directly without dedifferentiation, 2) but until that time they are restricted from being recruited to lateral organ primordia. Although most plant tissues maintain 'pluripotentcy', the ability to regenerate entire plants, they usually must lose structure and developmentally regress into callus before doing so.

The cells of the central zone, and their immediate derivatives are able to differentiate directly. The *Arabidopsis WUSCHEL* (*WUS*) gene is required to prevent central zone cells from being recruited into organ primordia, the meristems of *wus* mutants are consumed in leaf primordia (Laux 1996). Interestingly, *WUS* is not required for continuing meristematic growth, since primordia initiation re-occurs in mutant apices. Since *Arabidopsis STM* and the maize ortholog *KN1* are not expressed in the site of incipient organ primordia their meristem-promoting functions appear to be incompatible with differentiation. The *Arabidopsis CUP-SHAPED COTYLEDON 1* and *CUP-SHAPED COTYLEDON 2* (*CUC1* and *CUC2*) and the Petunia ortholog *No Apical Meristem* (*NAM*) are possibly involved in the separation of organ primordia from the meristem (Aida 1999, Souer 1996). These genes are expressed between the shoot meristem and lateral organ primordia *cuc* or *nam* mutations result in fused cotyledons and floral organs.

3) Populations of central and peripheral zone cells dynamically self regulate their numbers

Surgical studies have indicated even very small portions of dissected SAMs will regenerate entire meristems (reviewed in Steeves 1989). An interesting factor of these studies is that the meristem reorganises itself into central and peripheral zones before new organs are formed. These surgical studies are a dramatic example of a natural process, since cells of the central zone are continually being 'removed' once they are recruited into organs. The CLAVATA1/CLAVATA3 complex appears to be required for the correct regulation of central zone size since mutations in either gene result in fasciated meristems (Clark 1997). Both genetic and biochemical data indicate that these genes act as a complex with CLAVATA3 as a extracellular signal, and CLAVATA1 as a the receptor (Clark 1995, Trotochaud 1999). Since *CLAVATA3* mRNA is present in the L1 central zone, and *CLAVATA1* in the L2 and L3, this complex is an example of co-ordination between symplastic boundaries. Mutations in *MGOUN1* or *MGOUN2* (*MGO1*, *MGO2*) genes also develop abnormally large meristems, and fasciation of the inflorescence stem (Laufs 1998). A comparison of the cells accumulated in the individual mutants showed that *mgo2* apices accumulated cells in the peripheral zone, whilst *clv3-1* apices accumulated cells in the central zone (Laufs 1998). As noted by Lenhard (1999), this suggests that cells must progress through two 'checkpoints' that alter fate before they are able to be recruited into lateral organs: the progression from initial cell to peripheral cell, which requires the *CLAVATA* complex, and then a progression peripheral cell into primordia, which requires the *MGOUN* genes.

Leaf primordia also affect the maintenance of the meristem. When temperature sensitive mutants of the *Antirhinum* MYB gene *phantastica* are grown in the restrictive temperature they develop abaxialised leaves (ventralised), and cease to maintain a meristem (Waites 1998). In *Arabidopsis phabulosa-1d* mutants leaves are adaxialised and the meristems are enlarged. These two observations imply that the proximity to adaxial (dorsal) primordia cells promotes meristems, while proximity to abaxial primordia cells limits meristems.

4) Cells immediately beyond the central zone differentiate

The recent derivatives of central cells proliferate once they exit the central zone, and are then recruited into organ primordia. The factors that influence the position of organ primordia (phyllotaxy) are the subject of the next section. It appears some aspect of the progression out of the central zone through the peripheral zone induces organ formation in the absence of any other stimuli. The nature of the lateral organ formed depends on the identity of the meristem. Apices that have been cultured *in vitro* illustrate that the sub-apical portion of the plant is not required for organ formation (Steeves 1989). Of particular interest here are the recent experiments of Reinhardt *et al.* (2000) with vegetative tomato meristems that were cultured in the presence of the auxin transport inhibitor *N*-1-naphthylphalamic acid (NPA). The resulting apices grew as long 'pins' with an apparently normal apical meristem at their summit,and no lateral organs. Apices could be cultured to the extent that they contained no previous pattern of leaf primordia or sub-apical tissue. Once these apices were transferred to a medium lacking NPA, leaf primordia arose in a random phyllotaxy, in the absence of any predetermined pattern, or sub-apical tissues. The random phyllotaxy always stabilised to the normal tomato spiral pattern, and this observation serves to illustrate an important facet of phyllotaxy: the influence of existing leaf primordia upon the positioning of new primordia.

1.5. Initiation of lateral organs

1.5a. Similarity of lateral organ initiation

The types of lateral organs produced by shoot meristems has been diversified during the course of evolution. Molecular studies indicate that A, B and C group floral identity genes are necessary for floral organ identity: in *abc* triple mutants all lateral organs develop into leaf like structures (Bowman 1991). Ferns do not produce flowers, do not appear to contain the specific MADS-box genes that are involved with

flowering (Münster 1997). It appears that the commitment of a lateral organ to its identity as a floral organ, rather than a leaf, occurs after the events of initiation, since floral organ primordia develop leaf-like organs when shifted back to vegetative inductive conditions (Batty 1984). Floral organs appear to have evolved by a modification of existing leaf initiation mechanisms, and would therefore be expected to share some fundamental mechanisms.

The next two sections will be concerned with the initiation of lateral organs, firstly a discussion of the factors that determine the positions of lateral organ initiation (Section 1.5b), and then a discussion of the phenomena associated with organ initiation *per se*, and some early events in the process of leaf differentiation (Section 1.5c). Although much of the research in this area has examined the initiation of vegetative leaves, homologous mechanisms are probably involved in the initiation of other lateral organs.

1.5b. Phyllotaxy: specification of organ position

The precise arrangements of phyllotaxy have inspired mathematicians to produce models to account for the various spatial and volume relationships within the meristem (e.g. Richards 1951, Jean 1989). An important finding of these studies is that volume relationships produce more consistent models than surface areas (Richards 1951). The correct positioning of lateral organs is intimately connected with meristem function, particularly the dimensions of the meristem. The maize *abphyl1* mutant, for example has wider meristem, and produces leaves in opposite pairs (decussate phyllotaxy) rather than the usual distichous phyllotaxy of maize (Jackson 1999). Similarly *sho1, sho2* and *sho3* mutants of rice, produce wider and flatter meristems that are highly variable in shape, and develop malformed leaves in a random phyllotaxy (Itoh 2000). *Arabidopsis clavata* mutants develop fasciated meristems and produce many more leaves and flowers (Clark 1995).

Surgical studies have indicated that the position of a new organ initial is influenced by its proximity to older primordia (Snow 1931), specifically, that existing primordia inhibit the initiation of a new primordia in their proximity. The lateral inhibition of leaf primordia has given rise to the field theory of primordia positioning (Snow 1931, Wardlaw 1949). Presumably, morphogenetic gradients mediate the radial positioning of leaf primordia. The nature of the repressing morphogen is unknown, although auxin is a candidate (see below). Mutants that show altered phyllotaxy as the result of enlarged meristems probably do so because the ratio of the diffusion or

degradation of the lateral inhibitor is not in accordance with the extra volume of the meristem. The transient connection between the central and peripheral symplastic domains observed in birch SAMs that occurs at the start of a plastochron is a clue to some radial communication in the SAM at this stage (Rinne 1998).

Mutants of the maize *te1* also mis-regulate the positions of leaf primordia, both in the longitudinal with regard to the divergence angle, although the dimensions of the meristem appear to be unchanged (Veit 1998). The *te1* expression pattern and mutant phenotype will be further discussed in Section 1.6.

At least two phytohormones have been implicated in the control of phyllotaxy. Applications of exogenous auxin is able to induce organ formation and result in changes in phyllotaxy that remain for several phytomers (Schwabe 1971, Meicemheimer 1981). The culture of tomato meristems in the presence of the auxin transport inhibitor NPA has shed more light on the role of auxin (Reinhardt 2000). Meristems that were cultured in the presence of the inhibitor grew as extended leafless 'pin-like' structures. The application of exogenous auxin to the summit of such pins no with no subtending leaves resulted in leaf initiation at a site below the apex in the radial dimension, but constant distance from the apex irrespective of the quantity of auxin applied. Since there were no subtending leaves the meristem itself must mediate the longitudinal positioning. It appears that auxin influences the radial positioning of primordia.

Gibberellins (GA_3) have also been implicated in some aspects of meristem organisation. The application of exogenous GA_3 to axillary meristems of the ivy *Hedera helix* results in a stable reversion of phyllotaxy from the adult spiral pattern to the juvenile distichous arrangement (Marc 1991). An examination of the apical meristem by Marc *et al.* (1991) indicated that the GA_3 treatment increased both the radial and the longitudinal dimensions of the apical dome. The distichous arrangement of leaves formed on such meristems appears to result from an increased longitudinal displacement of meristems from the summit of the apex. The involvement of Gibberellins is also implicated by the finding that over-expression of the rice *OSH1* gene in tobacco results in reduced levels of gibberelin (Kusaba 1998).

The field theory assumes that diffusible morphogens determine the pattern of phyllotaxy. An alternative proposal is that biophysical factors generate pattern (Green 1996). This hypothesis states that the regular positioning of primordia is the result of regular patterns of tension imposed due to corpus cell turgor upon the tunica. While models based on material science predict that this theory is plausible (Green 1996), it is unclear how tension is transducer into a primordia initiation stimulus. *Expansin* gene expression is upregulated in the sites of incipient leaf primordia (Reinhardt 1998), but it has not yet been shown to be induced by tension. Application of exogenous EXPANSIN protein to tomato apices induced the first stages of primordia initiation but the resulting apices did not continue to develop (Fleming 1997), implying that EXPANSIN is necessary but not sufficient for primordium initiation.

1.5c. Leaf initiation and acquisition of identity

Once a population of cells is induced to alter its developmental course to that of a leaf a variety of developmental events occur. Clonal analysis had indicated that of the order of 100 cells are involved at the time of recruitment, depending on the species (Steeves 1989). Surgical studies indicate that the fate of primordia cells is determined before any morphological changes are visible (Snow 1931, Wardlaw 1949). The first morphological sign of primordia development is a relaxation is the planes of cell division in the tunica cell layers which is associated with a reoreintation of cortical microtubules; the frequency of division increases only transiently (Lyndon 1983). The promordia cells then enlarge; *expansin* mRNA is expressed in incipient primordia (Reinhardt 1998). Observations of cell enlargement at incipient primordia sites of gamma-irradiated seedlings indicate that cell enlargement occurs irrespective of cell divisions (Foard 1971), so expansin expression may be a very early event. Other molecular events associated with the sites of incipient primordia are the downregulation of *KN1*-like homeobox gene expression in maize (Jackson 1994), and in *Arabidopsis* an increase in *AINTEGUMENTA*, *YABBY3* and *FILAMENTOUS FLOWER* expression (Elliot 1996; Seigfreid 1999).

Once primordia are initiated they begin to differentiate in the dorso-ventral, lateral, and proximal/distal axes. This differentiation is a gradual process, that seems to involve different processes in different dimensions (reviewed in Sinha 1999, Foster 2000). The process requires communication between the SAM and the primordia, since surgical incisions between young primordia and the SAM can result in leaves without dorso-ventral patterning (Sussex 1955, Snow 1959). The dorsoventral pattern

appears to be mediated in part by YABBY genes, and to be linked to meristem propagation (Bowman 2000). A number of maize mutants have shown that the determination of lateral fate occurs quite early in maize, perhaps concurrent with the wave of recruitment into leaf primordia state that extends around the apical dome from the midrib precursor cells (reviewed in Foster 2000). There are a number of pea mutants that display homeotic transformations of the portions of compound leaves, indicating that domain-specification processes occur in leaves (Marx 1987).

1.6. An introduction to the *terminal ear1* gene and homologous plant genes

1.6a. The *terminal ear1* gene of maize

The *terminal ear1* mutant of maize appears to be involved with the process of leaf initiation. The first maize plant carrying a *te1* mutant was originally identified in 1957 (Mathews 1974). From the outset the authors recognised it was a valuable research tool. Nine recessive mutant alleles have now been recovered from a variety of sources (Table 2.1), all of which appear to be loss-of-function alleles (see Chapter 2 Discussion). Mutant plants have both a short stature and more leaves (Mathews 1974, Veit 1998). Internode shortening varies considerably with genetic background, and the lengths of successive internodes within a particular plant appear to be random, giving the impression of a disordered developmental process (N Alvarez, unpublished data). Other features of the mutant phenotype that may or may not occur include disordered radial positioning of leaves so that phyllotaxy may become approximately spiral, pattern defects in leaves, and partial feminisation of the tassel. Tassel feminisation has been anecdotally associated with short stature, and is probably a secondary effect of the shortened stature since inflorescence sex is positively correlated with the length of the primary shoot in maize-teosinte F2 populations (Doebley 1991). The remainder of the mutant phenotype clearly indicates that the *te1* gene functions to control some aspect of the positioning of leaf primordia.

The *te1* gene was cloned and found to encode a putative RNA binding protein with three RNA recognition motifs (RRMs) (Veit 1998). The most similar gene at the time was the *Schizosaccharomyces pombe* gene *Mei2* which is required for both pre-meiotic DNA synthesis and meiosis I, and whose localisation to the nucleus during meiosis is dependant on a specific RNA species (reviewed in Ohno 1999, see also Chapter 3. Discussion). A large number of plant *Mei2*-like genes have been

discovered since that time. These genes are introduced in the next section, and described more fully in Chapter 3.

In situ hybridisation experiments have revealed that *te1* transcripts are present in both pre- and post embryonic vegetative shoot meristems in semicircular rings of cells that laterally oppose the positions of leaf primordia (Veit 1998, N Alvarez, unpublished data). This data, combined with the unconstrained positions of leaf primordia in *te1* mutants, has lead to a model wherein the *te1* gene acts to repress or delimit the positions of leaf primordia. The downregulation of *knotted1* (*kn1*) transcript that is associated with the first stages of leaf primordia occur higher up on the apical dome in *te1* mutants (Veit 1998), and mutant leaves are narrower and shorter than normal leaves, suggesting that they have formed from fewer founder cells (N Alvarez unpublished data).

1.6b. Plant homologues of the *terminal ear1* gene: *Mei2*-like genes

The maize *te1* gene belongs to a family of RRM genes referred to here as *Mei2*-like genes. Chapter 3 of this thesis compiles a significant amount of unpublished data obtained from sequences databases, and no formal description has yet been published. While *Mei2* was cloned from the fission yeast *S. pombe* the vast majority of *Mei2*-like genes are present in plants. No *Mei2*-like genes have been identified in any animal phyla, or in prokaryotes. There are over 30 putative plant *Mei2*-like genes, the completed *Arabidopsis* genome contains 7 such genes. *Mei2*-like proteins contain three RNA recognition motifs, the third of which is highly distinctive and appear to contain some unique structures not present in RRM proteins whose structures have been solved (Chapter 3.). Very little functional analysis of plant *Mei2*-like genes has been published to date. Hirayama *et al.* (1997) described the cloning of an *Arabidopsis* cDNA, *AML1* (*Arabidopsis Mei2*-like 1) by its ability to rescue a meiosis-deficient pheromone receptor mutant of *S. pombe*. The genes disrupted in this strain act upstream of *Mei2*, so it is likely that the AML1 protein was able to mimic MEI2 *in vivo*. The *Arabidopsis Mei2*-like genes are beginning to be characterised in the Veit laboratory, preliminary data indicates:

- *Arabidopsis Mei2*-like genes fall into two classes the *te1*-like group, and the *AML1*-like group.
- *Arabidopsis Mei2*-like genes are expressed in the shoot apical meristem and in developing embryos, although the specific domains differ.

- Since single mutants have no obvious phenotype, *Mei2*-like genes may act redundantly.
- *Arabidopsis Mei2*-like genes act in a co-ordinated fashion to control some aspects of meristem maintenance, lateral organ specification, and/or embryogenesis.

There is currently no biochemical data available on any of the plant *Mei2*-like genes. It appears that plant *Mei2*-like proteins bind RNA, since the RNA binding domains are the only universally conserved regions of the peptide sequences (Chapter 3). The biochemical activity of the *S. pombe* gene is becoming well understood; this data is reviewed briefly in Section 1.7c, and more extensively in Chapter 3.

1.7 RNA recognition motif proteins

1.7a Origins and structures of RNA recognition motif proteins

Amino acid sequence analysis has indicated that the putative *te1* translation product contains three RNA recognition motifs[1] (RRM) (see Chapter 3). RRM proteins bind to RNA with high sequence specificity, and also exibit non-specific RNA binding. A large number of proteins containing between one and four RRM domains have been identified in eukaryotes (Birney 1993); the Pfam database of protein families contains over 1800 RRM proteins. An analysis of the motif by Birney *et al.* (1993) provided a comprehensive survey of the sequence motifs and indicated that there are two sub-motifs, RNP-2 and RNP-1, that can be recognised within the sequence.

The structures of several RRM domains have been solved (Allain 2000, Conte 2000, Crowder 1999, Deo 1999, Inoue 1999, Kranz 1999, Nagata 1999, Wang 2001, Xu 1997). All RRM domains consist of a four stranded anti-parallel β sheet with two α helices that contact the β sheet on one face. A comparison of the structure of seven RRM domains indicated that the length and orientation of the β sheet and the α helix is very highly conserved, and that the only variable portions of the structure are the loops between these elements (Figure 1.
3) (this comparison was produced by S Moore at Massey University).

[1] RNA recognition motifs are also referred to as the RNA binding domain (RBD), and also the RNP consensus-sequence domain (RNP-CS).

These variable loops have been found to be a major component of the specificity of RRM domains (Shamoo 1995). Structural analysis of RRM proteins in complex with their RNA substrates indicates that the motif forms a variety of hydrogen bonds and stacking interactions with the RNA (Handa 1999, Oubridge 1994). Aromatic and charged residues within the RNP sub-motifs include the residues that form interactions with RNA.

Figure 1.2 RRM overlay
one page

Individual RRM domains are separated by flexible linkers so that the RRM folds are essentially independent of one another (Shamoo 1995, Crowder 1999). The extent of the protein required for specific RNA binding varies. Some proteins require more than one domain, or act as homodimers, others seem to bind RNA with just one domain. For example, the two RRM motifs of *human hn*RNP A1 protein show different binding specificity's *in vitro* than combined action of the two domains in the full length protein (Burd 1994). In contrast, the *S. pombe* MEI2 protein retains normal RNA binding specificity and function, with only the third RRM (Watanabe 1997). The binding specificity's of RRM domains are made more complex by induced fit mechanisms. Both the conformations of RNA, and the relative positions of RRM domains with respect to each other are altered with substrate binding (Shamoo 1995). For example, the relative conformation of the *Drosophila* SEX-LETHAL proteins two RRM domains change with the binding of the RNA target (Handa 1999).

1.7b Functions of RNA recognition motif proteins

RRM proteins are involved in a wide variety of cellular functions in plants involving the processing of RNA. This includes splicing, polyadenylation, transport, translation, as well as the regulated alternatives of these processes, such as translational regulation and alternative splicing (Alba 1998). A wide variety of RRM proteins are known to regulate development in eukaryotes (Bandziulis 1989). A complex and well studied example in animals is the *Drosophila* SEX-LETHAL protein. This protein regulates the sex specific alternative splicing of both its own mRNA and the *transformer* transcript influencing splice site selection. The protein also regulates the translation of the *transformer* mRNA, and the RRM domains have been shown to be involved in protein:protein interactions (McKeown 1992, Gebauer 1997). The *Arabidopsis* FCA protein contains two RRM domains and controls flowering time (MacKnight 1997). The *FCA* transcript is alternatively spliced, and may bind to its own 5' untranslated region (MacKnight 1997, R MacKnight pers. comm.). As yet the mechanism of alternative splicing is not known.

1.7c Biochemical function of the *S. pombe* MEI2 protein

Initial tBLASTn searches indicated that the protein most similar to TE1 was *S. pombe* MEI2. The MEI2 protein is required for pre-meiotic DNA synthesis and meiosis I (Watanabe 1994). During meiosis I, the protein moves from a uniform cytoplasmic distribution to a specific location in the nucleolus (Yamashita 1998). This localisation

is dependent upon the association of the MEI2 protein with specific non-translated RNA. The function of the MEI2 protein is regulated by phosphorylation, although phosphorylation does not affect its RNA binding activity.

The MEI2 protein has been studied intensively (reviewed in Ohno 1999). A detailed description of the biochemical and functional activity of the MEI2 protein is presented in the Discussion of Chapter 3.

1.8. Aims of this project

The current state of knowledge of the *te1* gene when this project began included the expression pattern, mutant phenotype, and the position of three of the mutations in *te1* mutant alleles (Veit 1998). The most similar gene to *te1* was the *S. pombe* gene *Mei2* - no potential plant homologs had been identified.

This project aimed to further characterise *te1* gene and its activity at several levels. Molecular genetic aspects of the gene were studied with the aim of understanding the genomic structure of the normal gene, and the nature of the nine mutant alleles.

A more comprehensive analysis of the expression of the gene sought to extend our understanding of the structures that the gene may influence (Chapter 2). The biochemical function of the TE1 protein had not been analysed, and this project aimed to study the function of the protein *in vitro* (Chapter 4). Since the function of the gene was only understood from loss-of-function mutant alleles in maize, a gain-of-function phenotype was investigated by expressing the gene in *Arabidopsis* (Chapter 5). Chapter discusses an *in silico* analysis of the Mei2-like gene family.

1.9. Data produced by other researchers

This thesis discusses. a substantial amount of data that by other members of the Veit group at Massey University. A large number of materials were obtained or produced by Bruce Veit including *te1* genomic and cDNA subclones, and *te1* mutant lines. Carmel Gilman assisted with sequencing of the *te1* genomic sequence (Section 2.10), and produced the pSBET:*te1-BamHI* expression vector that was most successful in producing the TE1 protein in *E. coli* (Section 4.8), and produced the TEL2-glucicorticoid receptor fusion vector and transformed it into *Arabidopsis* (Section 5.2). Nena Alvarez conducted a morphological analysis of the *te1-1* mutant, and produced *in situ* hybridisation data for the *Arabidopsis* Mei2-like genes *TEL1*,

TEL2, and *AML1*. Vernon Trainor cloned the *TEL1* and *TEL2* cDNAs (discussed in Chapter 3) and produced TEL2:GFP the fusion plants.

CHAPTER 2: MOLECULAR GENETIC ANALYSIS OF THE *TERMINAL EAR1* GENE OF MAIZE

2.1. ABSTRACT

The genomic sequence of the *te1* gene is presented showing intron/exon structure. Levels of the *te1* transcript were quantified in embryo, SAM, tassel, ear, leaf and root tissues of normal plants by competitive RT-PCR. An analysis of a collection of nine mutant *te1* alleles by Southern blotting and PCR is presented. 5' RACE was used to predict the position of the transcription start site of the gene, but these results were inconsistent.

METHODS (2)

2.2 General *E. coli* methods

2.2a. Growth of *E. coli*

Liquid *E. coli* cultures were grown in Luria-Bertani Medium (1% bacto-tryptone, 0.5% yeast extract, 1% NaCl pH7.0) (LB) at 37°C, shaking at 200rpm. Solid LB media in petri dishes contained 1.5% agar. Antibiotics used in media were filter sterilised and added after autoclaving the media. A common shorthand for antibiotic concentrations in all media is used in this thesis, the antibiotic is listed with a superscript indicating the concentration in μg/ml.

Glycerol stocks were prepared by adding 0.2 ml of sterile glycerol to 0.8 ml of stationary phase liquid culture, vortexing, and freezing in liquid nitrogen. Glycerol stocks were stored at -80°C.

2.2b. Preparation of heat shock-competent cells

All materials used with bacteria cultures were either baked (glassware), or treated with 0.1M HCl overnight then rinsed in sterile Milli-Q water to remove traces of vectors. A stationary phase 2 ml of LB culture was used to inoculate 250 ml of LB in a 1 L conical flask. This culture was grown at 37°C, shaking at 200 rpm until mid log phase. Cultures were then chilled on ice in flasks for an hour. Cells were harvested by centrifugation at 1 000 xg for 15 minutes at 4°C, and resuspended in 80 ml of chilled FB buffer (100 mM KCl, 50 mM $CaCl_2.2H_2O$, 10% (w/v) glycerol 10 mM potassium acetate, pH6.2). Cells were dispensed in 50-250 μl aliquots into sterile 1.8 ml microcentrifuge tubes, frozen in liquid nitrogen and stored at -80°C.

2.2c. Transformation of *E. coli* by heat shock

Aliquots of cells were thawed on ice and dispensed into sterile 1.8 ml microcentrifuge tubes. The transforming DNA was gently mixed with the cells, and they were incubated on ice for 10 minutes. Cells were heat shocked at 42°C for 30 to 60 seconds[1] then 0.5 ml of LB was immediately added. Cells were incubated for 1 hour at 37°C, and 250 μl was

[1] The optimal time for heat shock was determined empirically.

plated onto LB plates containing the appropriate antibiotic. The remainder was stored at 4°C and plated as required.

2.2d. Preparation of electrocompetent cells

250 ml cultures were grown to mid log phase and harvested as described for heat-shock competent cells (Section 2.2a.). Cells were washed twice by resuspension in 500 ml of ice cold sterile Milli-Q water and harvesting by centrifugation at 2 000 xg for 10 minutes at 4°C. Cells were then resuspended in 250 ml of ice cold sterile 10% glycerol and harvested as above. Finally cells were resuspended in 0.6 ml of ice cold sterile 10% glycerol, dispensed in 50 µl aliquots into sterile Eppendorf tubes (Eppendorf tubes), and frozen in liquid nitrogen. Competent cells were stored at -80°C.

2.2e. Transformation of *E. coli* by electroporation

Aliquots (50 µl) of cells were thawed on ice. A sample of the transforming DNA not exceeding 2 µl volume was mixed with the cells. Cell-DNA mixtures were transferred to ice cold 2 ml electroporation cuvettes with a 2mm electrode gap (EquiBio Ltd), and incubated on ice for 10 minutes. Cuvettes were pulsed in a Biorad GenePulser electroporation device with the parameters recommended by Biorad for *E. coli* (25 µF capacitance, 2.5 kV potential, 200 Ω resistance). The cells were immediately resuspended in 500 µl of LB and transferred to a sterile Eppendorf tube. Cells were incubated for 1 hour at 37°C, and 250 µl was plated onto LB plates containing the appropriate antibiotic. The remainder was stored at 4°C and plated as required.

2.2f. Alkaline lysis of E. coli

This method was conducted essentially as described in Sambrook (1989). A 50 ml LB culture grown overnight with the appropriate antibiotics was harvested by centrifugation at 6 000 xg for 15 minutes at 4°C. The pellet was resuspended in 1 ml of Solution I (50 mM glucose, 25 mM Tris-HCl pH8.0, 10 mM EDTA). 100 µl of Freshly prepared 10 mg/ml lysozyme in 10 mM Tris-HCl pH8.0 was added and mixed. 2 ml of freshly prepared Solution II (0.2 N NaOH, 1% SDS) was added and mixed by gentle inversion, and the tube was stored at room temperature for 5 minutes. 1.5 ml of Solution III (5 ml of 5 M potassium acetate, 11.5 ml of glacial acetic acid, 28.5 ml of Milli-Q water) was added and the tube was mixed by shaking. The tube was stored on ice for 10 minutes,

and then centrifuged at 2 000 xg for 30 minutes at 4°C. The supernatant was removed, pouring it through cheesecloth wet with isopropanol into a new tube, taking care to avoid including any white precipitate with the supernatant. 0.6 volumes of isopropanol was added to the supernatant, mixed well, and stored at room temperature for 10 minutes. Nucleic acids were recovered by centrifugation at 3 000 xg for 15 minutes at room temperature. The supernatant was removed, the pellet washed with 80% ethanol, and resuspended in 0.5 ml of TE10/1 (10 mM Tris-HCl, 1 mM EDTA, pH 8.0). DNA samples were extracted with Tris-HCl pH 7.5 saturated phenol:SEVAG (1:1), and then with chloroform. SEVAG contains isoamyl alcohol:chlorform (1:24). Extracted samples were precipitated with 1/10th volumes of 3 M NaOAc and 2.5 volumes of ethanol, rinsed with 80% ethanol, and resuspended in 50 µl TE10/1 containing 1 µl of 10mg/ml RNase A.

2.2g. Lysis of *E. coli* by boiling using STET buffer

This method was conducted essentially as described in Sambrook (1989). A colony from a bacterial plate was spread with a sterile toothpick onto an area of approximately 1/8th of a plate of LB containing the appropriate antibiotics, and allowed to grow at 37°C overnight. The bacterial cells were scraped from the surface of the plate with a sterile toothpick, and mixed with 350 µl of STET buffer(0.1M NaCl, 10mM Tris-HCl pH8, 1mM EDTA, 5% Triton X-100) in a sterile Eppendorf tube. 25 µl of freshly prepared 10mg/ml lysozyme in 10 mM Tris pH8.0 was added, and mixed. Tubes were centrifuged at 12 000 xg for 10 minutes at room temperature. The pellet of cellular debris was removed with a sterile toothpick. 190 µl of 7.5 M NH_4OAc and 570 µl of isopropanol was added, mixed and the tubes stored for 5 minutes at room temperature. The tube was centrifuged at maximum speed (20 000 xg) at 4°C for 5 minutes. The supernatant was removed, and the pellet washed by centrifugation for 1 minute in 80% ethanol. The pellet was air dried and resuspended in 20 µl of TE10/1 containing 1 µl of 10mg/ml RNase A.

2.3 General DNA methods

2.3a. Quantification of DNA and RNA

Nucleic acids were quantified either by spectrophotometry, or by agarose gel electrophoresis and staining with ethidium bromide. For spectrophotometry quantification, DNA/RNA was diluted and dispensed into quartz cuvettes. Absorbance

readings at 260 and 280 nm wavelength light were recorded. The quantity of DNA was calculated according the formula C = OD_{260} x 50 x D for double stranded DNA, and C = OD_{260} x 40 x D for RNA, where C is DNA/RNA in μg/ml, D is dilution factor. 260/280 ratios were used as an indication of purity, readings of 1.8 were indicated relatively pure DNA samples, readings of 1.9 to 2.0 were expected for RNA samples.

To quantify DNA/RNA by electrophoresis gels were stained with ethidium bromide and compared by eye to the known masses of particular fragments in molecular weight markers according to their molar ratios, either λ *Hind*III digest or 100 bp ladder (Molecular Weight Marker XIV, Roche) for DNA, and 0.24-9.5 kb RNA ladder (Gibco BRL) for RNA. For example, a 211 ng loading of λ *Hind*III contains 100 ng of the 23 kb fragment, and 25 ng of the 9.4 kb fragment.

2.3b. Electrophoresis of DNA

DNA was electrophoresed in 0.8 to 2.5% TAE (40 mM Tris-acetate, 1 mM EDTA (pH8.0)) gels with 250 ng of 100bp ladder (Roche) and/or 211 ng of a λ *Hind*III digest as molecular weight markers. Gels were stained in ethidium bromide, and visualised with either an IS1000 Digital Imaging System (Alpha Innotech Corp.), or an Ultralum Integrating CCD Camera (Ultralum Inc.).

2.3c. Standard PCR Reaction Protocol

Most PCR reactions described in this thesis used very similar conditions, varying only in primers used, certain variables of the PCR program (annealing temperature, number of cycles and extension time), and the template(s). Standard conditions and procedures are described here, special cases referred to where appropriate. Standard reactions were performed in 50 μl final volume containing a final concentration of 1x PCR buffer (10 mM Tris-HCl pH8.3, 50 mM KCl, 1.5 mM $MgCl_2$) (Roche), 1x PCR Enhancer Solution (Gibco BRL), 400 pM of each primer, 250 μM of each deoxynucleotide triphosphate (Roche) (250 μM dNTPs), and 2.5 units of Taq polymerase (Roche). 'Hot start' reactions were standard; Taq polymerase was excluded from the initial reaction mix of 40 μl volume, a 10 μl volume enzyme mix containing 1x PCR buffer and 2.5 units of Taq polymerase was added to each tube during the first 95°C denaturation step.

Most PCR programs are described by a common formula $DJT_axN_c(L)$, where T_a denotes the annealing temperature, N_c indicates the number of cycles, and (L) is an optional

parameter that indicates a long extension time (see Appendix 2 for a more detailed description). Either an MJ Research PTC-200 DNA Engine, an MJ Research PCT 150 Minicycycler were used in most cases.

2.3d. Colony PCR

For rapid screening of *E. coli* transformants a sterile toothpick was touched to the colony and streaked onto an LB plate with the appropriate antibiotic, and then the small quantity of bacteria adhering to the toothpick were mixed with a PCR reaction mix of 40 µl volume (containing all reagents except Taq polymerase, see Section 2.3c). A PCR reaction was then conducted according to standard PCR protocol.

2.3e. General scheme for vector construction

In general, subcloning avoided the use of PCR. Both the cloning vector, and the vector containing the 'insert' DNA, to be cloned were digested with the appropriate restriction enzyme. Where there was potential for the vector to self ligate, the vector was phosphatased using a heat labile alkaline phosphatase (TsAP, Gibco BRL), followed by alkaline phosphatase deactivation as recommended by the manufacturer. Vector and insert were electrophoresed in 0.8 to 1.0% agarose TAE gels, gels were stained with ethidium bromide, and the desired fragments were cut from the gel with a sterile scalpel blade. DNA was extracted from gel slices with Concert gel extract kits (Gibco BRL) as recommended by the manufacturer.

Ligation reactions were set up based on the principles described in Sambrook (1989), and were usually allowed to proceed overnight at room temperature. Ligations were transformed into DH5α cells by heat shock (Section 2.2e). Well separated colonies were re-streaked onto LB plates with the appropriate antibiotic, and plasmid DNA isolated using STET buffer (Section 2.2g). Constructs were verified by restriction digests that indicated both the presence of the insert, and where necessary, the orientation. Where protein fusions were desired, vector-cDNA junction was sequenced as described in Section 2.4c.

2.4 Sequencing of *te1* genomic clone

2.4a. General scheme of project

This work was carried out in collaboration with Bruce Veit and Carmel Gilman. Bruce Veit cloned the original genomic fragment, and produced the subclones apart from DJ22. Daniel Jeffares and Carmel Gilman produced most of the sequence equally, and compiled and corrected the reads.

A 12.4 kb *Kpn*I genomic clone containing the *te1* gene was obtained by screening a genomic library (Veit 1998). All sequence positions are numbered from the 5' *Kpn*I site unless stated otherwise. This construct had been analysed by restriction digestion and subcloned into pBluescriptSK$^+$ (Stratagene) (subclones are described in Appendix 5). These subclones were used as templates for automatic sequencing using ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction kits (Perkin Elmer), which were electrophoresed and analysed on an Applied Biosystems 373A DNA Sequencing System by the Massey University Sequence Analysis Facility.
At first M13 forward and reverse primers, and a variety of primers that had been produced to the *te1* cDNA were used as sequencing primers. Further primers, the TG primer series (see Appendix 1), were designed as required. Primer positions are listed in Appendix 1., and the positions of the assembled sequencing reads is shown in Appendix 4.

2.4b. Construction of 5' end genomic subclone

One further subclone was constructed from the vector pBV306 (which contains the most 5' *Bam*HI fragment of the original 12.4 kb genomic clone) to facilitate sequencing of the 5' region. The 639 bp *Xba*I fragment of pBV306 was cloned into pBluescriptSK$^+$ using standards protocols to create pDJ22. This *Xba*I fragment corresponds to the region from 441 bp to 1080 bp of the completed sequence.

2.4c. Sequencing reactions

Plasmids to be used for sequencing were extracted from *E. coli* DH5α with Concert Miniprep plasmid kits as recommended by the manufacturer, except that plasmid was eluted from the column in 30 µl of sterile Milli-Q™ (Millipore) water and stored at -20°C.

Sequencing reactions were prepared to 10 μl volume and contained 200-500 ng of plasmid, 1.6 μl of 1 μM oligonucleotide primer and 4 μl of either Big Dye Terminator Ready Reaction Mix (Perkin Elmer), or for GC-rich regions that proved difficult to sequence, 4 μl of the dGTP Ready Reaction Mix. Thermocycling of the dye-termination products used the AUTOSEQ program; 25 cycles of 96°C for 10 seconds, 50°C for 5 seconds, 60°C for 4 minutes, after which reactions were held at 4°C. Temperature ramping was limited to 1°C/second.

Products were purified by ethanol precipitation. The reaction volume was increased to 20 μl with Milli-Q water and transferred to 1.8 ml microcentrifuge tubes, 2 μl of 3 M sodium acetate (NaOAc) pH5.2 and then 50 μl of ethanol was added. Tubes were incubated on ice for 5 minutes and centrifuged at maximum speed (20 800 xg) for 15 minutes at 4°C in an Eppendorf 5417R microcentrifuge. The supernatant was removed and 700 μl of 80% ethanol was added, tubes were again centrifuged at maximum speed (20 000 xg) for 15 minutes at 4°C. The 80% ethanol was removed and the tube was allowed to dry. Product were stored dry at -20°C until electrophoresis by the Massey University Sequence Analysis Facility.

2.4d. Assembly of sequence files

All sequence files were edited and assembled into contigs with the Wisconsin Package Version 9.1 (Genetics Computer Group) using the *Fragment Assembly System* series of programs. Where possible, ambiguities in the sequence was resolved by comparing the electrophoretograms produced from multiple reactions of the region concerned, this included other sources of sequence such as the published *te1* cDNA sequence and electrophoretograms from 5'RACE sequences (described in Section 2.6). Intron-exon positions were established with the GCG program *Gap*. The *Gap* output was examined by eye for monocotyledon consensus splice site sequences (Simpson 1996).

2.5 General RNA methodology

2.5a. Precautions to avoid RNase contamination

Solutions used for RNA work were not used for any other purpose, and were always handled with gloves. All solutions, except those containing primary amines (such as

Tris), were treated with diethylpyrocarbonate (DEPC) or dimethylpyrocarbonate (DMPC) (Sigma) prior to use. DEPC/DMPC to 0.1% by volume was added to solutions in baked bottles, bottles were shaken for a minimum of two hours at 37°C, then autoclaved twice. The pH of critical solutions was checked, and readjusted after treatment with DMPC/DEPC. The electrode used with the pH meter was treated with 50 mN NaOH for 10 minutes, then rinsed with sterile Milli-Q water prior to use with RNase-free solutions. Solutions that could not be treated with DMPC/DEPC (such as Tris) or could not be autoclaved (such as SDS) were prepared from new containers set aside for RNA work using sterile Milli-Q water and baked glassware.

Milli-Q water that had been collected directly from the Mill-Q™ apparatus, and autoclaved in baked bottles was found to be free of RNases. All glassware used for RNA work was baked overnight at approximately 160°C, all plastics were either new, or treated overnight in 0.3% H_2O_2 (Andrew Industrial Ltd.).

2.5b. DNase treatment of RNA preparations.

Two buffers were used for DNase I treatment of RNA samples. A Mn DNase buffer (final concentration 1 mM $MnCl_2$, 20 mM Tris-HCl pH 7.3), and a Mg DNase buffer (5 mM $MgCl_2$, 20 mM Tris-HCl pH 7.3). In both cases the RNA sample was incubated for 30 minutes at 37°C in DNase reaction cocktails containing,1x DNase buffer, 1 mM DTT, 40 units of RNase Inhibitor (Roche) and 1 unit of DNase I (Roche). The DNase I enzyme was then inactivated by incubation at 75°C for 5 minutes. Neither of these conditions degraded RNA, the Mn buffer was preferred[2].

2.6 5' Rapid amplification of cDNA ends (5' RACE)

2.6a. Summary of 5' RACE method

The method of 5' RACE allows cloning of the terminal 5' end of transcripts (see Schaefer 1995 for review). The procedure is summarised in Figure 2.1.

[2] DNase I cleaves faster with Mn cofactors, and produces blunt ended products. With a Mg cofactor DNase I produces random-length overhangs which can be amplified by PCR (Bauer 1994)

2.6b. Special precautions for 5'RACE experiments

Filter tips were used to prepare RNA extracts, and for all subsequent steps of 5'RACE . Because contamination of samples with was observed in several experiments, autopipette barrels were washed with hot water, and then with 80% ethanol prior to use. Benches were wiped periodically with 0.01 M HCl, and then with water. Trays used to hold microcentrifuge tubes and PCR tubes were periodically soaked overnight in 0.1 M HCl, and rinsed with water. These trays were not used with tubes containing plasmids or the products of PCR reactions.

1. Reverse transcription
First strand cDNA is synthesised from polyA+ mRNA.

2. Synthesis of homopolymer tail
Terminal transferase adds dATP to ends of first strand cDNA.

3. Second strand cDNA synthesis
PolyT of ANCHORT_{16}V primer anneals to polyA sequence. ANCHOR is attached.

4. Primary PCR
ANCHOR primes at one end, gene specific DJ201 primer at other end

5. Secondary semi-nested PCR
DJ203 primer adds more specificity to the PCR reaction

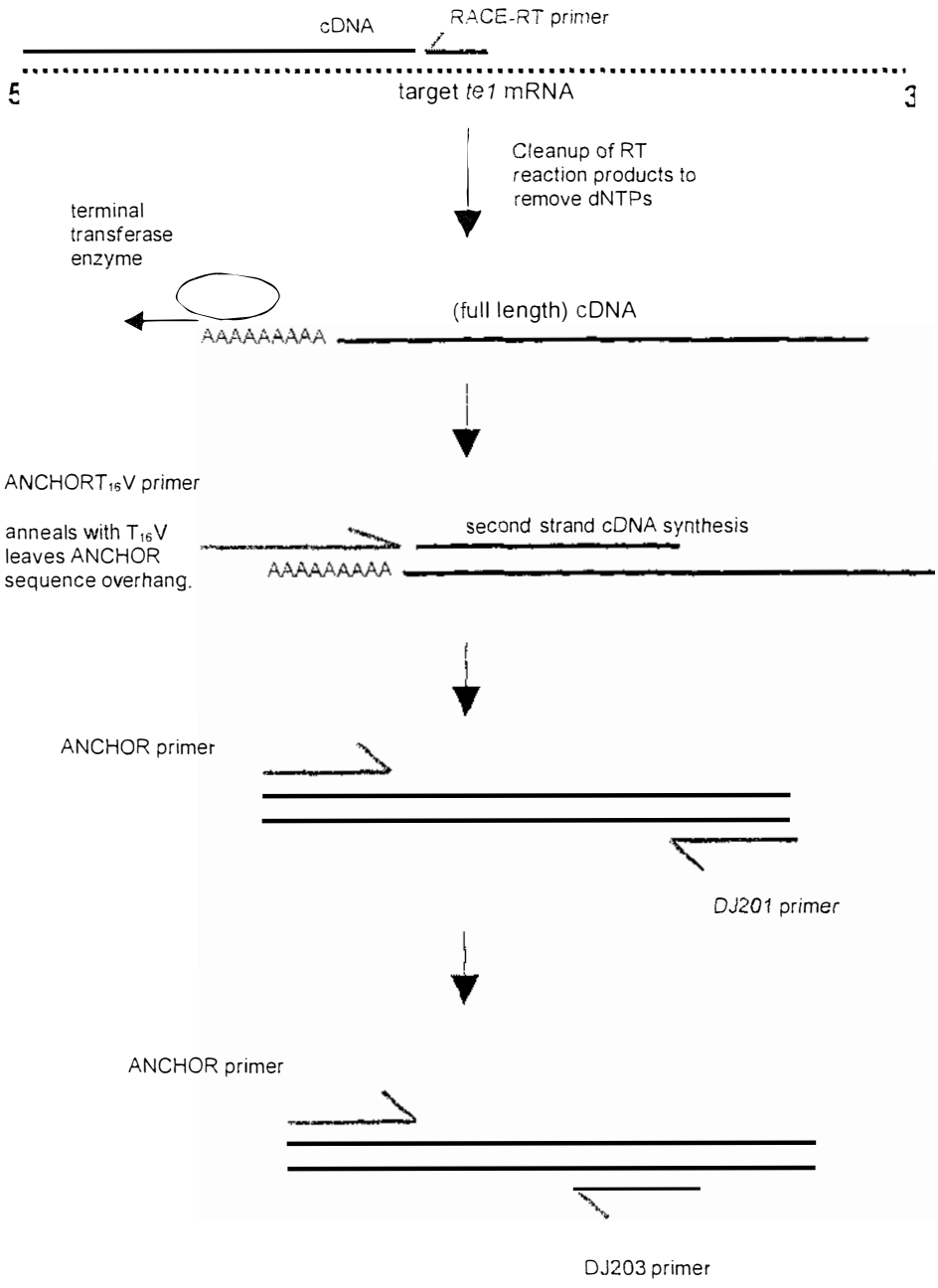


Figure 2.1. Overview of the 5' RACE protocol. The main steps of the 5' RACE protocol used are 1) reverse transcription primed with a gene specific oligonucleotide primer, RACE-RT, 2) synthesis of a homopolymer tail (dATP in this case) to the first strand cDNA with terminal transferase, 3) second strand cDNA synthesis with a chimeric ANCHOR T_{16}V primer, 4) primary PCR with ANCHOR primer and DJ201 gene specific primer, 5) secondary PCR with ANCHOR primer and a DJ203 gene specific primer, which is upstream of DJ201.

2.6c. Extraction of polyA+ RNA direct from tissue lysates

All solutions were freshly prepared and chilled on ice before use. 200 μg of Dynabeads Oligo$(dT)_{25}$ (Dynal) were equilibrated in lysis/binding buffer (100 mM Tris-HCl pH8.0, 500 mM LiCl. 10 mM EDTA, 1% SDS, 5 mM DTT) by collecting the beads with a magnetic particle collector (MPC), and resuspending them in 200 μl of lysis/binding buffer at least twice, or until the pH of the buffer removed was 8.0.

Approximately 100 mg of tissue was ground to a fine powder in liquid nitrogen with a baked mortar and pestle (see results for particulars of tissues used in each experiment). As soon as the liquid nitrogen had evaporated, but before the ground tissue had thawed, it was scooped into another baked mortar containing 1 ml of lysis/binding buffer, and ground again to mix. The lysate was then pipetted into a sterile microfuge tube and stored on ice until the remaining samples had been ground. Lysates that were particularly viscous were sheared by passing them through a 12 gauge needle several times. All samples were then centrifuged at 20 800 xg at 4°C for 2 minutes to pellet insoluble material.

The supernatant was added to the washed beads, gently mixed, and incubated for 3-5 minutes at room temperature. The beads were then washed in the following solutions by immobilising them with the MPC, removing the supernatant and adding the new solution: beads were washed twice in 500 μl of wash buffer (10 mM Tris-HCl pH 8.0, 150 mM LiCl, 1 mM EDTA), then three times with wash buffer containing 0.1% SDS. RNA was eluted from beads by addition of 30 μl of 10 mM Tris-HCl pH 8.0 that had been pre-heated to 65°C, gentle resuspension of the beads, and incubation at 65°C for 2 minutes. Finally , the beads were captured with the MPC, and the supernatant removed to another tube. An aliquot of the RNA was quantified by spectrophotometry (1.0 absorbence at 260 nm wavelength is 40μg/ml), the remainder was stored at -80°C.

Dynabeads were regenerated between re-use by resuspending the beads in 200 μl of 0.1 M NaOH, mixing, and incubating at 65°C for 2 minutes. Beads were recovered in the MPC for at least 30 seconds, and the supernatant was removed. Beads were washed twice more in this way. Beads were washed three times in 200 μl of storage buffer (250

mM Tris-HCl pH8.0, 20 mM EDTA, 0.1 % Tween 20, 0.02 % sodium azide). Beads were stored in 200 µl of storage buffer at 4°C between uses.

2.6d. Suitability of RNA samples for 5' RACE.

There were two criteria that RNA samples were required to satisfy before use in any 5' RACE experiment; they must not contain contaminating DNA, and they must not be degraded.

A PCR assay was used to exclude the possibility that any final products were the result of PCR amplification of genomic DNA in the RNA sample. A PCR reaction conducted according to the standard PCR reaction protocol described in Section 2.3c using the primers TG4 and TE15 (positions 3064 and 4694 respectively), and program DJ60x45L (see Appendix 2 for description of thermocycling conditions) with a 2:00 minute extension time produced a clear product from 50 pg (approximately 16 haploid genomes) of maize genomic DNA. To establish that polyA+ RNA preps that were to be used for 5'RACE were free of genomic DNA, 100ng of the RNA sample was required to produce no product when used as a template for this reaction. A positive control of 50 pg of maize genomic DNA was always included. Since initial polyA+ RNA samples prepared with Dynabeads were found to be contaminated with genomic DNA by this assay, samples were routinely treated with DNase I (Section 2.5b) using the Mn buffer, before use in 5'RACE experiments.

To establish that RNA samples were not degraded during extraction or DNase I treatment, RNA was examined by northern blotting (Section 2.8). Since the *te1* transcript itself was difficult to detect by northern blotting, membranes were probed with a ubiquitin cDNA (Christensen 1992), as described in Section 2.8c. A discrete band was taken as an indication that the RNA was not degraded.

2.6e. Reverse transcription reaction for 5' RACE

PolyA+ RNA was reverse transcribed in 30 to 40 µl volume using Moloney Murine Leukemia Virus (MuMLV) reverse transcriptase (Expand Reverse Transcriptase, Roche) primed with the RACE-RT primer (see Appendix 2). 100 ng of polyA+ RNA prepared with Dynabeads (Section 2.6c) was denatured in 20 µl volume containing 10 µM RACE-RT primer, 1x Expand Reverse Transcriptase buffer (50 mM Tris, 40 mM KCl, 5 mM $MgCl_2$) (Roche), 10 mM DTT, and 40 units of RNase Inhibitor (Roche) by heating to 75°C for 5 minutes, and then cooled to 45-48°C before reverse transcription. A reverse transcription

cocktail of 20 μl volume containing of 1x Expand Reverse Transcriptase buffer, 10 mM DTT, and 40 units of RNase Inhibitor (Roche), 500 μM dNTPs and 25 units of MuMLV (Expand Reverse Transcriptase, Roche) was then added, and the reverse transcription reaction was allowed to proceed for 30 minutes at 45-48°C. Note that the final concentration of dNTPs was 250 μM, and RACE-RT was 5 μM.

The RNA-cDNA hybrids were then denatured at 95°C for 2 minutes, and reactions were incubated at 37°C with 2 μl of 2 mg/ml DNase free RNase A. Samples of reverse transcription reactions were stored at -20°C. To remove excess nucleotides from the RT reaction products so that the subsequent tailing reaction product would be a homopolymer, samples were purified using Concert PCR Extraction Kits (Gibco BRL), as recommended by the manufacturer except that the cDNA was eluted in 30-50 μl of sterile Milli-Q water.

2.6f. Synthesis of homopolymer tail.

Since the *te1* genomic sequence is relatively GC rich, dATP was used as the substrate for the terminal transferase ('tailing') reaction. An important variation in the 5' RACE protocol was the buffer used for the terminal transferase reaction; initial experiments used the buffer supplied with the Terminal Transferase enzyme (final concentration 200 μM potassium cacodylate, 25 mM KCl, 25 mM Tris-HCl pH 6.6, 25 μg/ml bovine serum albumin) (Roche), later experiments used the Roche PCR buffer (final concentration 20 mM Tris-HCl pH8.4, 50 mM KCl, 1.5 mM $MgCl_2$). The PCR buffer was preferred to the buffer supplied with the Terminal Transferase (TdT) enzyme (see discussion Section 2.11). Tailing reactions were set up in 20 μl volume and contained between $1/5^{th}$ and $1/12^{th}$ of the purified RT reaction product. Reaction cocktails contained either 1x Roche PCR buffer or 1x TdT buffer, 200 μM dATP and 5 units of TdT enzyme. A thermocycling program was produced that allowed to the reaction proceed for 10 minutes at 37°C, and were then stopped the reaction by incubation at 65°C for 15 minutes.

2.6g. Second strand cDNA synthesis and primary PCR Reaction

Second strand cDNA synthesis occurred in 'Stage 1' of the RACE7 PCR program, followed by primary PCR during 'Stage 2' of the PCR program RACE7 (see Appendix 2 for a full description of the thermocycling conditions). Stage 1 contained only the

ANCHORT_{16}V, the ANCHOR primer and DJ201 primer were added at the start of Stage 2.

The tailing reaction product (5 μl) was added to a 30 μl PCR cocktail with a final concentration of 1x PCR buffer (10 mM Tris-HCl pH8.3, 50 mM KCl, 1.5 mM $MgCl_2$) (Roche), 1x PCR Enhancer solution (Gibco BRL), 250 μM dNTPs and 125 nM ANCHORT_{16}V. A Taq mix containing 1x PCR buffer and 2.5 units of Taq (Roche) was added at 95°C, after the first 98°C denaturation.
Either the RACE6, or the RACE7 PCR programs (Appendix 2.) were run until the first 95°C denaturation of Stage 2, when 10 μl of primer cocktail containing 1x PCR buffer, 2 μM ANCHOR primer and 2 μM DJ201 primer was added. The final concentrations ANCHOR and DJ201 primers were 400 nM as per a standard PCR reaction (Section 2.3c).

2.6h. Secondary semi-nested PCR reaction.

Secondary semi-nested PCR reactions used the ANCHOR 5' primer and the DJ203 3' primer which is 66 bp upstream of the DJ201 primer used for primary PCR (Section 2.6g). 0.5 μl or 5 μl of the primary PCR products as templates in a standard PCR reaction protocol (Section2.3c) using the primers ANCHOR and DJ203, the program DJ68x35 with a 1:00 minute extension time. The positions of all primers used in 5' RACE are shown in Figure 2.3.

2.7 Quantification of *te1* transcripts by competitive RT-PCR

2.7a. Summary of competitive RT-PCR procedure.

Total RNA was extracted from a variety of normal maize tissues and examined for contamination, equal loading, and absence of degradation. Suitable RNA samples were reverse transcribed using a polydT primer. To control for variability between samples in the reverse transcription and PCR reactions, a known quantity of an altered synthetic *te1* transcript was added to each RNA sample prior to reverse transcription. This altered synthetic transcript was differentiated from native transcripts by the absence of a *Bgl*II restriction site. A PCR reaction was used to amplify both native and synthetic *te1* cDNAs using a primer set that does not amplify DNA derived from maize genomic DNA. PCR products were cut with *Bgl*II and electrophoresed with conditions that allowed the native

products (cut with *Bgl*II) to be distinguished from the synthetic products (not cut with *Bgl*II).

2.7b. Extraction of total RNA with Qiagen RNeasy Plant Mini Kit

The following tissues were used for RNA extraction, all are from plants of normal B73 inbred;

1. entire embryos collected 21 days after fertilisation
2. SAM-enriched tissue collected 20 days after germination (see note below)
3. early tassels of approximately 1 cm in length collected 70 days after germination
4. late tassels of approximately 2 cm in length collected 95 days after germination
5. ears of approximately 2 cm in length collected 95 days after germination
6. root tips comprising the distal 1 cm of roots.

SAM-enriched tissue consists of the extreme upper region of the culm (approximately 0.5 cm) and the shoot apex covered by several young leaves. Several samples from a population were dissected, and the meristems examined with a dissecting microscope before collection. The maize inflorescence meristem is easily distinguishable from a vegetative meristem, only populations showing exclusively vegetative SAM morphology are collected. All tissue was dissected directly into liquid nitrogen, and stored at -80°C. Approximately 0.1 g of tissue was ground in liquid nitrogen to a fine powder in baked mortars and RNA was extracted using the RNeasy Plant Mini Kit (Qiagen), according to the manufacturers instructions using the supplied guadinium isothiocyanate lysis buffer ("RCL Buffer"). RNA was eluted from the column in 60 μl of RNase-free Milli-Q water, and stored at -80°C in four 15 μl aliquots. Since the yield of RNA from tassels was approximately 1/3rd that of all other tissues, RNA was also extracted from tassels using the acid guadinium thiocyanate-phenol-chloroform RNA extraction protocol described below.

2.7c. Acid guadinium thiocyanate RNA extraction methods

This protocol was modified from Chomczynski (1987). Approximately 1g of tissue was ground under liquid nitrogen to a fine powder in a baked pestle. As soon as the liquid nitrogen had evaporated, but before the ground tissue had thawed, it was scooped into another baked mortar containing 5 ml of guadinium denaturing solution (5 M guadinium thiocyanate (Sigma, SigmaUltra grade), 25 mM tri-sodium citrate (BDH, GPR grade), 1%

N-lauroylsarcosine (Sigma, SigmaUltra grade), 0.1% 2-mercaptoethanol (2-ME) (BDH, 99%), 1.75 mM diethyl dithiocarbamic acid (Sigma, SigmaUltra grade), 2-ME and diethyl dithiocarbamic acid were added immediately prior to use.

The lysate was then transferred to a new tube, and the following reagents were added sequentially; 0.33 ml of 3 M NaOAc pH4.5, 5 ml of water saturated phenol (Sigma) and 1.75 ml of SEVAG (1 volume isoamyl alcohol (Ajax, Analytical grade): 24 volumes chloroform (BDH, Analar grade)), mixing after each reagent. The tube was mixed by inversion for 30 minutes, then the organic and aqueous phases were allowed to separate before centrifugation at 4 000 rpm for 10 minutes at room temperature.

The aqueous phase was transferred to an RNase-free 50 ml Oakridge tubes (Oakridge tubes were soaked overnight in 0.1 M HCl then rinsed 3 times with sterile Milli-Q water), and an equal volume of chilled isopropanol was added. RNA was precipitated at -20°C for at least 2 hours, then recovered by centrifugation at 2,700 xg at 4°C for 20 minutes. The supernatant was removed and the pellet resuspended in 0.5 to 1.0 ml of resuspension solution (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.2% N-lauroylsarcosine) The resuspended RNA was extracted once with water saturated phenol:SEVAG (1:1), and several more times with Tris-HCl pH 7.5 saturated phenol:SEVAG (1:1), until the interface was clear. RNA was precipitated at -20°C with $1/10^{th}$ volumes of 3M NaOAc pH 7.5 and 1 volume of isopropanol for several hours or overnight.

The supernatant was removed, the pellet air dried, and resuspended in 50-100 µl of sterile Milli-Q water. RNA samples were quantified by spectrophotometry, typical 260/280 nm wavelength ratios were 1.8-2.0. In cases where RNA solutions were light brown coloured, or where 260/280 ratios were lower than 1.7, RNA was precipitated overnight at -20°C in 2 M LiCl, and resuspended in half the previous volume of sterile Milli-Q water.

A simplified version of this protocol, which is essentially as described by Chomczynski (1987), was used for RNA preparations that were intended for northern blotting rather than RT-PCR. Briefly, after the first isopropanol precipitation RNA was resuspended in 0.5 ml of guadinium denaturing solution, precipitated at room temperature with 0.5 ml of isopropanol, and collected by centrifugation at 20 000 xg for 20 minutes at 4°C. RNA was then resuspended in sterile Milli-Q water. Where necessary RNA was precipitated in 2 M LiCl as above.

2.7d. Suitability of RNA samples for competitive RT-PCR

As with 5' RACE, RNA samples were required to be free of contamination and not degraded before they were considered for competitive RT-PCR. Samples also needed to be accurately quantified so results could be compared between tissues.

A PCR assay was used to establish that RT-PCR products were not derived from contaminating DNA in the RNA samples. A standard PCR reaction protocol (Section 2.3c), except that primers DJ108 and DJ109 were used at 4 000 pM concentration rather than 400pM, produced a clear product from 1 fg of *te1* cDNA with PCR program DJ66x40.To establish that RNA samples were not contaminated, 1 μl (at least 200 ng) of each total RNA sample was required to produce no product when used as a template using these reaction conditions. This PCR assay did not produce any discrete product from 1 μg of B73 genomic DNA. RNA samples that were found to be contaminated were digested with DNase I using the Mn DNase I buffer as described in Section 2.5b.

RNA was first quantified by spectrophotometry. To establish that RNA samples were not degraded, and to confirm that absolute as well as relative quantification of RNA was accurate, an estimated 1 μg of each sample was glyoxylated, electrophoresed, and stained with ethidium bromide with a known quantity of mRNA molecular weight ladder as described in Section 2.8a. The mass of rRNA bands, which comprise the bulk of total RNA, was then compared both between samples, and to the known mass of the molecular weight ladder.

2.7e. Synthesis of altered *te1* transcript

The *te1* cDNA obtained by Bruce Veit as a pBluescriptSK$^+$ phagemid clone (pBV432) contains a unique *Bgl*II restriction enzyme site. This *Bgl*II site is 247 bp from the 5' end of the 291 bp PCR product produced by the DJ108/DJ109 primer set. So that a synthetic transcript could be produced whose RT-PCR product would lack the *Bgl*II site it was removed as follows; 2 μg of pBV432 was linearised with *Bgl*II, and the overhanging ends were filled in by adding 1unit of T4 DNA polymerase (Roche), 1 unit of Klenow enzyme (Roche) and dNTPs to 100 μM, and incubating at 37°C for 1 hour. Half this reaction was ligated with 2.5 units of T4 DNA ligase (Gibco BRL) in 1x T4 ligase buffer (50 mM Tris-HCl pH 7.6, 5 mM $MgCl_2$, 1 mM ATP, 1 mM DTT, 5% polyethylene glycol-8000) at room temperature overnight. The ligation reaction was then digested again with 2 units of *Bgl*II to remove unwanted recircularised products and transformed by heat shock into DH5α

cells. Plasmids were isolated by the STET method, and screened by *Bgl*II digest. This vector is referred to as pDJ16.

A synthetic transcript was produced from pDJ16 as follows. 2 μg of pDJ16 was linearised with *Xho*I, which cleaves only in the pBluescriptSK$^+$ polylinker at the 3' end of the *te1* cDNA. To remove RNases from the linearised plasmid, digestion was extracted twice with Tris-HCl pH 7.5 saturated phenol, and once with chloroform, and precipitated overnight with 1/10th volume of DMPC-treated 3 M NaOAc and 2.5 volumes of ethanol. The vector was then resuspended in 10 μl of sterile Milli-Q water. A 20 μl transcription reaction containing 1x transcription buffer (40 μM Tris-HCl pH8.0, 6 mM $MgCl_2$, 10 mM DTT, 20 mM spermidine), 1 mM nucleotide triphosphates (NTPs, Roche), 1 μg of linearised pDJ16, 40 units of RNase inhibitor (RNasin, Promega), and 40 units of T3 RNA polymerase (Roche) was incubated for 1 hour at 37°C. The pDJ16 DNA template was removed from half the transcription reaction by adding 10 units of DNase I (Roche), Mn DNase I buffer to 1x concentration (1 mM $MnCl_2$, 20 mM Tris-HCl pH 7.3), and 20 units of RNase inhibitor (RNAsin, Promega), and incubation at 37°C for an hour. The DNased transcription reaction was precipitated overnight at -20°C with 1.25 μl of 8 M LiCl, 0.5 μl of 0.5M EDTA, and 75 μl of ethanol, and resuspended in 20 μl of sterile Milli-Q water. The pDJ16 transcript was quantified by spectrophotometry, as described in Section 2.6c. A sample of the transcript was examined for degradation by electrophoresis and ethidium staining as described in section 2.8a.

2.7f. Reverse transcription reactions for RT-PCR

RNA samples prepared with either RNeasy kits (Section 2.7b), or the modified acid guadinium thiocyanate protocol (Section 2.7c) that had been shown to be suitable for RT-PCR (Section 2.7d), were reverse transcribed as follows.

1 μg of total RNA derived from tissues described in Section 2.7b (embryo, SAM, early tassel, late tassel, ear, leaf, root) was denatured in 10 μl volume with either 100, 10, or 1fg of the altered synthetic *te1* transcript (Section 2.7e), and 1 μl of 100 μM OLIGOdT-20 by heating for 5 minutes at 75°C. A tube containing only the OLIGOdT-20 primer at the same dilution, and no synthetic transcript was included as a negative control. The RNA was cooled to 37°C, and a reverse transcription mix of 15 μl was added to each tube to bring the final contents (in 25 μl volume) to 1x MuMLV reverse transcriptase buffer (50 mM Tris, 40 mM KCl, 5 mM $MgCl_2$) (Roche), 10 mM DTT, 40 units of RNase Inhibitor

(Roche), 1 mM dNTPs. 30 units of MuMLV reverse transcriptase (Expand Reverse Transcriptase, Roche). The final concentration of the OLIGOdT-20 primer was 5 μM. The reaction was allowed to proceed for 1 hour at 37°C, reactions were stored at 4°C until use in PCR. This reverse transcription procedure was performed three times with the full complement of RNA samples/controls (embryo, SAM, early tassel, late tassel, ear, leaf, root, no RNA). The volumes and solutions were identical, except for the levels of synthetic transcript added: 100 fg/tube of 1μg of total RNA in the first experiment, 10 fg/μg in the second, 1 fg/μg in the third.

Reverse transcription reactions used to confirm 5' RACE experiments used 100 ng of polyA+ RNA (prepared as in Section 2.6c.) that had been digested with DNaseI (Section 2.5b). The reaction was performed as described above except that 100 pmol of random hexamer primer was used and the reaction was allowed to proceed at 42°C for 15 minutes.

2.7g. Competitive PCR

Each reverse transcription reaction product (5μl of the 25 μl total RT reaction volume) was amplified in a 20 μl PCR reaction using the GC-Rich PCR System (Roche). Final reagent concentrations were; 0.5 M betaine, 1x GC-Rich PCR buffer (includes1.5 mM $MgCl_2$, 5% DMSO[3]) (Roche), 4 μM DJ108 and DJ109 primers (10x the concentration of the standard PCR reaction protocol), 250 μM dNTPs, 1 unit of GC Rich enzyme mix (Roche). Taq polymerase was excluded from the initial reaction mix of 15 μl volume, a 5 μl volume enzyme mix containing 1x GC Rich PCR buffer and 0.5 μl (1 unit) of GC-Rich enzyme mix (Taq polymerase and Tgo polymerase) (Roche) was added to each tube during the first 95°C denaturation step. A negative control sample, containing water, and a positive control sample containing 1 fg of native *te1* cDNA were also included.

PCR products were cut with *Bgl*II by adding 5 μl of a *Bgl*II mix containing 4 units of *Bgl*II (Roche), and 1x restriction enzyme buffer H (50 mM Tris-HCl, 10 mM $MgCl_2$, 100 mM NaCl, 1 mM DTE, pH 7.5) (Roche) directly to each PCR reaction, and incubating for an hour at 37°C. 5 μl of samples were electrophoresed in a 2.5% TAE gel at 50V, stained with ethidium bromide, and visualised with an IS1000 Digital Imaging System (Alpha Innotech Corp.).

---

[3] Roche will not release details of this buffer.

2.8. Northern blotting

2.8a. Electrophoresis of glyoxylated RNA.

A 0.1 M sodium phosphate buffer (NaPB) stock was prepared by mixing 1M NaH_2PO_4 (approximately 42 ml), and 1 M Na_2HPO_4 (approximately 58 ml) until the pH is 7.0, diluting 1:10 with water to 0.1 M, and DMPC/DEPC treatment.

RNA was glyoxylated in a glyoxylation mix containing 1 M deionised glyoxal (Sigma), 50% DMSO (Fluka Molecular Biology grade), 10 mM NaPB pH7.0 for 1 hour at 50°C. Glyoxal was deionised as described in Sambrook (1989).

The gel apparatus and silicon tubing for a peristaltic pump (Cole Parmer Instrument Co.) was treated with approximately 0.3% H_2O_2 (Andrew Industrial Ltd.) overnight to inactivate RNases. Molten 1.0% agarose was prepared in baked glassware in 10mM NaPB buffer. To inactivate RNases in the molten agarose sodium iodoacetate (Merk-Schuchardt) to 10 mM concentration was mixed into the molten agarose once it had cooled to approximately 70°C. Once the agarose had set, 10 mM NaPB buffer was circulated over the gel for an hour prior to electrophoresis.

Glyoxylated RNA was mixed with 1 µl of RNA loading buffer (50% glycerol, 10 mM NaPB, 0.25 % bromophenol blue, 0.25% xylene cyanol, DMPC treated) per 15 µl of sample and loaded onto the agarose gel. RNA was electrophoresed at 45 mA for 30 minutes without buffer circulation, and for approximately 4 hours at 45 mA with vigorous circulation with the peristaltic pump (approximately 150 ml/minute, total buffer volume was 1.2 L). 3 µl of Gibco BRL 0.24-9.5 kb RNA molecular weight ladder was run in the first lane of the gel.

The portion of the gel that was to be stained with ethidium bromide (including the molecular weight marker) was cut from the remainder of the gel with a sterile scalpel blade, those lanes that were to be transferred to a membrane were not stained with ethidium bromide. Staining was carried out in trays protected from light with aluminium foil, gels were soaked for 10 minutes in 50 mM NaOH, stained for 30 minutes in 10 µg/ml ethidium bromide in 0.1 M ammonium acetate (NH_4OAc), and destained for 30 minutes in 0.1 M NH_4OAc. Images of ethidium stained gels were captured under ultraviolet light with an IS1000 Digital Imaging System (Alpha Innotech Corp.), with a ruler included for molecular weight calibration.

2.8b.Transfer of RNA to membranes

RNA was transferred to positively charged nylon membranes (Nylon+, Roche) with a downward transfer blotting stack (as described in Section 2.9e) wet with DMPC-treated 20 x SSC (3 M NaCl, 0.3M sodium citrate pH 7.0) without further treatment of the gel. Stacks were usually left to transfer overnight. RNA was fixed by crosslinking with 120 000 μJ/cm^2 ultraviolet light in a UV Stratalinker 2400 (Stratagene). Membranes were then dried at room temperature, sealed in a plastic bag, and stored at 4°C.

2.8c. Synthesis of radiolabelled probes.

Random-primed α^{32}P dCTP labelled probes were synthesised from templates of either PCR products (*ROC1* and *Ubiquitin* probes), or the entire *te1* cDNA cut from the vector and gel purified.

The following templates were used for probes,

a) *Ubiquitin*. A PCR product was synthesised from the pSKUBI vector (Green 1994) using a standard PCR protocol (Section 2.3c), PCR program DJ55x35, M13 forward and reverse primers, and 100 pg of pSKUBI vector template.

b) *ROC1*. A PCR product was synthesised from a vector containing the *ROC1* cDNA (Lippuner 1994) using a standard PCR protocol (Section 2.3c), PCR program DJ55x35, M13 forward and reverse primers, and 100 pg of *ROC1* vector template.

c) *te1*. The 2.4 kb cDNA insert cloned into pBluescriptSK$^+$ (pBV432), was cut from the vector by digestion with restriction enzymes *Eco*RI and *Xho*I. The digest was electrophoresed, and stained as described in Section 2.3b, and the insert was cut from the gel with a sterile scalpel blade, using a long wavelength UV lamp to visualise the band. The DNA was extracted from the gel slice with a Concert Gel Extractions Kit (Gibco BRL), and eluted in 30 μl of sterile Milli-Q water.

In each case radiolabelled probes were synthesised as follows; probe templates were quantified by ethidium stained 100 ng of template DNA was diluted to 7 μl volume in sterile Milli-Q water, boiled for 5 minutes in a 1.8 ml micorcentrifuge tube, then immediately chilled on ice for 5 minutes. Templates were mixed on ice with a priming reaction mix to a final concentration of 75 μM random hexamer primer, 375 μM of dATP, dTTP and dGTP, 2 units of Klenow enzyme (Roche) to a volume of 15 μl. Finally, 5 μl (50 μCi) of α^{32}P dCTP (NEN Research Products) was added and mixed. The reaction

was incubated at 37°C for 1-2 hours. Probes were then diluted to 50 µl volume with 30 µl of sterile Milli-Q water, and the unincorporated nucleotides were removed with Sephadex G-50 spin columns (ProbeQuant™ G-50 Micro Columns, Pharmacia). Probes were used immediately or stored at -20°C.

2.8d. Hybridisation of RNA membranes.

All hybridisation procedures used H_2O_2 treated hybridisation tubes, or H_2O_2 treated plastic boxes. Immediately prior to pre-hybridisation, RNA was de-glyoxylated by treating the membrane with 20 mM Tris-HCl pH8.0 at 65°C for 5 minutes. Membranes were then pre-hybridised for an hour in at least 20 ml of Church/Gilbert hybridisation buffer (0.5 M $NaHPO_4$ pH 7.2 [2 M $NaHPO_4$ stock is composed of 35.6g of Na_2HPO_4 and H_3PO_4 to pH 7.2], 7% SDS, 1 mM EDTA) (Church 1984) at 65°C. Probes were boiled for 5 minutes in 500 µl of sterile Mill-Q water, chilled on ice for 5 minutes, and added to 10 ml of 65°C Church/Gilbert hybridisation buffer. Membranes were hybridised overnight at 65°C. Membranes were washed twice in 2 x SSC, 0.5% SDS in hybridisation tubes at 65°C, then twice in 2 x SSC, 0.5% SDS in boxes at room temperature to remove excess probe. Stringentcy washes were performed in 0.1 x SSC, 0.1% SDS at 65°C for 5 to 15 minutes. Membranes were sealed in plastic bags before being exposed to film.

2.9. Analysis of *te1* alleles

2.9a. Summary of *te1* allele analysis

Nine recessive mutant alleles of the *te1* gene had been identified (Table 2.1). The exact positions of *Mutator* elements in the *te1-mum1* and *te1-mum2* alleles, and the approximate position of the *te1-mum3* allele were known (Veit 1998). DNA was extracted from homozygous mutant plants of each allele (Section 2.9b), the normal B73 inbred, and several other inbred lines, and was used for analysis by PCR (Section 2.6c) and Southern blotting (Section 2.6d).

Table 2. 1 *te1* mutant alleles

Allele	Mutation by	Notes
te1-1	not known	
te1-mum1	Mu8 element	Mutator insertion after 4651 bp, in exon three.
te1-mum2	Mu8 element	Mutator insertion after 4395 bp, in exon three.
te1-mum3	Mutator element?	Approximate position previously identified as intron 1. This study found a Mutator element 3' occurs after position 4017 in the *te1* genomic sequence. Previous nomenclature is *te1-mum4*
te1-mum4	Mutator element?	Previous nomenclature is *te1-mum3*
te1-mum5	Mutator element?	S Briggs, unpublished.
te1-mum6	Mutator element?	S Briggs, unpublished.
te1-mum7	Mutator element?	S Briggs, unpublished.
te1-ems	ethylmethane sulfonate (EMS)	Identifed in screen for modifiers of *liguless2* L Harper, unpublished.

2.9b. Extraction of maize genomic DNA

Filter tips, and baked mortars and pestles were used for this procedure when DNA was to be used for PCR. All phenol/SEVAG solutions were from stocks set aside for use with genomic DNA that were not used with plasmids. Approximately 1 g of leaf tissue was ground to a fine powder in liquid nitrogen, and scooped into a 15 ml falcon tube containing 5 ml of urea extraction buffer (6.9 M urea, 350 mM NaCl, 50 mM Tris-HCl pH8.0, 0. 20 mM EDTA, 1% N-lauroylsarcosine) and mixed. 5 ml of Tris-HCl pH 7.5 saturated phenol:SEVAG (1:1) was then added and mixed by inversion for 15 minutes. Phases were separated by centrifugation at 4 000 rpm for 20 minutes at room temperature. The aqueous phase was removed to a new tube, and the nucleic acids were precipitated with by adding $1/10^{th}$ volume of 4.4 M NH_4OAc, and an equal volume of isopropanol and mixing gently. The strands of genomic DNA were spooled with a Pasteur pipette with a hook formed under a Bunsen burner flame, rinsed on the hook with 80% ethanol, and transferred to a 1.8 ml tube. The DNA pellet was air dried and resuspended in 100 µl of TE50/5 (50 mM Tris-HCl pH7.5, 5 mM EDTA).

2.9c. PCR analysis of *te1* mutant alleles

Two types of primer sets were used in attempts to clone PCR products. Firstly, one of two primers (MUEND and MUSEL, see Appendix 1.) designed to *Mutator* element ends (facing outward), was combined with a primer designed to the *te1* genomic sequence. Products indicated the juxtaposition of a *Mutator* element and the *te1* gene, and these were cloned and sequenced. Since the position of the *te1-mum1* allele was known, this allele was used to establish which *te1* primers and conditions produced products with MUEND and MUSEL primers.

Secondly, pairs of primers designed to the *te1* gene were used. Any PCR products obtained from *te1* mutant alleles were compared to the product from the normal allele (B73 inbred line). Products that differed in size from the B73 line were cloned and sequenced. Analysis of the cases where a primer pair produced a single, abundant product from the B73 line but no product from a mutant line also served to delineate the position of the lesion.

PCR reactions utilised standard PCR reaction protocol (Section 2.3c), except in cases where long (~ 3 kb) PCR products were expected when 2.5 units of Expand Long Template enzyme mix (Roche) and 1x Expand Long Template buffer replaced Taq and the standard PCR buffer. 100 ng of genomic DNA extracted as described in Section 2.9b was used as templates for reactions. Optimal annealing temperatures for particular primer sets were established with either *te1-mum1* DNA (where MUEND and MUSEL primers were used), or B73 inbred DNA (where two *te1* primers were used), these samples were used as a positive control for subsequent reactions. Typical PCR programs used the standard $DJT_axN_c(L)$ template described in Appendix 2. with 35 to 45 cycles, and 1:00 min to 4:30 extension times. Exact conditions are described with results. PCR reaction products were electrophoresed as described in Section 2.3b.

2.9d. Cloning and sequencing of PCR products

PCR products were either cut from agarose gels, where the product was heterogeneous, and purified using Concert gel extraction kits (Gibco BRL), or for homogeneous products, purified directly from the PCR reaction with Concert PCR cleanup kits (Gibco BRL). DNA was eluted in sterile Milli-Q water. PCR products were then analysed by restriction digests, or cloned into pGEM®-T Easy (Promega) using pGEM®-T Easy Vector System I (Promega), according to the manufacturers instructions, and transformed into *E. coli* DH5α by heat shock (Section 2.2e). Transformants were screened by colony

PCR (Section 2.3d), or by extracting plasmids using the STET method (Section 2.2g), and analysis by restriction digest to confirm the insert. Once vectors with the desired insert had been identified they were sequenced as described in Section 2.4c.

2.9e. Southern Blotting with the DIG System

Samples of DNA that were to be used for Southern blotting were carefully quantified by spectrophotometry before restriction digestion, according to the formula 1.0 A_{260} = 50 μg/ml dsDNA. 1-2 μg of maize DNA was found to be optimal for detection of digoxygenin labelled probes with the DIG System (Roche), 100 ng of *Arabidopsis* DNA was sufficient. DNA was typically digested with 5 units of restriction enzyme, and 0.5 μl of 10 mg/ml RNase A at the recommended temperature for the restriction enzyme for at least two hours. When recommended temperature exceeded 37°C, digestions were performed in a thermocycler with a heated lid.

Digested DNA was electrophoresed in 0.8% TAE gels at 70V overnight in an 20 cm x 25 cm gel tray, in a buffer circulating gel apparatus (Owl A5 Buffer Puffer™). 4 μl of DIG-labelled λ *Hind*II digest (Roche DNA Molecular Weight Marker II, DIG labelled) was run at both edge lanes. The entire gel was then stained with ethidium bromide, and an image captured as described in Section 2.3b.

Gels were soaked with gentle agitation twice for 15 minutes in 500 ml of denaturing solution (0.5 M NaOH, 1.5 M NaCl) and then twice for 15 minutes in 500 ml of neutralisation solution (1 M Tris-HCl pH7.2, 1.5 M NaCl). The gel was then transferred onto positively charged nylon membranes (Nylon+, Roche) by capillary transfer using a downward blotting stack[4] using 20x SSC (3 M NaCl, 0.3M sodium citrate pH 7.0) as the transfer buffer. Blotting stacks were left to transfer overnight, and then both sides of the membrane were crosslinked with 120 000 μJ/cm^2 ultraviolet light in a UV Stratalinker 2400 (Stratagene). Membranes were then dried at room temperature, sealed in a plastic bag, and stored at 4°C.

[4] Downward transfer blotting stacks consisted of (from the bottom up); a 10-15 cm stack of paper towels, two sheets Whatman 3MM Chromatography paper the same size as the paper towels, nylon membrane, gels, four sheets of Whatman 3M paper the same size as the gel, a wick of two sheets of Whatman 3M paper that overlays the stack and feeds into two buffer reservoirs (containing 1 L of 20x SSC) on either side of the stack, the gel tray, a weight of approximately 200 g. All layers apart from the paper towels were pre-wet with 20x SSC.

Digoxigenin probes were synthesised by PCR according to the standard PCR protocol (Section 2.3c) except that a DIG labelling mix was used in place of dNTPs. Final concentrations of nucleotides were 250 μM dATP, 250 μM dCTP, 250 μM dGTP, 166 mM dTTP, 40 μM DIG-dUTP. Conditions for particular primer sets were optimised with standard (non labelled) PCR reactions, and 50 to 100 pg of the PCR product from the optimised reaction was used as templates for the labelling PCR reaction. Labelled PCR products were examined by gel electrophoresis for quality, expectations were of a single band of higher molecular weight than the unlabelled PCR product, and quantity by comparing to a known mass of a molecular weight marker.

Membranes were pre-hybridised for an hour in 20 ml of DIG Easy Hyb (Roche) at 65°C. Approximately 50 ng of DIG labelled probe was diluted in 0.5 ml of sterile Milli-Q water, boiled for 5 minutes, then immediately chilled on ice for 5 minutes. The probe was added to 10 ml of fresh DIG Easy Hyb, and hybridised overnight at 65°C.

Membranes were washed twice in 2x SSC, 0.5% SDS in hybridisation tubes at 65°C, then twice in 2x SSC, 0.5% SDS in boxes at room temperature to remove excess probe. Stringentcy washes were performed in 0.1 x SSC, 0.1% SDS at 65°C for 5 to 15 minutes. Detection of DIG labelled probes was performed as recommended by Roche. The CDP-Star™ substrate was used, Anti-Digoxigenin-AP antibody was used at 1:20 000 dilution. Membranes were exposed to film for 30 min to 2 hours.

RESULTS (2)

2.10. Genomic sequence of the *te1* gene

All sequence positions are numbered from the 5' *Kpn*I site unless stated otherwise. The 8039 bp of sequence produced from the *te1* genomic clone (pBV304) is represented on the following page (Figure2.2), showing the positions of introns, the three RRMs, and putative transcription start sites. The sequence is shown in Figure 2.3. The *te1* gene contains five introns, four of which are in comparable positions to the *Arabidopsis* orthologs *TEL1* and *TEL2* (conservation of intron structure is discussed in more detail in Section 3.6). Like all seven the *Arabidopsis Mei2*-like genes, *te1* has an intron after the exon encoding the absolutely conserved IKNIPKY sequence at the N-terminus of RNA recognition motif three. Most introns are of the order of 100 bp in length, the exception is intron 4 (1037 bp). *Arabidopsis Mei2*-like genes do not have a particularly long intron in the corresponding position. This sequence was submitted to Genbank, Accession number AF348319.

Monocot consensus splice sites were present at most boundaries of genomic sequence and *te1* cDNA sequence, the exception was the 5' splice site of intron 3 (see Table 2.2). Sequencing reactions were at fourfold coverage of this region, two independent sequences from each of two pBV304 subclones.

Table 2. 2 Splice sites in the *te1* gene

	5' Spice site *(exon/intron)*	3' Spice site *(intron/ exon)*
Monocot consensus*	AG/**GT**AAGT	TGC**AG**/GT
intron 1	CG/**GT**ACGC	CGC**AG**/GG
intron 2	AG/**GT**AAAC	CCT**AG**/GC
intron 3	AG/**G**GTATC	CGC**AG**/CC
intron 4	AA/**GT**GAGT	TGC**AG**/CA
intron 5	AA/**GT**ACTG	CGC**AG**/GG

*Absolutely conserved nucleotides are in bold type. Splice site consensus from Simpson (1996).

figure: 2.2 te1 bar diagram
one page

Figure 2.3. Genomic sequence of the *te1* gene.

The genomic clone was obtained by Bruce Veit from a genomic library, and consisted of a 12.4 kb KpnI fragment. This genomic sequence consists of 8039 bp of this clone, from the 5' KpnI site to the XhoI site of the subclone pBV312. Lower case letters indicate that multiple coverage did not achieve an absolute consensus. IUPAC symbols are used for ambiguous sequence. The Genbank accession of this sequnce is AF348319.

Where the sequence is part of the transcript, the corresponding sequence is shown in bold. The amino acid coding is shown underlying exons. Exon positions are (1) 2971-3940, (2) 4043-4211, (3) 4297-4783, (4) 4891-5013, (5) 6050-6191, (6) 6310-7095. The transcription start site predicted by 5' RACE experiments (Section 2.11) is at 2971 bp.

The positions of primers that were used for 5' RACE (Section 2.3) (RACE-RT, DJ201, DJ203) are indicated.

The exact positions of *te1* alleles are indicated by triangles. ▽ Positions are; *te1-mum1*, 4653 bp (exon 3) (Veit 1998); *te1-mum2*, 4396 bp (exon 3) (Veit 1998); *te1-mum3*, 4017 bp (intron 1).

```
1       GGTACCCCTAATTATGGTCCCCGACAGAAATAAAGTCGAATCCGAAAGAGAGGCGAAAAC
        AATCACAAGCAAATAGAGCGGATGACACGGTGATTTGTTTTACCGAGGTTCGGTTCTTGC
        AAACCTACTCCCCGTTGAGGTGGTCACAAAGACCGGGTCTCTTTCAACCCTTTCCCTCTC
        TCAAACGGTCACCTAGACCGAGTGAGCTTTTCTCCTTAATCAACGGGTCACTTAGACCCC
        TTACAAGGACCACCACAACTTGGTGTCTCTTGAGTTTATTACAAGTTGCTTGAGAACAAG
        AATGGAGGAAGAAGAAAAGCGATCCAAGCGACAAGAACTCAAATGAACACAAATATCTCT
        CTCTCACTAGTTACTAAATGTTTGGAGTGATTGTGGACTTGGGAGAGGATTTGATCTCTT
        GTTTGTGTCTTGGAGTGAAGTCTAgAGCTCTTGTATTGAATGCAATGGCTGAAAACTTGG
        ATGCCTTGAAGTGGTGGTGGTTGGGGGTATTTATAGCCCCAACCACCAAAATGGCCGTTG
        GGGAGGCTGTCTGTCGATGGGCCGCACCCGGACAGTCCGGTGnnCnGGACACTGTCTGGT
        GCGCCACCACCGTCACCCAACCGTTAGGTTCTGACGGTTTCGACCGTTGGAGCTCTGACA
        GCTGGGACCACCAGACAGTCCGGTGGTGCACCAAgACAGTCACTGTTCACTGTCCGGksC
        GCCTTCTGGsckGCTCTtGACTCTGCGcGCGCTGTCCGCrgcATTGTCCACGTTCACTGT
        TCACTTTTGCArACGACCGTTGGCGCAGTAGCCGGTTCGCCGGCATGGCACACTAGAGAG
        TCCGGTGCTACACCGGACAGTCCAGTGAATTATAGCAGAGTGGCTCTCCAAAAACCTGAA
        GCTGAGCAGTTCAGAGTTGATCTCCCTtGGTGCACcGGACACTGTCCGGTGGTGCACCGA
        ATAGTCCGGTGCgCcAGaCCAgGGCaAcCTTCGGTTTCTTTTGCTCCTTTCTTTTTGAaC
        CCTATCTTGGACTTTTTATTAGTTTGTGTTTGAaCCTTTGGCACCTGTAGAACTTATAAT
        CTAGAGCAAACTAGTTAGTCCAATTATTTGTGTTGGgCAATTCAaCCACCAAAATCATTT
        AGgAAAAGGTTTTACCCTATTTCCCTTTCATGCATGTCTTGGTAATTCATTAAAAATGGG
        ACCGGTCGCCGTTGGCCCGGACGGTCCGTGCTCATGTGCGGATGGTCCGGACATGCGTAG
        ATCGACGAATTTATCACCGATGTGTGGAGGAGGTTGCGGTTGCCCAGGGCATGTGTCTAT
        CGACATCCCATAAAGGGGtTATAACTGGTCGTGACAACCTGTAGCCGATGAATTACACGT
        GTTTTCCCCAAATTCATCCTCGCGCGAAGgAAAATTTGCACCAGTAGATTTATCAAACGC
        ACGGTACTAGCCTCCTATAATCATTTTGCATACCCCCTATGATATTTTGCATTTGTTCTC
        GCTGTTCATCTACATAATTTTTAAGAGATTATAGCTCGTTTGTGTTGCTTACCTCGGGGT
        AGTCGTGGTAGGTCGAAGTGAAGTCAGATCCGTCGCCAGTTGTCGAACGACCTTGTTATT
        CCTGTCCACCTTGAAGTTGGCCAGGAACTTTGCCTTTGCCTCTTGGATCAGCTGCTCCTT
        GTGCTCCTCGAACTGGAGCTGTTCGTCAGCCGGCAAGGTTTCCCAAGTCGACTCTATGAT
        ATTGCTGGGGGAGACTTCAGAACTATCCCTTGAATCGGCCATTGAGGGCCGATTTGATGG
        GTCTATATGTGTCGTCCCCAGCGGGGTCACCAAAAAGTGTGTTGGCGCTTTTTCTGAGTG
        CCAATCACTGCGTGAGAACCGGCGGCGGTGCTCACTGCACAGGCGCGGACGGTCCGCGGC
        CAGGGGCCGGACGGTCTGCGACCTGGCGCAGGGCTTAGGGTTTCCTGCTTGACGGTCGGA
        CGATCCGCGCCTATAGGCCGGACGGTCCGCACGTGTGCAGGGCAGCGAAGGTCGTCGGTG
        GCGCCTGGATCTCGCTCCCGGGAGGGACCCCGTCGGGGAGGAGAGATCCTAGGTGTTGTC
        TAGGCTTGGCAGGCCGACCTAGACTCCTATCGATATAGAGTCGAAGAGAAGCGGAGAATT
        TGGGGATTGGAAGGCTAAACTAGAACTACTCCTAAATATACAAGAAATAAATACGAGATA
        AACTGGTATTGATTCGATTGATGGTGTTTAATCGGTCgGTATTCCTCTGTATTTATAGAG
```

```
        GAGGGGGGCTGGACCCGTTACAAACATATTTTTCGAGCTAATTCTGTGAATCTAGCCAAC
        AACTATAGCAAGAAACTCAGAATTCTAACTGGTTCTGCGCGCGCAAACCGTCCGACCCTT
        TATTTGATGCTCACCAGAACCCTGTCTTACCAATGTATTGAAATTGATTTATGGTGAAGA
        TAGAAATCTAAAGTTGTAATTCTTAAAATAAAGCACTCCTAAACAGAACATTAATCTTTA
        TAATTGATATTCGAATATTTTAGTCCGGTAACCGAACGCCCAGTAATTTGAAGAATATGG
        TCTTTATTTGTAGCCGGCGGCATCTCGAAAAGCTAGATCTAACTCCGAAAACAAACACGC
        AAAATCTACCGGAAAAATCTCCCCAGCAAAnAAGCAGACGTGGGGCCTGTCATGTACGCA
        CGGGGCATGACTCGTGCAAGAGCAACAAGTGCTGTCGTTGCAGAGGATCCGAGCTAAAAC
        AATCCAGCACACGGCCACTCTCGTCTCTCCTCTTCCTCCTCTTCCCTCCGTACCGCCCGC
        CCTTCACCTCCTCGACTCTTCTCCACCGCCCCCTCCCTCCTCGCGTCTCTCTCCCCCTTT
2881    GTCCCTGTAGAGCTCCAAATAAAACCCCACCGGTTTTCTTTTCCCGCGGCAATTACCATA

                  Start site predicted by RACE (GGCCC . )          DJ204 primer >>
2941    AAAAGAGCTCCCAATCTCTCTCCTCTCTCCGGCCCCTCTCTCTCGTTTCTGGCAGCAGTG
a       K  R  A  P  N  L  S  P  L  S  G  P  S  L  S  F  L  A  A  V  -
b        K  E  L  P  I  S  L  L  S  P  A  P  L  S  R  F  W  Q  Q  W  -
c         K  S  S  Q  S  L  S  S  L  R  P  L  S  L  V  S  G  S  S  G -

3001    GTGCGGTACTACCACCGCTCTCACTCCACACACACACACCGAGTATACGGTTAAGCAGGA
a       V  R  Y  Y  H  R  S  H  S  T  H  T  H  R  V  Y  G  *  A  G  -
b        C  G  T  T  T  A  L  T  P  H  T  H  T  E  Y  T  V  K  Q  E  -
c         A  V  L  P  P  L  S  L  H  T  H  T  P  S  I  R  L  S  R  R -

3061    GAGAGAACGGGAGAGTGAGACTGAGACCGGGTCCCAAGTACAATTCTCGCCTGGTTCTTG
a       E  R  T  G  E  *  D  *  D  R  V  P  S  T  I  L  A  W  F  L  -
b        R  E  R  E  S  E  T  E  T  G  S  Q  V  Q  F  S  P  G  S  *  -
c         E  N  G  R  V  R  L  R  P  G  P  K  Y  N  S  R  L  V  L  D -
                  I
3121    ATCGAAGGCATGATCAAGAATCACCAGAAAGTTCTAGATCTTTAGACgGCAGTCTTCTTT
        I  E  G  M  I  K  N  H  Q  K  V  L  D  L  *  T  A  V  F  F  -

                                        5' end of cloned cDNA (GGGTGG ..)
                                  II                                  III
                                                  DJ101 primer >>
3181    GGACTCCTCGGTTTCTTTTGTTCTGAGCTCTAGCCATGGAGGGTGGGGGAGGGAGTGGAA
          T  P  R  F  L  L  F  *  A  L  A  M  E  G  G  G  G  S  G  M -

                                                          << DJ203 primer
3241    TGGGTGGGTTCCCGGAAGCCACGGGTAACCTTCTCGATGCCGCAGCTCAGGAGTTCCACC
          G  G  F  P  E  A  T  G  N  L  L  D  A  A  A  Q  E  F  H  P -
<< DJ203 primer (continued)                    << DJ201 primer
3301    CTACGGTCTGTGCCCCCTATCCTCTACAGCCGCTTCCGCAACAGCTATACTGCCCCCACC
          T  V  C  A  P  Y  P  L  Q  P  L  P  Q  Q  L  Y  C  P  H  P -
                                                << RACE-RT primer
3361    CATATCCAGCCATGCCGGTGCCTCCGCCGCCGCAAATAGCCATGTTACAGCCAGTGCCTC
          Y  P  A  M  P  V  P  P  P  P  Q  I  A  M  L  Q  P  V  P  P -

3421    CGATGGCGATGGCCATGGCGCCGCAGCCGGGGTACACCTTGCCAACGACGACGCCGGTGG
          M  A  M  A  M  A  P  Q  P  G  Y  T  L  P  T  T  T  P  V  V -

3481    TCAATGGCCCGTCGAGCCGCGTCGTGGTGCTGGGCCTTGTCCCGCCGCACGCGCAGGAGG
          N  G  P  S  S  R  V  V  V  L  G  L  V  P  P  H  A  Q  E  A -

3541    CCGACGTGGCGCAGGCGATGGCGCCATTCGGCGCGATCCGCTCGGTCGACGCGTGCGCGG
          D  V  A  Q  A  M  A  P  F  G  A  I  R  S  V  D  A  C  A  V -

3601    TGGCGTCCGAGGGCGTGGCCACCGTCCATTTCTTCGACATCCGCGCCGCCGAGCTCGCCT
          A  S  E  G  V  A  T  V  H  F  F  D  I  R  A  A  E  L  A  L -
```

```
3661  TGACCTGTGTCCGCGAGCAGCACATGCGCCAGCAgAGCCGCCTCGGGCAGCTcTACGCGG
        T  C  V  R  E  Q  H  M  R  Q  Q  S  R  L  G  Q  L  Y  A  A -

3721  CGGCCGCCGTAGCCCCGGCGTGGGCTCCTGCACCGACGCCCCAGGCCTGGGACTGGCCCC
        A  A  V  A  P  A  W  A  P  A  P  T  P  Q  A  W  D  W  P  H -

3781  ACCCCAACGACGACGGCCGCGGCCTCGTCCTCGGGCACGCCGTGTGGGCCCACTTCGCCA
        P  N  D  D  G  R  G  L  V  L  G  H  A  V  W  A  H  F  A  T -

3841  CCGGCGCCGACGACGGCGACAACCGCGGCTCCCTGGTGGTCCTGAGCCCCCTGCCCGGCG
        G  A  D  D  G  D  N  R  G  S  L  V  V  L  S  P  L  P  G  V -

3901  TCTCGGTCGCTGACCTCCGCCAAGTCTTCCAGGCCTTCGGTACGCGCCACCGACCGAGCC
        S  V  A  D  L  R  Q  V  F  Q  A  F  G                    -

                                                    ▽ te1-mum3
3961  GACCAACCAGGCATTTCGTTTTCCCACGCTCGTTTCTTGTGTTTCCCGCAGCAATTGTAT

4021  TGGcCCGTCCTCGTGTTCGCAGGGGACTTGAAGGATGTGAGGGAGTCGGCGCAGCGGCCC
                              D  L  K  D  V  R  E  S  A  Q  R  P  -

4081  AGCCACAAGTTCGTGGACTTCTTCGACACGCGCGACGCCGCGCGCGCGCTCGCCGAGCTC
      S  H  K  F  V  D  F  F  D  T  R  D  A  A  R  A  L  A  E  L  -

4141  AACGGCCAGGAGCTTTTCGGCCGCCGCCTCGTCGTCGAGTTCACGCGCCCTTCCGGCCCC
      N  G  Q  E  L  F  G  R  R  L  V  V  E  F  T  R  P  S  G  P  -

4201  GGGCCCCGCAGGTAAACAACAACGCAATTTCAGTCAGCTAGCCTTCCCATCACCCTGTGG
      G  P  R  R                                                  -

4261  GCGGCTGAATTTTTGCCGTGTGCCTGTCGGTCCTAGGCGCGGGTACGCACCCCACCAGCA
                                          R  G  Y  A  P  H  Q  H  -

4321  CCGGCCCACCGCGCCGACTCCGCCGAGGCTTCAAGCGACGTGGCGACCGTCCCAACCGAC
       R  P  T  A  P  T  P  P  R  L  Q  A  T  W  R  P  S  Q  P  T -

                  ▽ te1-mum2
4381  GTCGTCTCAGCCGCCGGCATCCTCGTCGTCGTCCGGTTCCGTAAGGGCGAGGGAAGGAGT
       S  S  Q  P  P  A  S  S  S  S  S  G  S  V  R  A  R  E  G  V -

4441  GGTGCTTCTGAGGAGGAGCTCCTGTAAGTCTAGCGCGGGCAGCGACCAGTCGTCCAAGGG
       V  L  L  R  R  S  S  C  K  S  S  A  G  S  D  Q  S  S  K  G -

4501  AGGCAATGCCGGAACGAGCCATGAGCGCAAGACCAAGGGCGGCAAGATCGTGGTGGCGGC
       G  N  A  G  T  S  H  E  R  K  T  K  G  G  K  I  V  V  A  A -

4561  GGCGGCGGCATCCTCGTCGACCCCGACAGCGTCCGGGAAGCAAACCCAGAAAGGCGTCGG
       A  A  A  S  S  S  T  P  T  A  S  G  K  Q  T  Q  K  G  V  G -

                                          ▽ te1-mum1
4621  GAGCAGCGGCGGCGGGAGCTGGAAAGGACGAAAGAGCGGGTGGGAGGCGCGCTTCCTGTT
       S  S  G  G  G  S  W  K  G  R  K  S  G  W  E  A  R  F  L  F -
      << TE15 primer
4681  CAAGGAGCCCGAGCCGGCGGCGGCGCCGACACGCAAGCAACGCCGGCTTCGGAGATGGA
       K  E  P  E  A  G  G  G  A  D  T  Q  A  T  P  A  S  E  M  D -

4741  TACGAGGACCACCGTCATGATCAGGAACATACCGAACAAGTACAGGGTATCCCTGCCAAT
```

```
          T  R  T  T  V  M  I  R  N  I  P  N  K  Y                    -

4801      CTCCACGTTTGTTGCAACAATGCTATATTTCATGGGCGACGCACAGAATTGCGATTCTGC

4861      TAGCTGATGCCCGAGGCGTTTGGGTCGCAGCCAGAAGCTGCTGCTCAACATGCTGGACAA
                                        S  Q  K  L  L  L  N  M  L  D  N  -

4921      CCACTGCATCCAATCCAACGAGTGGATCGTGGCGAGCGGCGAGGAGCAGCCCTTCTCCGC
           H  C  I  Q  S  N  E  W  I  V  A  S  G  E  E  Q  P  F  S  A  -

4981      CTACGATTTCGTCTACCTCCCCATAGATTTCAAGTGAGTTGCGCGCTCGGCTTCCTTCTC
           Y  D  F  V  Y  L  P  I  D  F                                -

5041      ACATAATTGCATTTAGCTGTTCGCATTGACTTATATTTTATGGTGCCCAATGCAAGAACA

5101      CACACAAGCTTCCCAGCCTGAATTCGGGTTGTTTTCCTGGGAATTTGTTGCCTTGAATAA

5161      TGGTTGGAGTGGGAAATGTGGGCGGGGGGGTGGGAGGCGAGGAGCTGGCTGTGGAAGCGT

5221      GAGCTTCAGGATCGAGGTATCTTATCTTTGGGTGGGTAGAGTACAGTGCAGCAGGCATAG

5281      ATGGGCTATGGCGTTACTAGCCCAAATGGAAATAAATTTGGCAGCGGCAAGCAAGACCAT

5341      GCCAGATCCTTCTCGTTTTGATGCCCCCCCATGGATTCTTTGGTTGCATGCGCGCATCTG

5401      CTGGCTGCTGCTGCCGCTGCCGATGGCGAGGGAGCCGGCCTGCCTTCCTTTCTCGCGGCT

5461      GCTGCCGCCCCTGGGATCGCCTCCCTGCCTGCCTTTCTTGCTTTGCACAGCCGTGCACTT

5521      CCATAACGAGGCCAACACATACAAAAGGTGCGGCACCGATGGGGTTGTAGTGTTGAGACT

5581      AGACAGTAGAGATAGACGCTTAGCTATGGAGATGGAGCTGTGAAAAAAAAAGAGTTGCAT

6641      GAGTATGAAAAGGGGGAAAAAGTCAGCACTTTCGTTCTAAAGAGAAAATTTGAACTTCAG

5701      ATCAAAGTAGTTTTCTATATGCAAACAAAATGTATGCTTTCAAGTTTGCTTATTTAAAGC

5761      ACTTGCTTTTATATATACCGCCGGCAGCTTTTCTCTTGCACGAAGCAAGTTTGAGAAAAG

5821      ACCTATCAATGCCCAATGACGAGAGGACATTAGTGGCTGTCAAAAGGCATGCAGCTTCAA

5881      ACATACTCATGTGTAAAGCAGCTGCAGCATAATGCCAAAGTTTTTTTTTGATTGCAAGTC

5941      TTTTTAGTTATTGTTCCTCCATGCCATTACCATTCTGAATTTCTTGCCCTTGCATGAAAT

6001      GAGAATTTTTCTTCCTTATGCAAAAATGAAAAAAAACAAAAAAATGCAGCAACAAGTGTA
                                                      N  K  C  N  -

6061      ATGTGGGCTACGGCTTCGTCAACCTGACATCGCCGGAGGCTCGCGTGCGGCTGTACAAGG
            V  G  Y  G  F  V  N  L  T  S  P  E  A  R  V  R  L  Y  K  A  -

6121      CGTTCCACAAGCAGCCATGGGAGGTGTACAACTCGCGCAAGATCTGCCAAGTGACATACG
            F  H  K  Q  P  W  E  V  Y  N  S  R  K  I  C  Q  V  T  Y  A  -

6181      CGCGCGTACAAGTACGTACTCCACTCCyGCAGCTGCCCTTCTCCTACATCTACATCCGTC
            R  V  Q                                                    -

6241      GAAGACATTTTAATGCTACGTACGTACTTATTGCTCACATGTCAATGTTACAATTGTGCG
```

```
6301    GCGGCGCAGGGCCTGGAAGCGCTGAAGGAGCACTTCAAGAACTCCAAGTTCCCGTGCGAC
                G   L   E   A   L   K   E   H   F   K   N   S   K   F   P   C   D   -

6361    AGCGACGAGTACCTGCCCGTGGCGTTCTCGCCGGCGCGCGACGGCAAGGAGCTTACGGAT
        S   D   E   Y   L   P   V   A   F   S   P   A   R   D   G   K   E   L   T   D   -

6421    CCAGTGCCCATCGTGGGCCGCTCGCCCGCGGCGTCGTCCGCGTCGTCGCCTCCCAAGAGC
        P   V   P   I   V   G   R   S   P   A   A   S   S   A   S   S   P   P   K   S   -

6481    CGGGCGGCTAGCGTGGACCGGCTTGGGCAGGAGCTGATGCCGGCGCCGTCGTCATCCGCG
        R   A   A   S   V   D   R   L   G   Q   E   L   M   P   A   P   S   S   S   A   -

6541    GACGGCGCGTCGTCGACCACTACGTCCACCCACGCGCCGTCCGAACACGACGAGGAGGAG
        D   G   A   S   S   T   T   T   S   T   H   A   P   S   E   H   D   E   E   E   -

6601    GAGGAGGGAGACATCAGGCTCGCAGGCGAGCTGCGGCGGCTTGGCTACGACGACTAGCTG
        E   E   G   D   I   R   L   A   G   E   L   R   R   L   G   Y   D   D   *       -

6661    GCTCCGATCCTAGCTGCAGCTAGGCGGGCGACGCCTAGCTCAGCAACGATCGGCATGGCG

6721    TGGCGTCGCTTCGAGAAGCGCGGAGGCAGTAGTGACTTCGGCCAAGCTTTTCACAGATCC

6781    GTAGCGCTCGCTCCTGCACGCGCAGGCGAAGCAAGTGAAGTGGACTGGGACTTGACTTTG

6841    GAGAAGCAAAAGTGAGCTGATGTTCGTTGGTTCGTAGACGCTGCGCACAGCAGGTGTGGA

6901    GTCGTGTTCTTTTGGATTTTGGTTTGAAGTTTCCTACTACTTGTTTCTTTGAAACTTGGA

6961    GCTATAGCTAGCTTAGGCTTAGCTAGGACTGCTGGGAGGGGAAAGGGCATGCAGGGGACT

7021    ATCAATTCGTGGTTCGTCATGCGCTCGTTCTTTACATTAAAATAAGATTGTTATTAGTAC

7081    AGCTTTACAATACTTCTGCATATTCGGAGAGAGACTCTGATCTCATATCATCACGAGTTG

7141    CTAGATTTCggtCgGATCTCTTCsATTGATGCGAACCAACAGATCTTTCTCTATCGCTGC
        AcGGCTCCACCGCGCTATCTTTGCAGACTCACACTGTCCTCTCCTTTGTAGACTCTAGTC
        TTCTCATCCTTCTATTTTTCTCAATTTACATTCGATTTGTTCAATTTACATTTnGTTATC
        CCATCGATCTGTGCTCACTCGTGCTAAAAAACAAAGATGTAAGCGTTGCGTAATATTnAT
        ACAGTACAAGCATATATCTAGAAAATTCAAATTATCTTGTGATTTGAAATGGAAAACTGT
        CTCACTAATAAAAAGATATTTGTATCAAatAtATATtTtaTaTtTtTttkkATTTatTtc
        CtaAcataATAgAtataaTAggwwcTatatagTtTaTATcTATttAtcaaaAcTtattgT
        cgTTTAtttAaAAhcaATATaTatttATccATATATaTaTTtttTcCATGcatggcAtcT
        CTATATATTTAAATATGtAATTATTTTGAAATCTTACCCTTAATCATGGTGCCTCTTCTT
        GCATCATGTGTCCTCAGTCATGAATGAATCTATTTTCTCTCTCCTCAAGTTAAGTTGAGA
        CTCCTGAGTTTATTGTATCTTTCATTAACTCTAGTTTTTTGTATTTTCTTATAAATAAAT
        ATTTTAGGTTATAGAATAAGTTCTTAAGTACTACAATATCATAAAACCATTGTTTTAAAA
        ACATGTATCATATTATTCAAAATAAGTCATAGAGCTTGAAGTGAGATCTTAAATAGGGGG
        GTAATTCAAACCTCACGACAATTTTAATATTTGAAGCTTATAATGATCAATTTTTTACTC
7981    TAAAGTATTCTTAAATAGCCATAAAATATAGCCACGAGTAGTCTCTCATACTTCTCGAG
```

2.11. Prediction of transcription start site with 5' RACE

2.11a. The 5' RACE protocol was inconsistent

Two RACE experiments were carried out to completion starting with separate reverse transcription reactions. The gels of secondary PCR products are shown in Figure 2.4a. These experiments were inconsistent. The two ~ 400 bp products were produced from shoot meristems samples in the first experiment. They were cloned and sequenced ; the sequences were identical apart from the length of the polyT tail, and corresponded to the *te1* genomic sequence from 2971 bp (not shown). Three SAM products, and one embryo product from the second experiment were cloned and sequenced, but the sequence did not correspond to any region of the *te1* genomic sequence, or to each other when compared with the GCG program *gap*.

2.11b. RT-PCR did not confirm that the transcription start site was as predicted by 5' RACE

The predicted start site of the first experiment was unexpected since the 5' end of the transcript contained three ATG start codons (I, II and III in Figure 2.3), the first was followed in frame by a stop codon shortly afterwards. RT-PCR was used in an attempt to confirm that that the *te1* transcript extended as far 5' as predicted by RACE. A reverse transcription reaction was performed from both leaf and SAM polyA+ RNA samples prepared as described in Section 2.7f. This reaction was amplified in two separate PCR reactions with different primer sets (Figure 2.4b).

The first primer set (DJ101/RACE-RT) extended from within the previously cloned *te1* cDNA to187 bp downstream, within the first exon of the *te1* gene. PCR from these primers produced a product of the expected size from the SAM RT sample but not from the leaf sample (data not shown), indicating that the transcript included this region.

The second primer set (DJ204/TE15) extended from very close to the predicted 5' end of the transcript to beyond intron three. This primer set was designed to detect spliced mRNA products whose 5' ends were close to that predicted by 5' RACE. This primer set produced no products from the same quantity of the same RT reaction. A positive control included in the DJ204/TE15 reaction produced a product of the expected size from 2 ng of maize genomic DNA. While these results do not prove that the prediction of

transcription start site obtained from 5' RACE experiments is incorrect, they do not support the prediction.

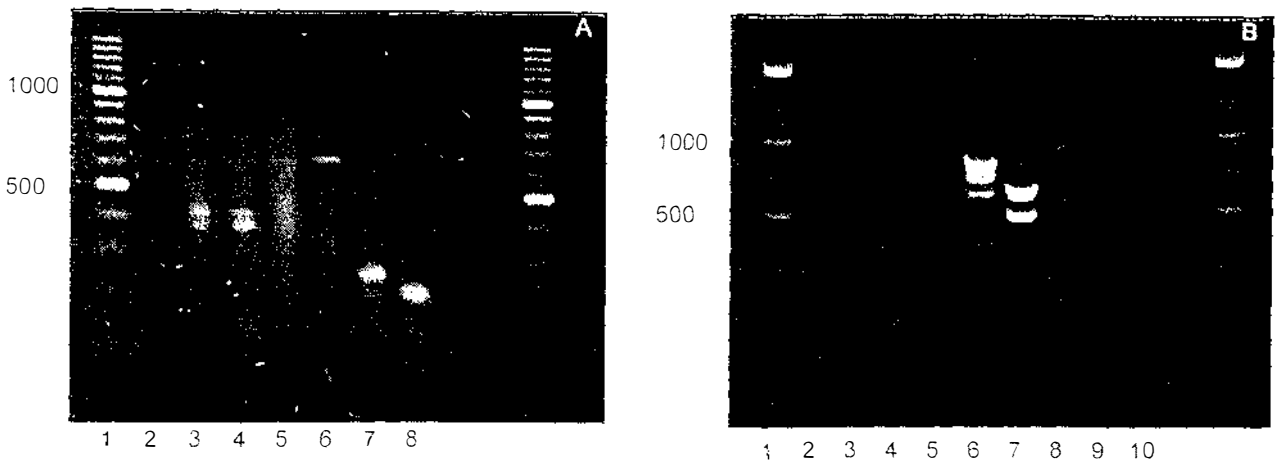


Figure 2.4a. 5' RACE gels. 5' RACE produced inconsistent results and was prone to contamination. Images A) and B) show the secondary nested PCR gels from two separate experiments. **A**) included two SAM samples. **1**) 100 bp ladder (Roche), **2**) blank, **3**) SAM1 5µl of primary PCR, **4**) SAM1 0.5 µl of primary PCR, **5**) SAM2, **6**) leaf, **7**) DNA positive control, **8**) synthetic transcript control. **B**) included SAM and embryo samples, RT reactions were purified both by precipitation (*ppt*) and Concert Kit cleanup (*kit*) **1**) 100 bp ladder, **2**) SAM (*ppt*), **3**) embryo (*ppt*), **4**) no RNA (*ppt*), **5**) leaf (*ppt*), **6**) SAM (*kit*), **7**) embryo (*kit*), **8**) DNA positive control, **9**) PCR negative, **10**) PCR positive.

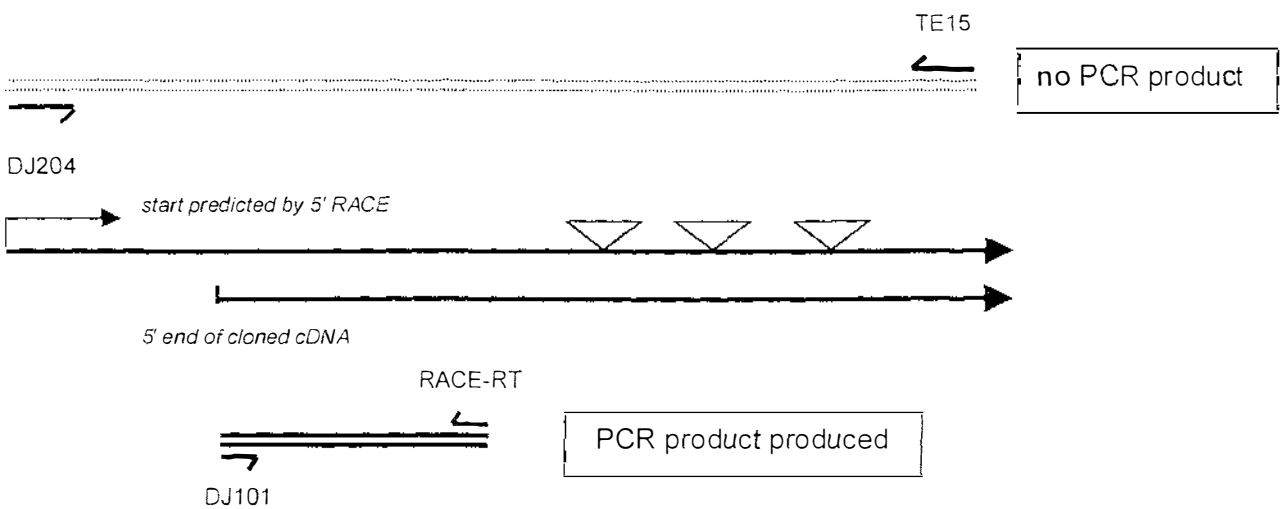


Figure 2.4b. RT-PCR used in attempts confirm the transcription start site predicted by 5' RACE. Both PCR reactions used the same RT products as a template. The DJ204/TE15 primer set produced no product indicating that there were no cDNAs in the RT reaction that spanned the regions from the predicted start site to beyond intron three (introns are represented by triangles). Some form of the *te1* cDNA was reverse transcribed, since the DJ101/RACE-RT primer set produced a product from SAM sample, but not from the leaf sample. Previous RT-PCR reactions indicated that SAM tissues contain the *te1* transcript, but leaf tissues do not (Section 2.12).

2.12. Quantification of *te1* transcript levels

RT-PCR results indicating the levels of the *te1* transcript in various wild type tissues are shown in Figure 2.5. Figure 2.6 shows an RNA gel stained with ethidium indicating that the samples used for RT-PCR were equally loaded. Competitive RT-PCR results were interpreted in the following way; when only one band was present in the RT-PCR gel this was taken as an indication that this sample (either altered synthetic or native transcript) was present in much larger quantities, when two bands were present the transcripts were present in approximately the same order of magnitude, and differed by the intensity of the band. All tissues except leaf contained some detectable level of *te1* transcript. The estimated levels of transcript were; embryo > 100fg/μg total RNA, SAM > 100fg/μg, early tassel <100fg/μg >10fg/μg, late tassel ~1fg/μg, ear ~100fg/μg, leaf << 1fg/μg, root ~100fg/μg. The lower level of detection for this assay was at least 1 fg/μg. RNA was extracted only once from each tissue sample, and three consecutive reverse transcriptions reactions were performed for each (with 100, 10 and 1 fg synthetic altered te1 transcript).

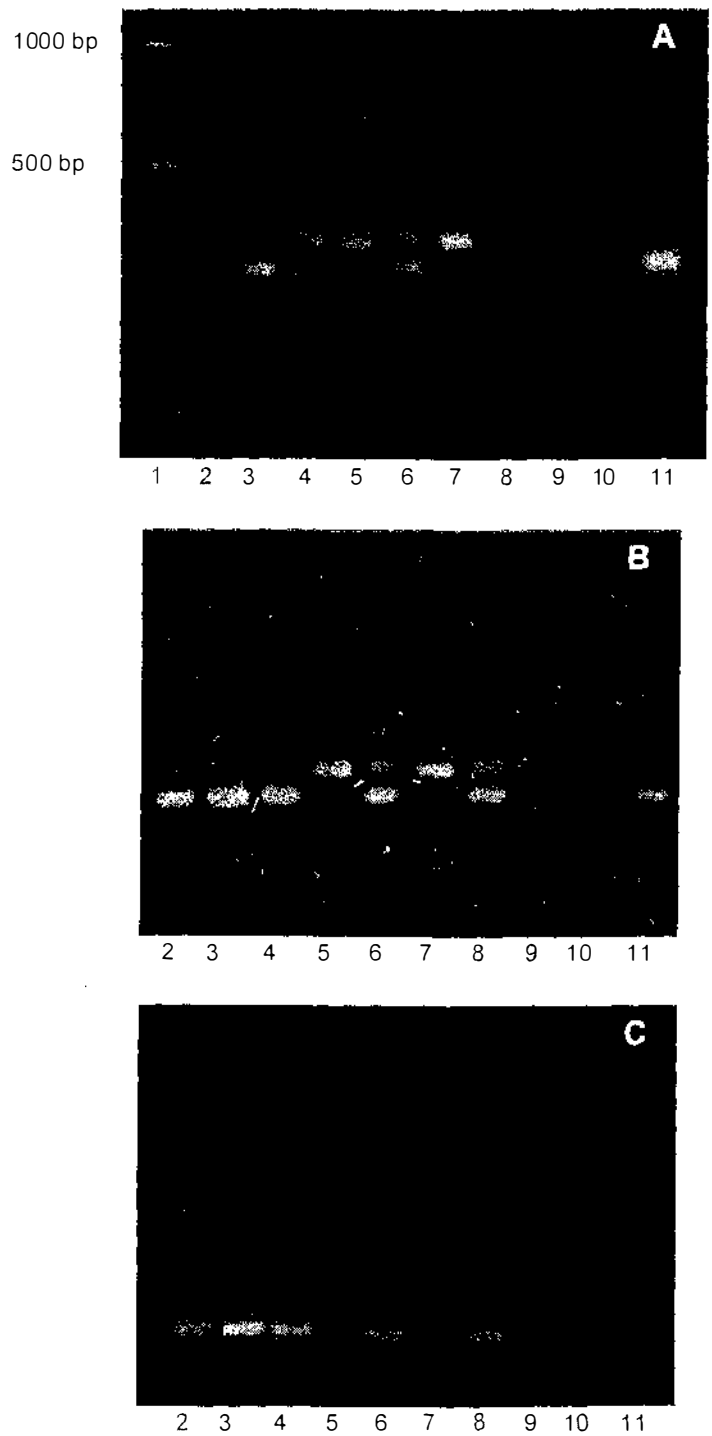


Figure 2.5. Competitive RT-PCR. A synthetic *te1* transcript lacking the *Bgl*II restriction site was added to each sample. PCR products derived from this transcript are distinguished from native transcripts by digestion with *Bgl*II, the synthetic transcript runs at the original size of 291 bp, the native transcripts are cut and run at 247 bp. **A)** spiked with 100 fg of altered synthetic *te1* transcript/μg total, **B)** spiked with 10fg of altered synthetic *te1* transcript/μg RNA Lanes **C)** spiked with 1fg of altered synthetic *te1* transcript/μg. Lanes are **1**) 100bp ladder (Roche), **2**) embryo, **3**) SAM, **4**) early tassel, **5**) late-tassel, **6**) ear, **7**) leaf, **8**) root, **9**) RT mix control, **10**) PCR negative control (H_2O), **11**) PCR positive control (1fg of pBV432 insert).

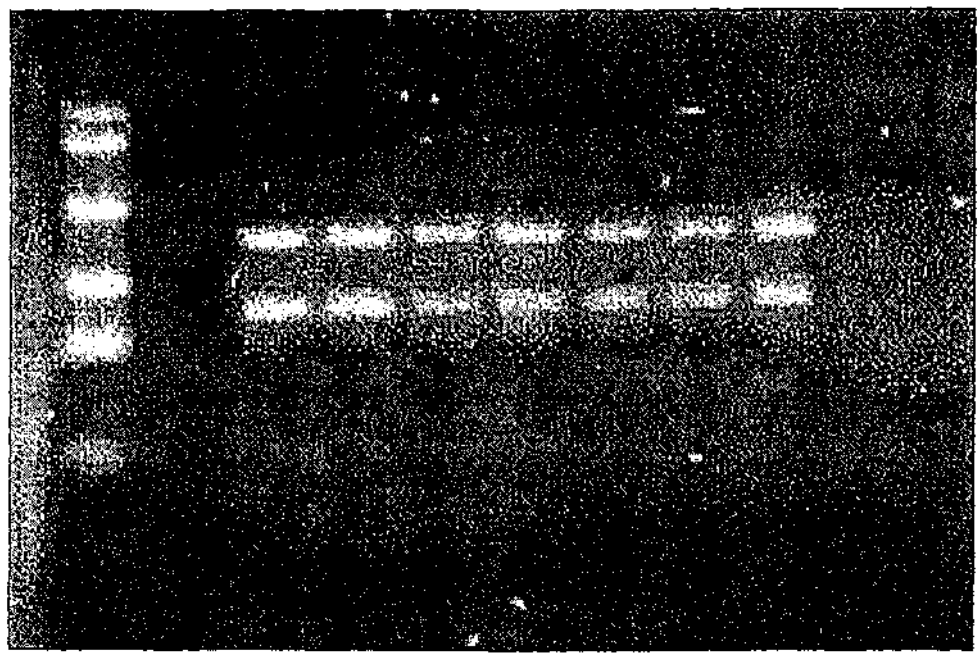


Figure 2.6. RNA samples used for competitive RT-PCR are equally loaded. RNA was glyoxylated and electrophoresed as described in section 2.8a. Lanes are **1**) 0.24-9.5 kb RNA ladder, 3μg total loading, **2**) no sample, **3**) 1μg embryo sample, **4**) 1μg SAM samples, **5**) 1μg early tassel, **6**) 1μg late tassel, **7**) 1μg ear, **8**) 1μg leaf, **9**) 1μg root. Assuming that each band of the RNA molecular weight ladder has equal mass, they should contain 0.5μg of RNA each. Only the rRNA of plant tissue samples are visible.

2.13 Analysis of *te1* mutant alleles

The mutations in the *te1-mum1* and *te1-mum2* alleles had been identified previously (Veit 1998). PCR and Southern blotting was used to characterise the remaining alleles by the following criteria.

1) Only 'reliable' primer sets are included in these summaries; those primer sets that consistently produced a single abundant product from B73 DNA.
2) When a PCR product was produced using two primers complementary to the *te1* sequence from a mutant allele DNA that was identical in size to the B73 normal inbred, the region bound by these primers was assumed to be normal. A larger product defined an insertion, a smaller, or no product indicated a deletion or rearrangement. This allowed the normal and abnormal regions of the alleles to be determined.
3) PCR products resulting from mutator element primer (MUSEL or MUEND) paired with a primer complementary to the *te1* sequence produced a product were used to infer the position of the insertion, and cloned.
4) Southern blot data was compared to PCR data, and hyptheses were formed about the nature of the mutations by a largely intuitive process.

The position of the *Mutator* insertion in the *te1-mum3* allele was determined by cloning and sequencing the PCR product. Positions of these insertions are; *te1-mum1*, 4653 bp (exon 3); *te1-mum2*, 4396 bp (exon 3); *te1-mum3*, 4017 bp (intron 1) (Figure 2.3).
The results of PCR analysis of the remaining uncharacterised alleles (*te1-1*, *te1-ems*, *te1-mum4*, *te1-mum5*, *te1-mum6*, *te1-mum7*), are summarised in Section 2.13a.

2.13a. Analysis of mutant alleles by PCR

Appendix 8 shows a minimum data set for the PCR analysis: where a mutation was defined by a particular PCR reaction, only that reaction is shown.

2.13b. Southern blots of *te1* alleles

The results of Southern analysis of *te1* alleles are shown in Figures 2.8, 2.9 and 2.10. Expected molecular weights were calculated from the genomic sequence, and the molecular weights of bands on blots were estimated using the molecular weight markers. This data is tabulated in Appendix 9. Primers used for synthesis of DIG labelled PCR

probes from a *te1* cDNA template (and corresponding genomic region covered) were; Southern 1, BV34/TE19 (4597-6345); Southern 2, SW1/TE1 (3327-4537), BV34/TE19 (4597-6345), SK/TEC10 from pBV412 (the 3' *Bam*HI fragment subclone of the *te1* cDNA vector pBV407) (6417-6963); Southern 3, SW1/TE1 (3327-4537).

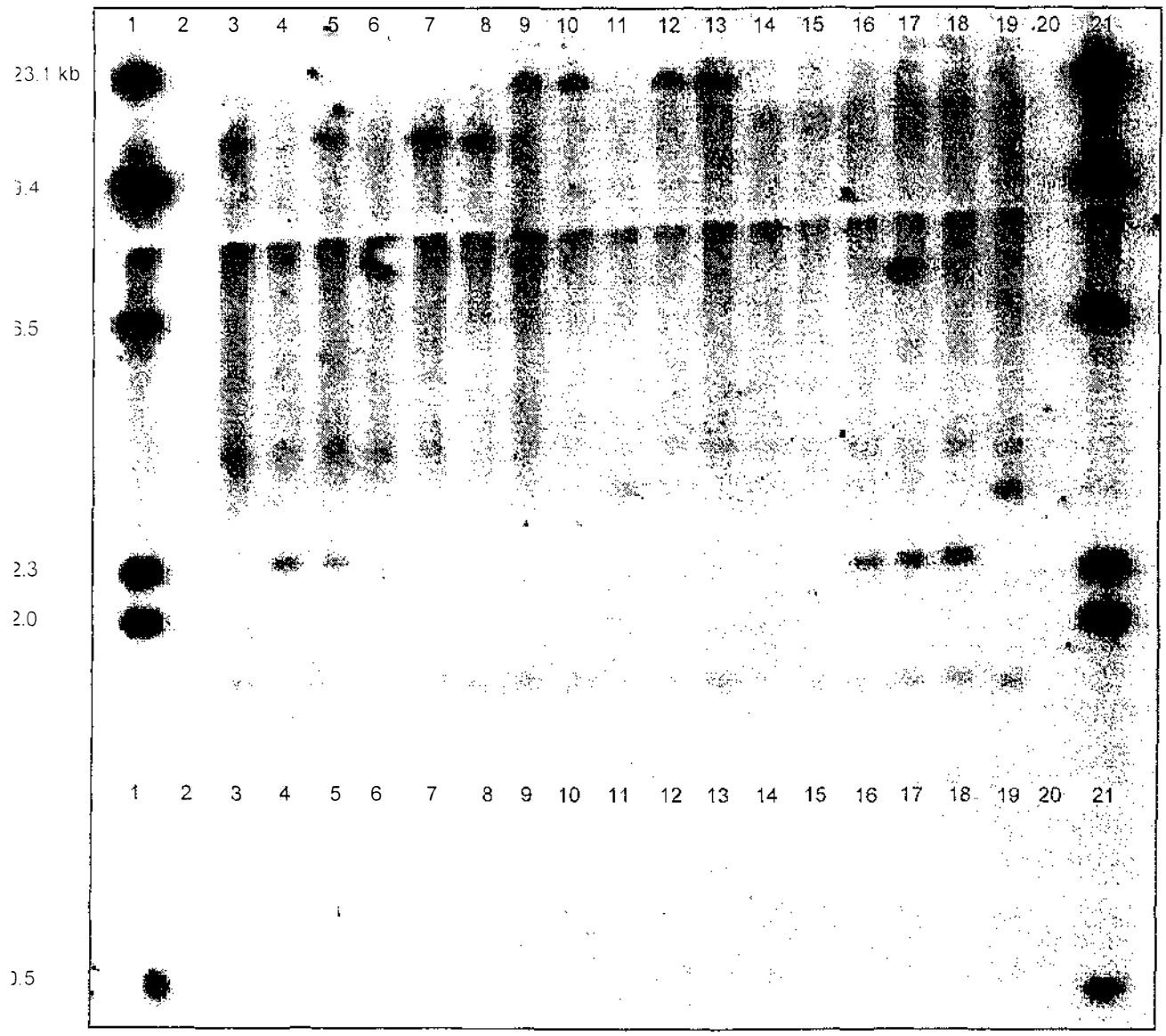


Figure 2.7. *te1* allele Southern blot 1. DNA was extracted from homozygous plants as described in Section 2.9b, cut with *Hind*III and subjected to Southern blotting (Section 2.9e). The membrane was probed with a DIG labelled PCR probe synthesised from the *te1* cDNA with the primers BV34 and TE19. This portion of the cDNA covers the region from 4597 to 6362 of the *te1* genomic sequence. Note that some contaminating DNA was present, probably in the gel loading buffer, that produces a band in all samples including the molecular weight markers (arrow). All alleles are in B73 inbred background unless stated otherwise. Lanes are described as, *te1* allele (family of origin); **1**) DIG labelled λ *Hind*III molecular weight marker, **2**) blank, **3**) normal (B73), 4) *te1-1* (A188) (DJ96:17). **5**) *te1-1* (DJ96:22), **6**) *te1-1* or *–mum7* (DJ96:16), **7**) *te1-ems* (A188) (DJ96:15), **8**) *te1-ems* (DJ96:26), **9**) *te1-mum1* (DJ96:23). **10**) *te1-mum1* (DJ97:18), **11**) *te1-mum1* or *2* (DJ97:51), **12**) *te1–mum2* (DJ96:25), **13**) *te1–mum2?* (DJ97:57), **14**) *te1–mum3* (DJ97:43), **15**) *te1–mum3* or *6* (DJ96:4), **16**) *te1–mum3* or *6* (DJ96:3), **17**) *te1–mum3* or *6* (DJ96:2), **18**) *te1–mum3* or *4* (DJ97:22), **19**) *te1-?* (W22, DJ96:40), **20**) blank, **21**) DIG labelled λ *Hind*III molecular weight marker.

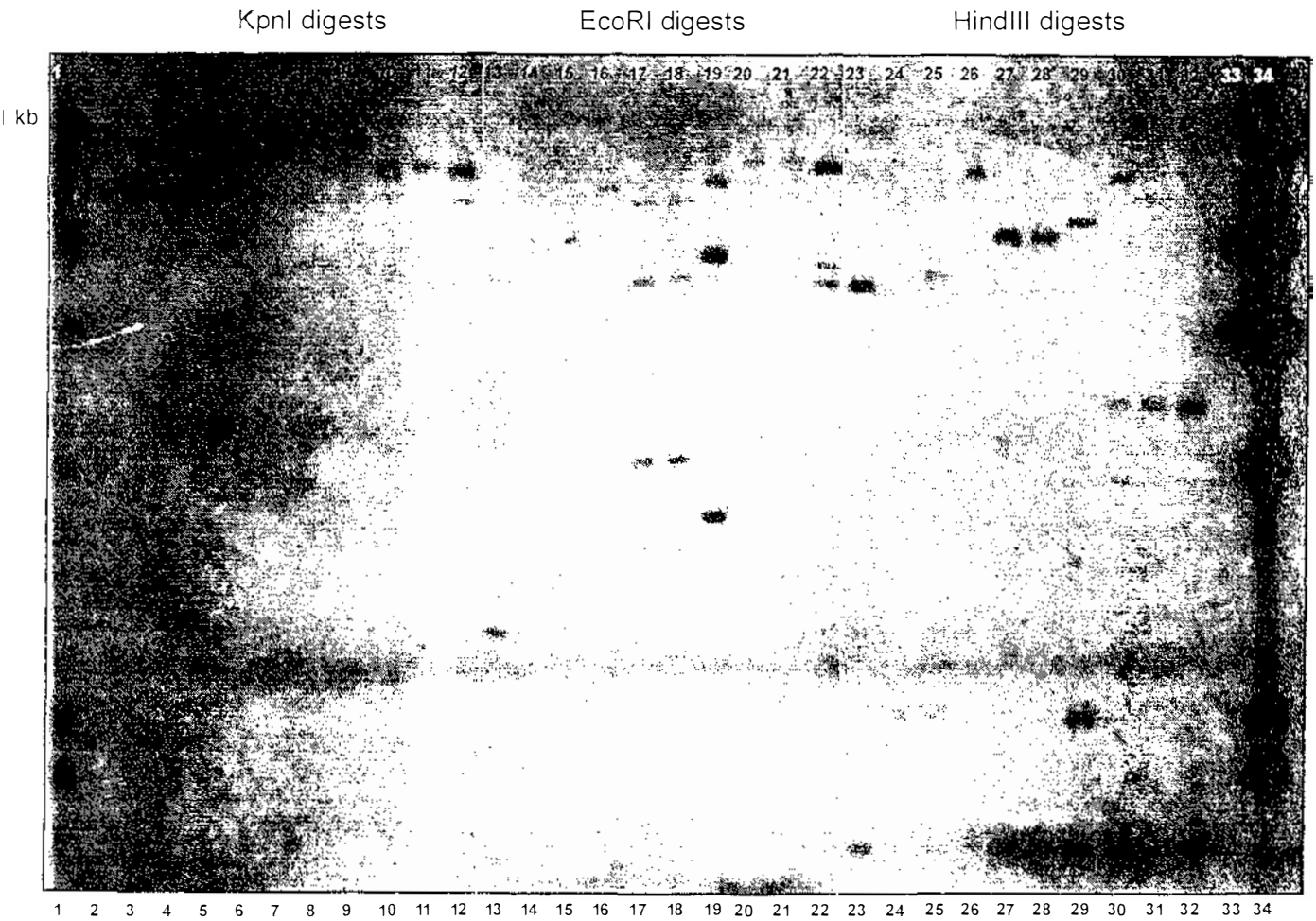


Figure 2.8. *te1* allele Southern blot 2. DNA was extracted from homozygous plants (Section 2.9b), cut with either HindIII, EcoRI, or KpnI and subjected to Southern blotting (Section 2.9e). The membrane was hybridised with three DIG labelled probes that included the entire *te1* cDNA. All alleles are in B73 inbred background unless stated otherwise. Lanes are; **1**) DIG labelled λ HindIII molecular weight marker (Roche), **2**) blank, **3-12**) **KpnI digests**; 3) normal (B73), **4**) *te1-1* (BV97:32), **5**) *te1-1* (A188) (DJ96:22), **6**) *te1-1* (NA97:32), **7**) *te1-mum1* (DJ96:23), **8**) *te1-mum2* (DJ96:25), **9**) *te1-mum3* (DJ96:2), **10**) *te1-mum3* or *5* (DJ98:17), **11**) *te1-1* or *–mum7*, **12**) *te1-mum7* (NA2331), **13-22**) **EcoRI digests**; **13**) normal (B73), **14**) *te1-1* (BV97:32), **15**) *te1-1* (A188) (DJ96:22), **16**) *te1-1* (NA97:32), **17**) *te1-mum1* (DJ96:23), **18**) *te1-mum2* (DJ96:25), **19**) *te1-mum3* (DJ96:2), **20**) *te1-mum3* or *5* (DJ98:17), **21**) *te1-1* or *–mum7*, **22**) *te1-mum7* (NA2331), **23-32**) **HindIII digests**; **23**) normal (B73), **24**) *te1-1* (BV97:32), **25**) *te1-1* (A188) (DJ96:22), **26**) *te1-1* (NA97:32), **27**) *te1-mum1* (DJ96:23), **28**) *te1-mum2* (DJ96:25), **29**) *te1-mum3* (DJ96:2), **30**) *te1-mum3* or *-mum5* (DJ98:17), **31**) *te1-1* or *–mum7*, **32**) *te1-mum7* (NA2331), **33**) blank, **34**) DIG labelled λ HindIII molecular weight marker.

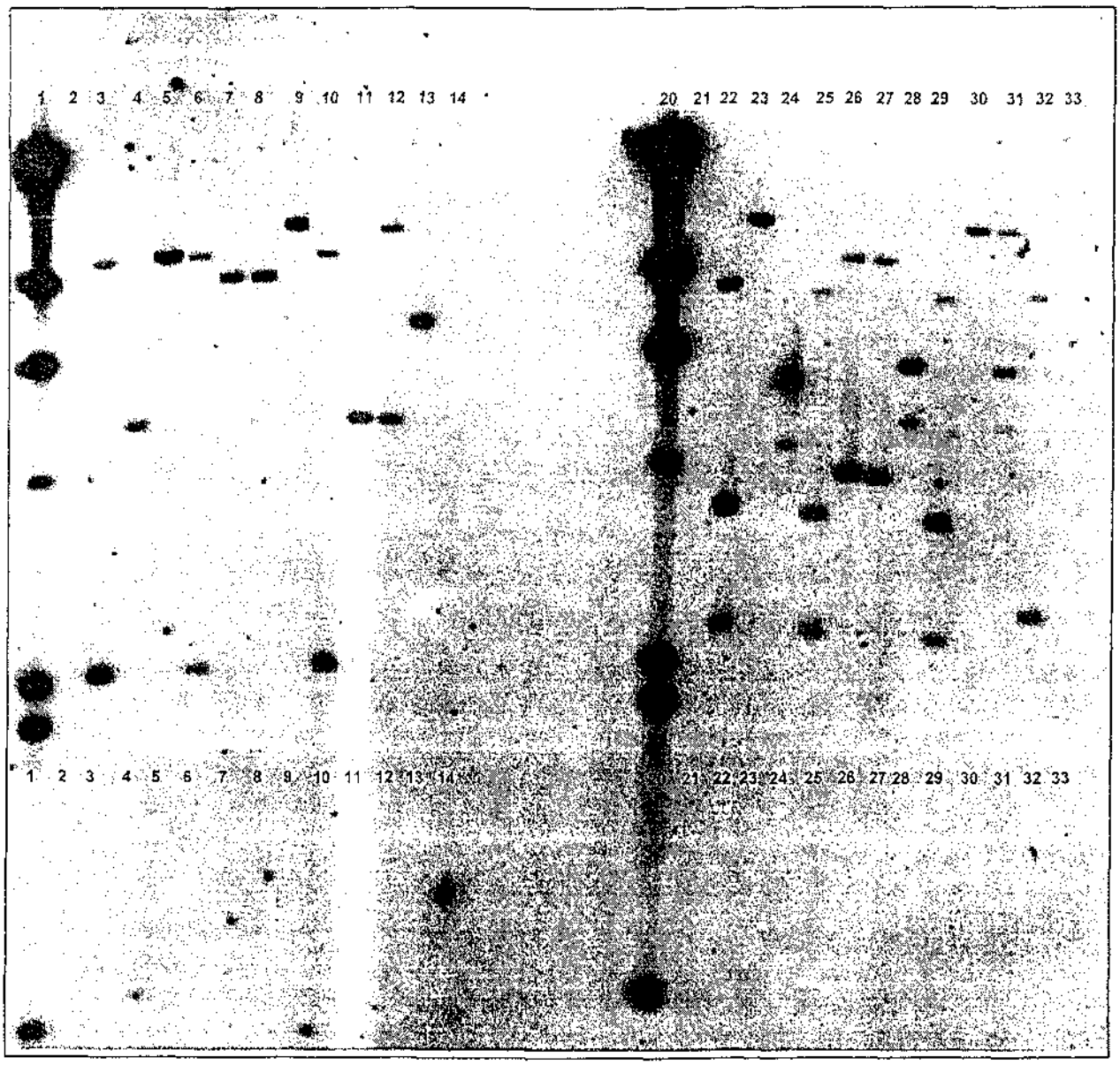


Figure 2.9. *te1* allele Southern blot 3. This Southern blot was produced by Carmel Gillman, methodology is essentially the same as for Southern blots one and two. The membrane was probed with a DIG labelled PCR probe made from the 5' region of the *te1* gene with primers SW1 and TE1. This region of the cDNA encompasses from 3328 to 4537 of the *te1* genomic sequence. Lanes are; **1**) DIG labelled λ HindIII molecular weight marker (Roche), **2-13**) **HindIII digests**; **2**) *te1-mum3* (BV99:5), **3**) *te1-mum6* (A188) (BV99:9), **4**) *te1-mum7* (BV99:13), **5**) A188 (BV99:16), **6**) *te1-?* (BV99:20), **7**) *te1-mum1* (BV99:21), **8**) *te1-mum2* (BV99:23), **9**) *te1-mum3* (BV99:25), **10**) *te1-mum6* (BV99:26), **11**) *te1-mum7* (BV99:29), **12**) *te1-mum3 or -mum5* (BV99:32), **13**) normal (B73), **14-19**) blank, **20**) DIG labelled λ HindIII molecular weight marker. **21-33**) **EcoRI digests**; **21**) *te1-mum3* (BV99:5), **22**) *te1-mum6* (A188) (BV99:9), **23**) *te1-mum7* (BV99:13), **24**) A188 (BV99:16), **25**) *te1-?* (BV99:20), **26**) *te1-mum1* (BV99:21), **27**) *te1-mum* 2 (BV99:23), **28**) *te1-mum3* (BV99:25), **29**) *te1-mum6* (BV99:26), **30**) *te1-mum7* (BV99:29), **31**) *te1-mum5* (BV99:32), **32**) normal (B73).

2.13c. Analysis of combined Southern and PCR data

te1-mum3* **allele**

The 500 bp product from BV allele *te1-mum4* from (line 2154-5) was cloned in to pGEMT-Easy, and sequenced; the *Mutator* elements 3' occurs after position 4017 in the *te1* genomic sequence. The position of this insertion is shown in Figure 2.3.

te1-mum5 **allele**

The *te1-mum5* lesion is a ~700 bp insert between 4170 and 4899 bp. The 2.4kb PCR product from *te1-mum5* was compared by SmaI and StuI restriction digestion to the B73 product from the same primer set (not shown). This analysis indicated that there is a ~700 bp insert between the SmaI site (4170) and SW4 primer (4899 bp). This product has not been cloned.

te1-1 **allele:** see Figure 2.10

te1-ems **allele:** see Figure 2.11

te1-mum4* **allele:** see Figure 2.12

te1-mum6 **allele:** see Figure 2.13

te1-mum7 **allele:** see Figure 2.14

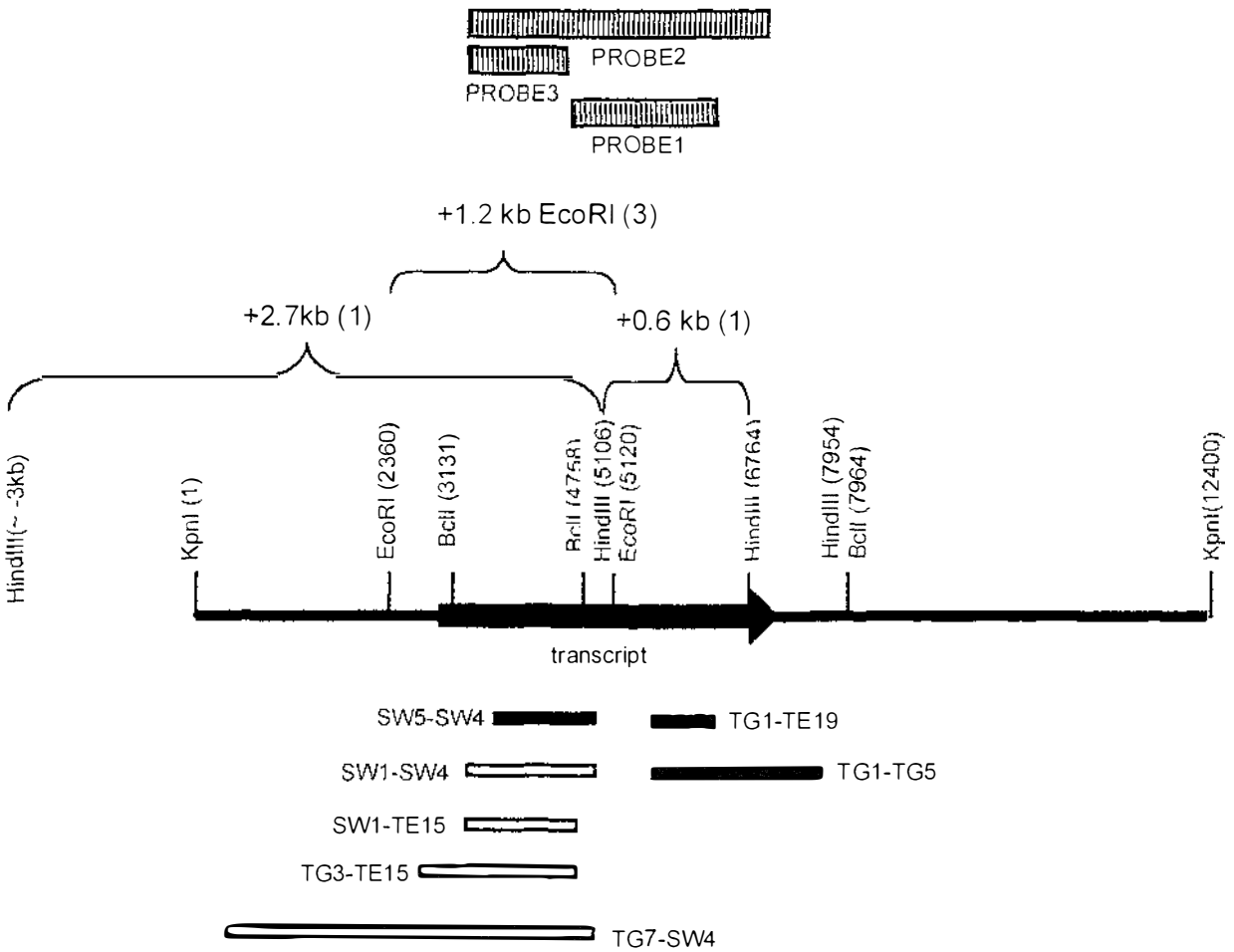


Figure 2.10. Summary of *te1-1* allele PCR and Southern analysis.

The 12.4 kb *te1* genomic clone is indicated as a thin line containing an arrow indicating the extent of the nascent *te1* transcript. The positions of probes for Southern blots (1,2,3) are indicated above as striped bars. The range covered by PCR primers that produced a product (filled line) or no product (open line) in experiments where B73 (wild type allele) produced a single abundant product are indicated below the gene line. Differences in sizes are bands from Southern blots are shown as brackets ({), the number in parenthesis indicates the Southern from which they derive. For the *te1-1* allele any primer set whose 5' primer is more 5' than SW5 (position 3653 bp into the *te1* genomic sequence) produced no product, indicating that this region is absent or not juxtaposed to the down stream region as in wild type. The mutation extends at least to 389 bp (TG7 site). Since *both* HindIII fragments in Southern 1 are larger than B73 the data are best interpreted as the result of a rearrangement, a large insertion that contains a HindIII site, or a deletion.

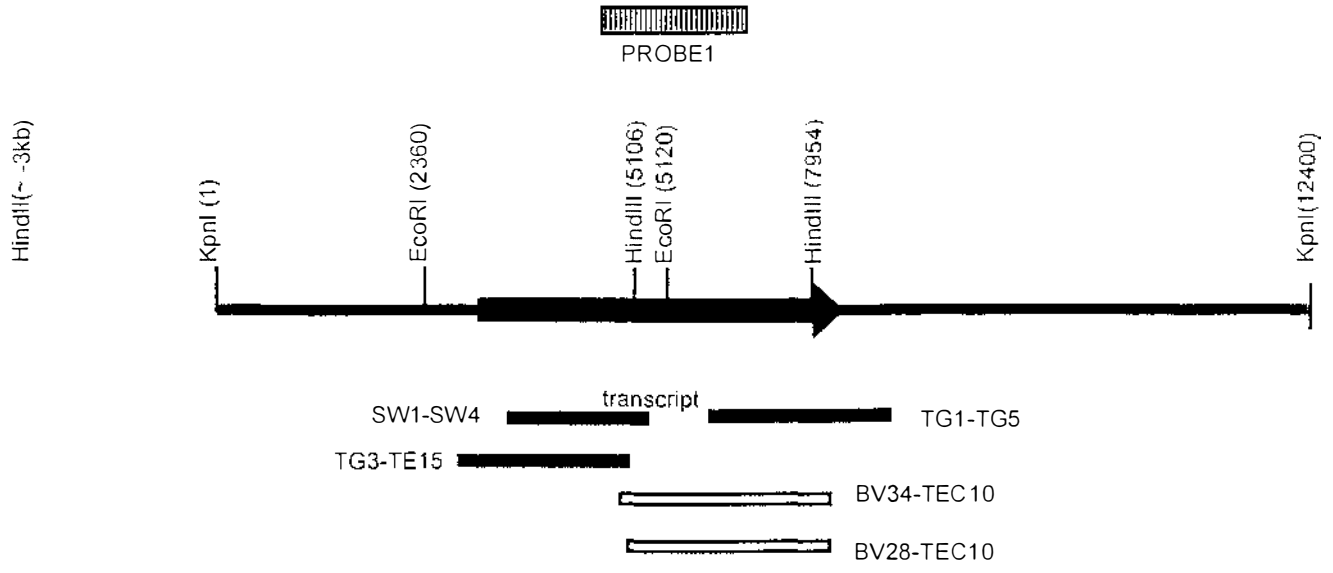


Figure 2.11. Summary of *te1-ems* allele PCR and Southern analysis. Symbols are as described for Figure 2.11: striped bars above gene line represent probes for Southern blots, open bars below line represent successful PCR, open bars below line represent unsuccessful PCR. The *te1-ems* allele was only examined in Southern 1 - band sizes were identical to B73. This allele was produced in an EMS mutagenesis screen for *liguless* (by whom?). Since EMS would be expected to induce point mutations the lack of PCR products from the BV34/TEC10 and BV29/TEC10 is surprising. Since the preparation of *te1-ems* allele genomic DNA produced inconsistent results from PCR and the lack of amplification from these primer sets is interpreted an artefact.

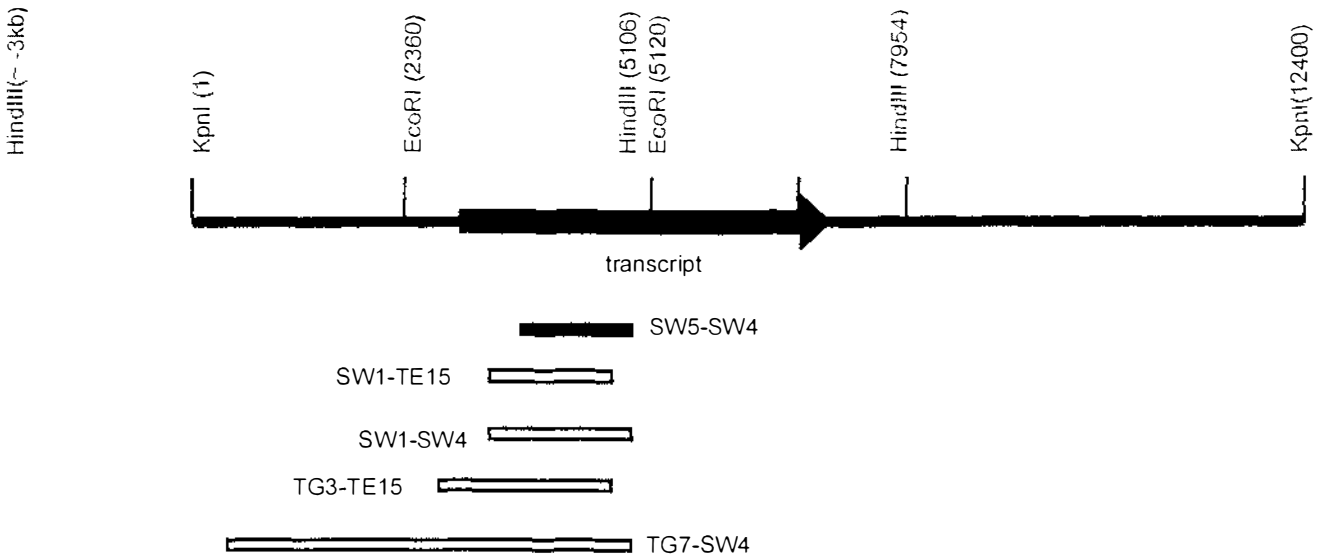


Figure 2.12. Summary of *te1-mum4* allele PCR and Southern analysis. Symbols are as described for Figure 2.11: open bars below line represent successful PCR, open bars below line represent unsuccessful PCR. The *te1-mum4* allele was not included in any of the Southern blots. PCR experiments indicate that the mutation extends from the TG7 primer to the SW1 primer (389 to 3327).

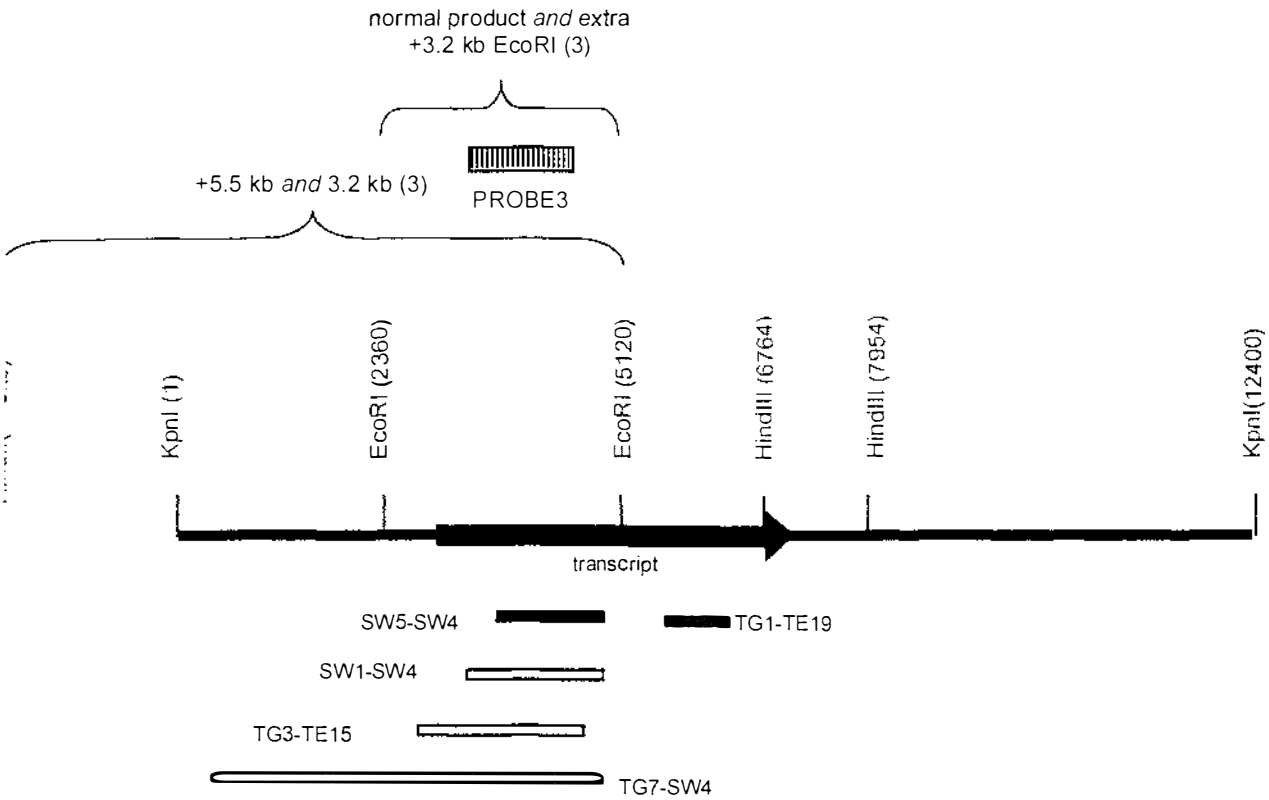


Figure 2.13. Summary of *te1-mum6* allele PCR and Southern analysis. Symbols are as described for Figure 2.11: striped bars above gene line represent probes for Southern blots, open bars below line represent successful PCR, open bars below line represent unsuccessful PCR. Only Southern 3 includes the *te1-mum6* allele. Because EcoRI digests produces both the normal sized and an extra 3.9 kb band, and HindIII digestion produces two aberrantly sized bands (Figure 2.10) this mutation does not appear to be a simple insertion/deletion. PCR experiments indicate that up to the primer SW5 (3653) is normal, but from SW1 to TG7 (389 to 3327) has been rearranged.

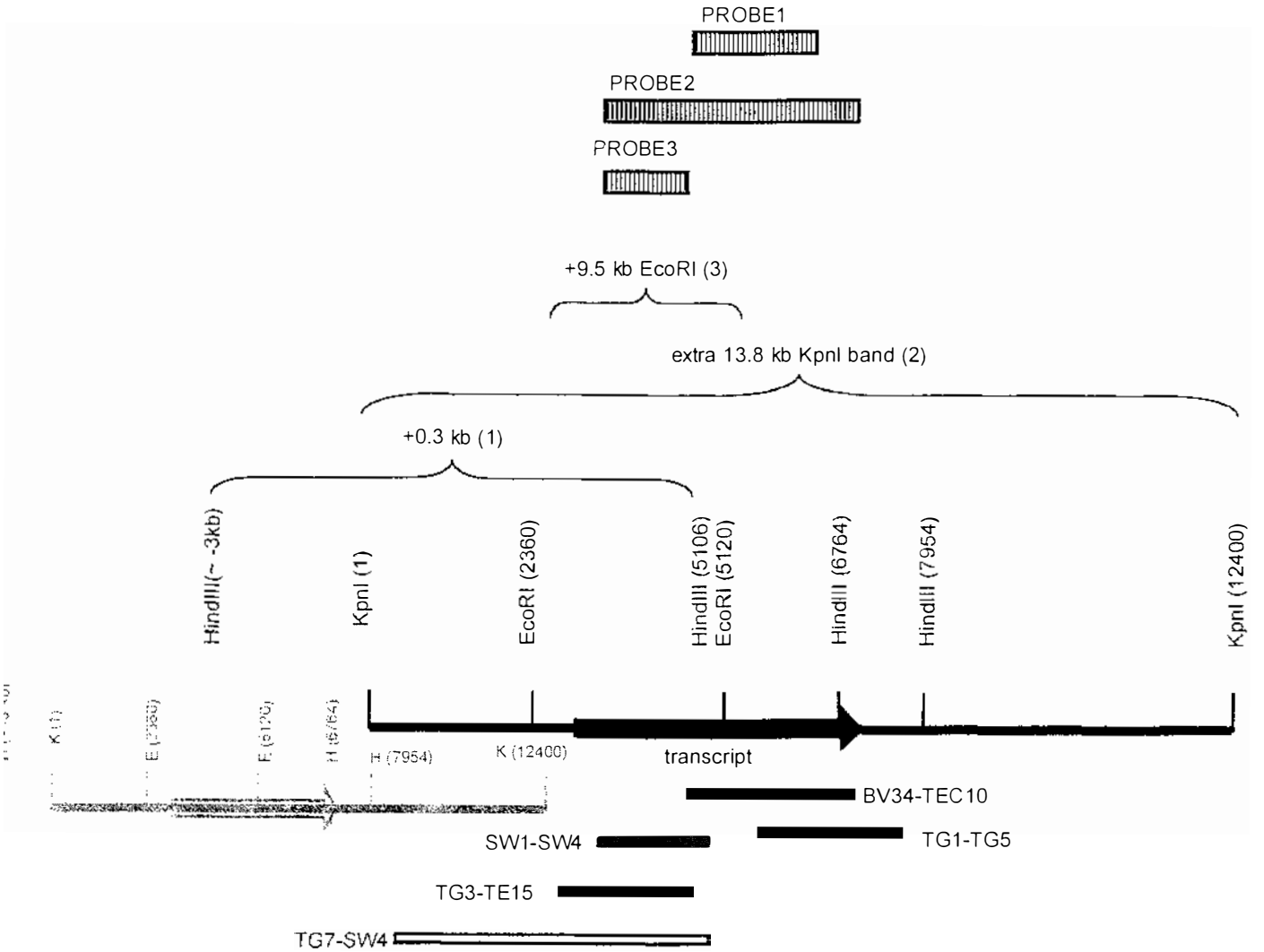


Figure 2.14. Summary of *te1-mum7* allele PCR and Southern analysis. Symbols are as described for Figure 2.11: striped bars above gene line represent probes for Southern blots, open bars below line represent successful PCR, open bars below line represent unsuccessful PCR. One hypothesis for the position of a duplication of the *te1* gene that is consistent with the data is shown in grey (not to scale). PCR experiments indicate that the entire coding region of the *te1-mum7* allele is intact and that from the primer TG7 to the EcoRI site is abnormal (389 to 2360). The extra 13.8 kb KpnI fragment seen in Southern 2 indicates that a large region has been duplicated. The increased size of the EcoRI fragment seen in Southern 3 (9.5 kb larger than normal), and the positioning of the successful PCR amplifications indicates the approximate 5' position of the duplication point.

DISCUSSION (2)

2.14. Analysis of the *te1* genomic sequence

The only irregularity in the *te1* genomic sequence was the apparently non-canonical 5' splice site of intron three (see Figure 2.3), which does not conform to the consensus of either the most common GT-AG type introns or the rare AT-GC type introns (Lorkovic 2000). Although sequence of this region was identical at fourfold coverage from two subclones this irregularity could well be due to a cloning/sequencing artefact and should be confirmed from maize genomic DNA, perhaps from more than one inbred line (the current genomic sequence is from the B73 inbred). The genomic organisation is essentially the same as the putative *Arabidopsis* orthologs *TEL1* and *TEL2*, although the *Arabidopsis* orthologs appear to have lost an intron (see Figure 3.6). Exon structure is discussed in more detail in Chapter 3.

Of the three possible AUG translation initiation sites (see Figure 2.3), the second is most similar to the monocot consensus c(a/c)(A/G)(A/C)cAUGGCG (Joshi 1997), being consistent at 9/11 sites, the others are consistent at 6/11 sites.

2.15 Expression studies

2.15a. 5' RACE

The 5' RACE experiments were not conclusive, since a) the first result, which predicted the transcription start site to be at 2971 bp (see Figure 2.3 for genomic sequence context and Section 2.11 for 5' RACE results), and b) an RT-PCR experiment that used a primer set that extended 5' to the predicted end of the transcript and 3' to beyond the first intron produced no product (Figure 2.4 Section 2.11).

There are two possible sources for of amplification products not derived from a full length *te1* cDNA. Firstly, genomic DNA contamination could have been present in the polyA+ preparations prepared with magnetic beads (Section 2.6c). These RNA samples were treated with DNaseI to the extent that the PCR assay would not amplify a 1.6 kb portion

of the *te1* gene, but since the cleaved DNA was not purified by precipitation it could still have been present in the RNA samples. To avoid this artefact polyA+ samples should be precipitated after DNaseI treatment to remove any contaminating genomic DNA fragments. Also, primers for the reverse transcription and 3' priming of PCR designed to beyond the first intron will distinguish between cDNA and genomic DNA.

Secondly primers could have amplified some other cDNA. The best solution to this problem may be to use one of the reverse transcription enzymes available that are stable at high temperatures. However, experiments not shown in this thesis indicated that, in my hands, the reportedly heat stable Roche 'C-therm' reverse transcriptase actually produced less *te1* cDNA at 50°C than Moloney Murine Lukemia Virus enzyme. Other products may perform differently, and enzyme/primer sets should be optimised as much as possible.

Methods that treat the RNA sample with calf intestinal phosphatase, which removes the 5'-phosphate from molecules with uncapped 5'ends (particularly fragmented/degraded mRNA), should be more reliably obtain 5' RACE products exclusively from full length mRNAs (Schaefer 1995). In this method, the RNA is then treated with tobacco acid pyrophosphatase (TAP) which removes the cap structure leaving behind a 5'-monophosphate, and a synthetic RNA adapter is then ligated to the 5'-monophosphate. Since 5' capped mRNAs are not dephosphorylated by calf intestinal phosphatase they still retain the ability to ligate to the RNA primer. Several kits are commercially available that utilise this method (Ambion RLM-RACE Kit, for example).

2.15b. Expression of *te1* and meristem function

The initial model (Veit 1998) focused on the role of *te1* in vegetative SAMs. Mutant alleles of *te1* show irregularly shortened internodes, abnormal spiral or disordered phyllotaxy, and leaf pattern defects such as leaves with two midribs. The accumulation of the *te1* gene in normal vegetative SAMs is positioned in a semicircle that brackets the positions of leaf primordia. This data is consistent with a model for *te1* function wherein the gene product is required for correct positioning of leaf primordia, and acts to repress organ formation. Mutants phenotypes can be envisaged as being the result of the organ formation process being de-repressed. The lack of expression in mutant apices disrupts the normal partitioning of cells into organ *vs.* internode fate, causing short internodes.

These short internodes are not due to lack of cell expansion in the shoot, in fact the short internodes have larger cells with disordered cell division patterns (N Alvarez, unpublished data). This authors interpretation of the fused midribs that occur occasionally in *te1* mutants is that they are the result of a primordia initiating too close to an existing primordia, and recruiting some of the same cells.

This project found that the *te1* transcript was present in embryo, early tassel, ear and root tissues at approximately the same level as SAMs (Section 2.12), implying that the *te1* gene product functions to repress organ formation in other meristems. A role for *te1* in developing tassels is not unexpected since *te1-1* mutant plant have shorter tassel internodes than normal plants (N Alvarez, unpublished data). Mutant root and ear morphology have not been examined systematically. In this context it is interesting to note that *te1* expression levels generally correlate with internode lengths (Figure 2.15). Leaf primordia initiate closer to the summit of the *te1-1* mutant (Veit 1998). An interpretation of this observation is that the *te1* genes activity limits the longitudinal positioning of primordia. Spiral phyllotaxies and fused are probably the result of leaf 'crowded' leaf initiation.

The *te1* transcript is present in root tips at about the same level as it is in shoot apices. RNA extracts used in this analysis were from the terminal 0.5 cm of roots and excluded zones that were initiating lateral roots. The expression of *te1* in this zone is therefore consistent with a role for *te1* in repression lateral organ formation, although homology of internodes to root extension zones is less clear. All three RT reactions derived from root RNA were from the same RNA extraction, so its possible that this signal is an artefact. Several facts argue against this a) all controls showed that the signal was not genomic DNA, or contaminating vector containing the *te1* cDNA, b) the related Arabidopsis *TEL1* gene was also shown to be expressed in the root (Nena Alvarez, unpublished data).

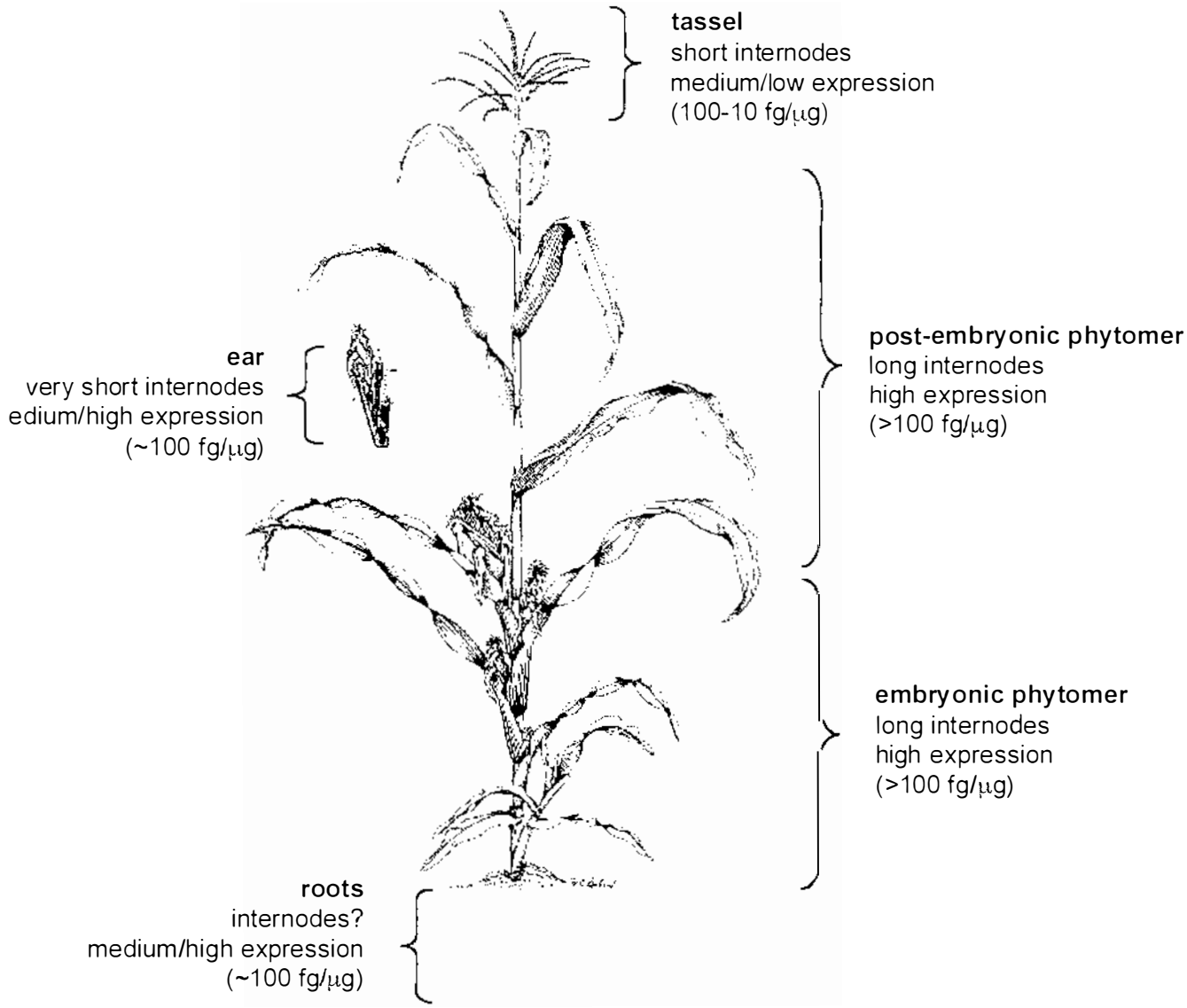


Figure 2.15. Levels of the *terminal ear1* transcript and the extension of maize internodes. The *te1* gene is expressed in all maize meristems. Broadly speaking high expression is associated with longer internodes. This data is consistent with a general role for TEL genes in restricting organ initiation in the peripheral zone to a limited number of cells (see text). This model specifies that the activity of TEL genes determines the positions of organ primordia, and the relative proportions of cells attributed to organ primordia and internode. The correlation between *te1* expression is not entirely consistent, but differential cell expansion could contribute to the irregularities. Source for maize image: Galinat (1994).

2.16. Mutant allele analysis

Analysis of the *te1* mutant alleles is not complete but should yield some interesting insights into the molecular genetics of the gene. It must be noted that there was some uncertainty as to the identity of some seed stocks, and so any further analysis should begin with Southern blots checked against these results to confirm the identity of the stocks used.

The structure of the *te1-1* allele is not well understood, but probably represents a null as far as transcription is concerned (Bruce Veit, unpublished data). The *te1-ems* allele (generated in an EMS mutagenesis screen for *liguless2*) is probably best analysed using RT-PCR to detect and clone transcripts. The mutation is probably a point mutation (the allele looks normal by Southern analysis), and so will be difficult to characterise without sequencing genomic PCR products with the aim of detecting a single base pair change in several kb of sequence. Further characterisation of the *te1-mum7* allele will also be particularly interesting because the entire coding region appears to be normal (Figure 2.9): PCR with TG3 produces a normal product, this primer is at position 2761, 455 bp upstream from the first uninterrupted in-frame ATG.

CHAPTER 3. *IN SILICO* ANALYSIS OF *MEI2*-LIKE GENES

3.1 ABSTRACT

Mei2-like genes encode proteins with three RNA recognition motifs (RRM), of which the third is highly distinctive. The family has been named after the *Schizosaccharomyces pombe Mei2* gene, which is required for meiosis, and was the first to be characterised (Watanabe 1988). However, *Mei2*-like proteins are most abundant in plants. Several other examples from of Mei2-like gene have been detected from fungi, and alveolates, which are eukaryotes that are divergent from both plants and fungi. No *Mei2*-like genes have been identified in metazoans.

Nine full length *Mei2*-like genes are presented in this study, the six *Mei2*-like genes that had been cloned (*S. pombe Mei2,* maize *te1,* and Arabidopsis *AML1, AML2, TEL1* and *TEL2*) and three putative *Arabidopsis* genes (*AML3, AML4, AML5*). Combining predictions with cloned genes, the full compliment of seven Arabidopsis Mei2-like gene is described. Finally, about 30 putative Mei2-like genes from plant, alveolate and fungal ESTs are presented.

Structural predictions of all known full length proteins of this class are compared with the known structures of two RRM proteins, *Drosophila* SEX LETHAL, and human U1A. These comparisons indicate that *Mei2*-like proteins share some novel structures not present in the RRM proteins crystalised so far.

A phylogenetic analysis of *Mei2*-like proteins is presented which clearly identifies two groups in plants; one including the maize TE1 protein and *Arabidopsis* orthologs TEL1 and TEL2, and another including *Arabidopsis* AML1. The evolution and function of *Mei2*-like genes in general is discussed.

METHODS (3)

3.2 tBLASTn searches to identify putative *Mei2*-like genes

The class of *Mei2*-like genes was first identified by tBLASTn (protein query against nucleotide sequence database dynamically translated in all reading frames) using either the TE1 or TEL1 protein sequences. Such searches indicated that *Mei2*-like genes share low overall similarity in RRM1 and RRM2 with canonical RRM proteins, and much higher similarity in the third RRM with a smaller group of proteins (Figure 3.1).

A data set of putative *Arabidopsis Mei2*-like genes was identified from the completed *Arabidopsis* genome (The Arabidopsis genome Initiative, 2000). tBLASTn searches were conducted against *Arabidopsis* genomic sequences using the TEL1, and then AML1 protein sequences as queries. To ensure that that this approach had identified all Arabidopsis *Mei2*-like genes, all sequences whose tBLASTn E values were less than those of canonical RRM proteins (such as splicing factors) were analysed, E values of approximately 10^{-6}, tBLASTn results were also inspected for the presence of the highly conserved third RRM (Section 3.2). At least 7 kb of each accession, starting from approximately 5 kb upstream of the easily recognisable third RRM3 was used for further analysis. To ensure that the *Arabidopsis* data set was non-redundant, all sequences that were on the same chromosome were compared using the GCG program *gap*.

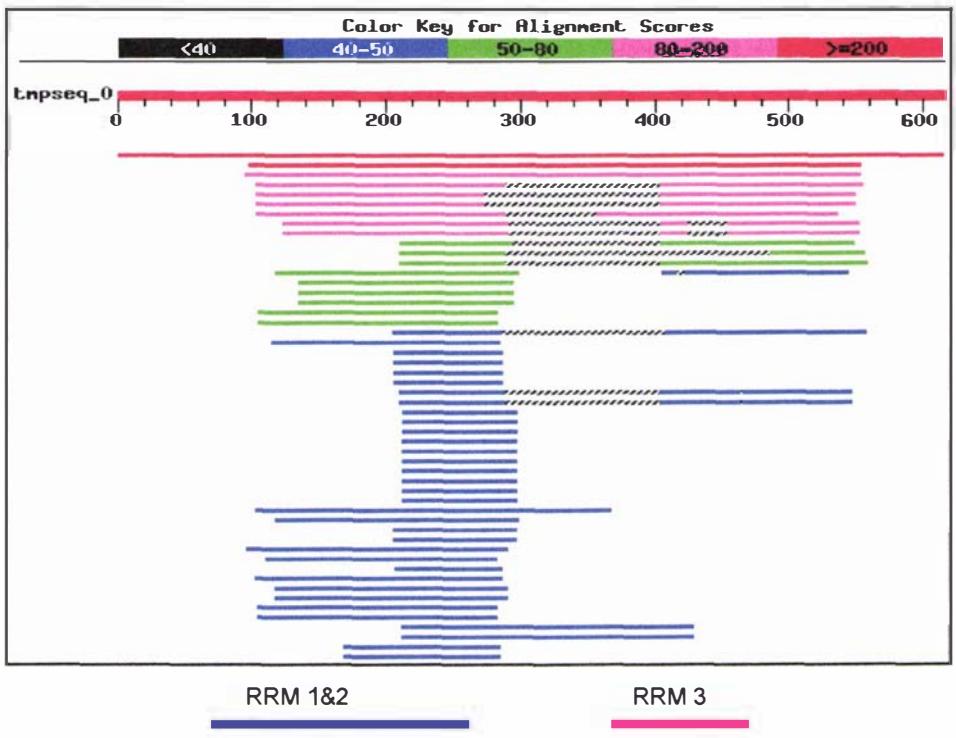


Figure 3.1 tBLASTn searches indicate that *Mei2*-like genes are most similar at the third RRM. A typical result from a tBLASTn search of the non redundant genbank dataset using the TEL1 protein as query. Most genes are similar only at the N-terminus, corresponding to RRM 1 and RRM 2 of the TE1 query (blue line); a small subset of RRM genes (*Mei2*-like genes) are similar over the entire length, including the third RRM (pink line). Experience with such BÆAST searches has indicated that Expect values (E values) from *Mei2*-like genes range from 10^{-164} to 10^{-10}, canonical RRMs usually yield E values of less than 10^{-10}.

3.3. *Arabidopsis* gene prediction methodology.

Gene predictions that had been published in genbank were used as starting points where available. The gene prediction programs Genscan (Burge 1997, Burge 1998), Netgene2 (Brunak 1991, Hebsgaard 1996), and Genefinder (Solovyev (1997), unpublished) (see References for all internet addresses) were used to identify possible splice sites. To determine which exons of those predicted were the most likely to be actual exons, the peptides they encoded were compared to the protein sequences of TE1, *Mei2*, and AML1, for which cDNAs had been cloned (Veit 1998, Hirayama 1997,

Watanabe 1994). The program *pileup* in the Wisonsin Package (Wisconsin Package Version 9.1, Genetics Computer Group (GCG)) with gap weight 6, gap extension weight 1 would usually align peptides from 'real' exons with regions from known Mei2-like proteins. This method made the assumption that *Mei2*-like proteins were similar along their entire length, and thus has the limitation that it will not identify cases where exons have recently been added/omitted during the course of evolution.

In most cases continuous open reading frames could be assembled within splice sites that conformed to the *Arabidopsis* consensus[1]. In a few cases the proteins predicted would not have been produced as described without assuming single nucleotide sequencing errors. Two regions of the *Arabidopsis* genome (genbank accessions AB005249 and AC007505) produced marginal tBLASTn E values, of the order of 10^{-7} to 10^{-3}, depending on the query. Both these regions contained regions whose translation products were similar to the highly conserved region of RRM 3. However, automated gene prediction programs did not assemble these regions into long open reading frames. Closer scrutiny, and manual attempts to assemble open reading frames from possible exons using predicted splice sites revealed the presence of stop codons in many predicted exons (not shown). The conclusion is that these two regions are either pseudogenes, or too distantly related to assemble by joining exons that are similar to other *Mei2*-like genes.

3.4. Promoter analysis

To ensure that the entire promoter of the *Arabidopsis* genes was used for analysis, Netgene2 was used to identify the probable position of the next gene to 5' of *TEL1* and *TEL2* as follows. Each genes genomic sequence was submitted to Netgene2 (Brunak 1991, Hebsgaard 1996). The output from this server gives a graphical view of the probability that a sequence is coding (Figure 3.2). The Netgene2 output was examined by eye for the first (or last) exon of the next gene upstream. Predicted genes precede *TEL1* by ~1799 nt, and *TEL2* by ~2400 nt. So that results could be compared between *te1*, *TEL1* and *TEL2*, the minimum length of 1799 nt upstream from the ATG start codon was used for promoter analysis. From RT-PCR cloning of the *TEL1* and *TEL2*

[1]. Consensus from the TAIR internet site. (see References).

transcripts, it was known that the transcripts extended *at least* as far as the ATG that had been predicted. These sequences (*te1-p*, *TEL1-p* and *TEL2-p* in supplementary files) were submitted to the internet search tool PLACE (Version 6.01, see References for internet address) which searches a database of plant promoter elements. Promoter element sequences were also compared by a multiple alignment with the GCG programs *pileup*, and *pretty* (consensus plurality 2).

RESULTS (3)

3.4. Seven *Arabidopsis Mei2*-like genes have been identified

3.4a. Prediction of *Arabidopsis Mei2*-like genes

The completed *Arabidopsis* genome data was used to compile a set of *Arabidopsis Mei2*-like genes. The aim was to identify all the *Mei2*-like genes of *Arabidopsis* (see Table 3.1). The *AML1* (*Arabidopsis Mei2*-Like) cDNA had been identified from a screen of *Arabidopsis* cDNAs that could complement a *Schizosaccharomyces pombe* meiosis mutant (Hirayama 1997). Eight additional genomic regions were identified by tBLASTn searches. Gene prediction programs (Section 3.3) were used to identify putative proteins from six of these genomic regions, and two further regions could did not yield meaningful predictions. The corresponding genes are referred to as *TEL1, TEL2* (*terminal ear 1-like*), and *AML2*, *AML3*, *AML4*, *AML5* (*Arabidopsis Mei2*-Like) according to their similarity to the corresponding cloned genes (see Section 3.6 and 3.7). Annotated genomic sequences for *AML3*, *AML4* and *AML5* are included in supplementary files. These files show the predicted exons, their coding potential, and the positions of all ESTs identified by BLASTn searches (section 3.4b).

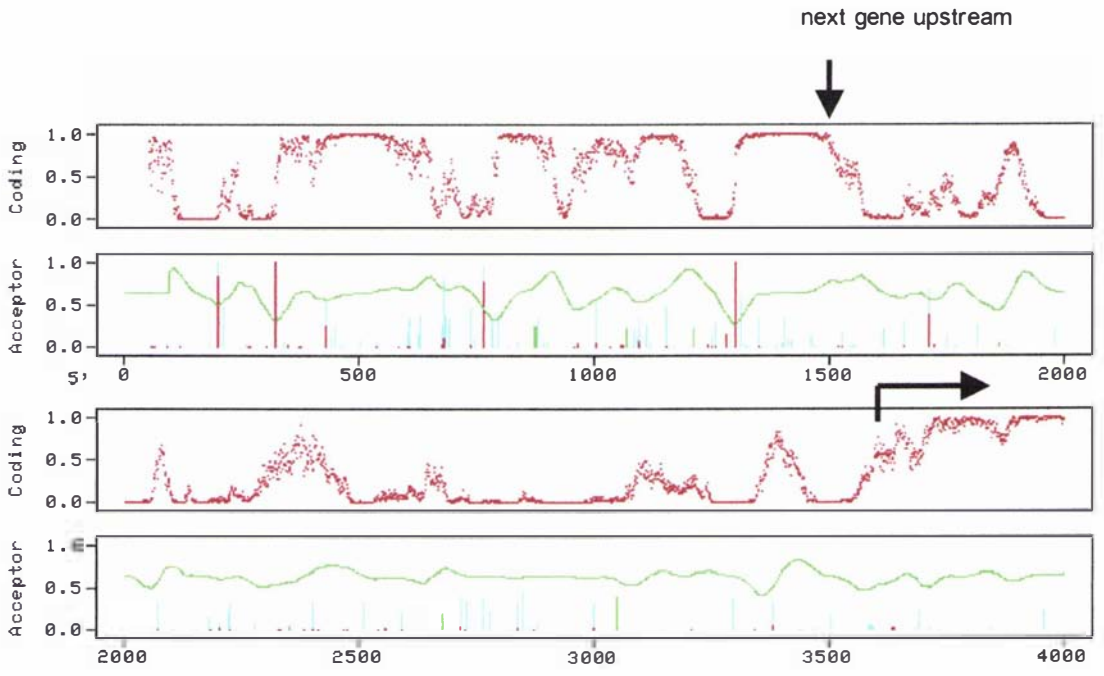


Figure 3.2. Netgene2 graphical output predicting probably of coding in the *TEL1* gene. The entire 9642 nt of sequence (AP000601) was submitted to Netgene2. Only part of the graphical output of the 5' region is shown. The start ATG is positioned at 3677 nt, and it is clear from the graph of coding probability (top boxes, red) that another gene present up to about 1500 nt (arrow). The orientation of this upstream gene was not determined.

3.4b. *Arabidopsis Mei2*-like cDNAs and ESTs

Genetic and molecular analysis of the maize *te1* gene has indicated that this gene is required for the correct positioning of leaf primordia in the shoot apical meristem (Veit 1998). Because our interests were to extend our understanding of developmental processes of the shoot apex, we were most interested in the two *Arabidopsis* genes that were most similar to the maize *te1* gene; *TEL1* and *TEL2*. RT-PCR was used to clone cDNAs from these two genes (V. Trainor, unpublished). The cDNAs were found to conform to the predictions. A cDNA for the gene we refer to as *AML2* was also cloned from seedling tissue (Bourdon 1998).

Table 3. 1 Arabidopsis Mei2-like genes

Gene	Genbank accession	BAC clone	C/some	Position[3]	Expression[4]	cDNA cloning (Accession)	Notes
TEL1	AP000601 (800-7800, D)	MJL14	3	Start 1.05322197 End 1.0541861E7	In situ: Embryo SAM	Trainor[5]	
TEL2	AC008113 (29000-36000)	F12A21	1	Start 2.974668E7 End 2.9860872E7	In situ: Embryo SAM	Trainor	
AML1	ABO22212 (23200-30200)	K22G18	5	Start 3.147718E7 End 3.1522632E7	Northern: All adult tissues	Hirayama[6] (D86122)	Cloned by complementing S. pombe meiotic mutant
AML2	AC006931 (21900-28900)	F7D19	2	Start 1.8887852E7 End 1.9004056E7	cDNA: seedlings	Bourdon[7] (AF101056)	
AML3	AL110123 (23000-30000), also AL161548	F15J5	4	Start 1.4874061E7 End 1.4932488E7	EST: Green siliques, Seedling hypocotyls, Adult rosettes	-	
AML4	AL163652 (6800-13800) also AL163912	T28J14	5	Start 4060816.0 End 4165423.0	EST: Roots, Developing seed, 2-6 week seedlings	-	
AML5	AC068667 (29,000-33,000 R)	F15D2	1	Start 1.1176639E7 End 1.1311108E7	EST: Seedling hypocotyls, 2-6 week seedlings	-	Annotated in genbank as F15D2.30
AML6?	AC007505 (43,000-53,000)	F28L22	1	Start 1.6019605E7 End 1.6154104E7	EST[8]; No tissue information	-	
AML7?	AB005249 (50,000-57000)	MXM12	5	Start 4342753.0 End4426352.0	No ESTs found.	-	Pseudogene? (see text)

[3] These positions refer to the position in nucleotides of the BAC clone. Positions for BAC clones can be located at http://www.arabidopsis.org/servlets/mapper

[4] Sources of information for expression is stated: EST data, Northern blots, or *in situ* hybridisation, or source from which cDNA was isolated.

[5] Cloned by RT-PCR from floral mRNA extract. Veit laboratory, Trainor V., unpublished.

[6] Hirayama T, Ishida C, Kuromori T, Obata S, Shimoda C, Yamamoto M, Shinozaki K, Ohto C. Functional cloning of a cDNA encoding Mei2-like protein from Arabidopsis thaliana using a fission yeast pheromone receptor deficient mutant. FEBS Lett. 1997 Aug 11;413(1):16-20.

[7] Bourdon,V. and Tinland,B. Unpublished, Direct Submission to Genbank.
Isolation and characterisation of MEI2, the Arabidopsis homolog of the *Schizosaccharomyces pombe* mei2 gene. Submitted (22-OCT-1998) Department of Developmental Biology, Institute of Plant Sciences, ETH Zurich, Universitatstrasse, 2, Zurich 8092, Switzerland

[8] *Arabidopsis* ESTs that correspond to AC007505 by BLAST search range from 6254 to 3400 nt 5' of the region encoding the region identified by tBLASTn searches (VMVKNIPN).

Table 3.2 EST hits from Arabidopsis BLASTn searches of genomic sequences containing putative Mei2-like genes

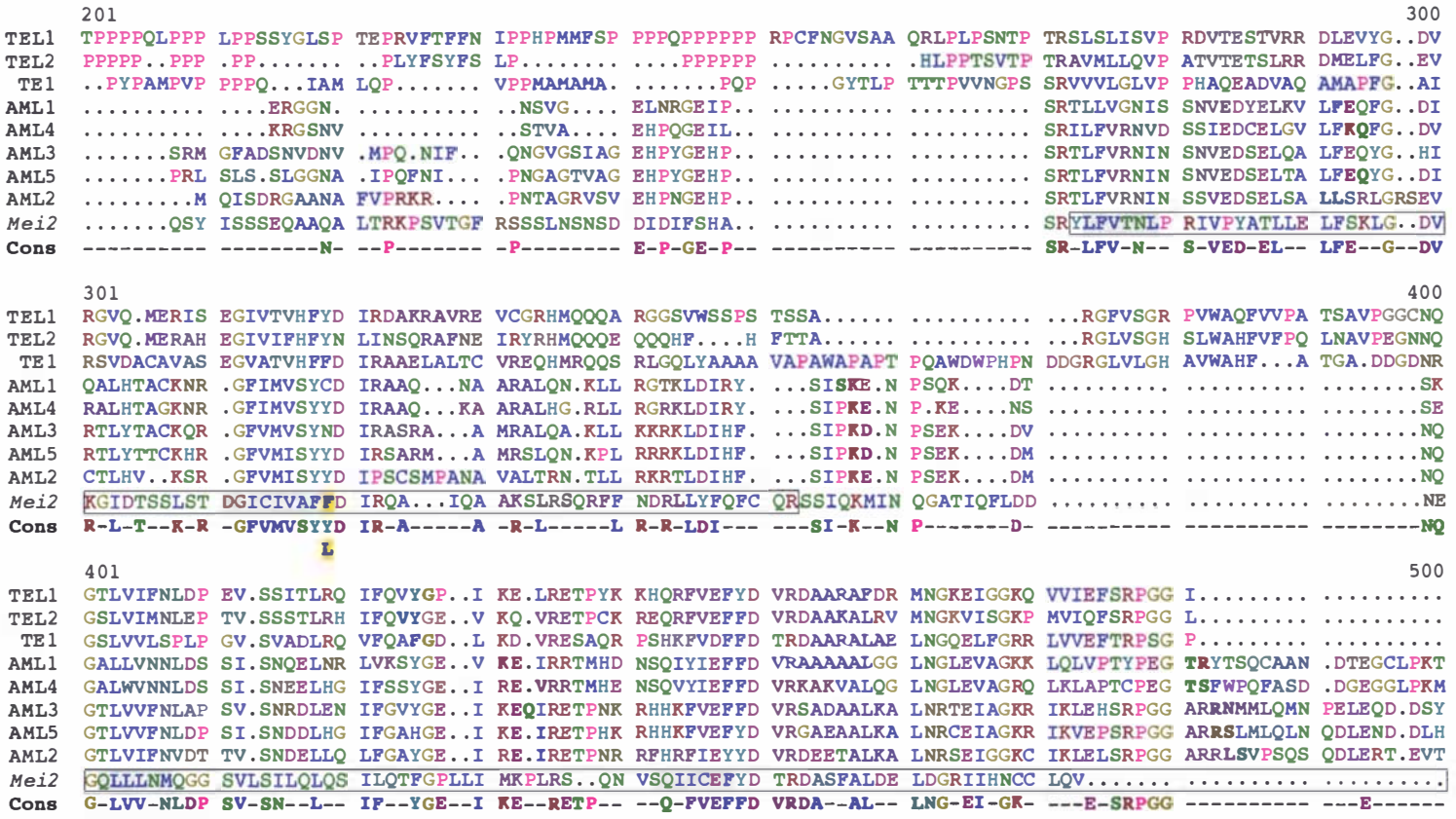
Gene query	EST Accession	E value	tissue
AML3	AV555397	0.0	green siliques
	W43784	0.0	3 day seedling hypocotyl
	AA395462	0.0	mixed source
	AI998063	10^{-180}	rosette 4-7 weeks
	T21020	10^{-9}	mixed source
AML4	AV545804	0.0	roots
	AV543746	10^{-110}	roots
	BE528694	10^{-74}	developing seed
	AV523936	10^{-53}	above ground organs, 2-6 week old seedlings
	AV552144	10^{-40}	roots
	AV529706	10^{-37}	above ground organs, 2-6 week old seedlings
AML5	AA650745	0.0	mixed source
	AA006464	0.0	withdrawn from genbank
	AV529784	10^{-145}	2-6 week old above ground organs
	BE524842	10^{-124}	developing seed, 5-13 days after flowering
	BE527698	10^{-90}	developing seed, 5-13 d.a. flowering
	BE530712	10^{-84}	developing seed, 5-13 d.a. flowering
	W43781	10^{-79}	3 day old seedling hypocotyl
	BE522161	10^{-69}	developing seed, 5-13 d.a flowering
	AV528631	10^{-62}	2-6 week old above ground organs
	T04404	10^{-54}	mixed source
AC007505 (*AML6*?)	AW004409	10^{-22}	mixed source
	AA006249	10^{-19}	this record was removed at the submitters request
	AA605391	10^{-19}	mixed source
	AI994379	10^{-10}	not stated
AB0053249	No significant hits	≥0.67	

3.5. Three RRMs can be identified in all complete *Mei2*-like genes.

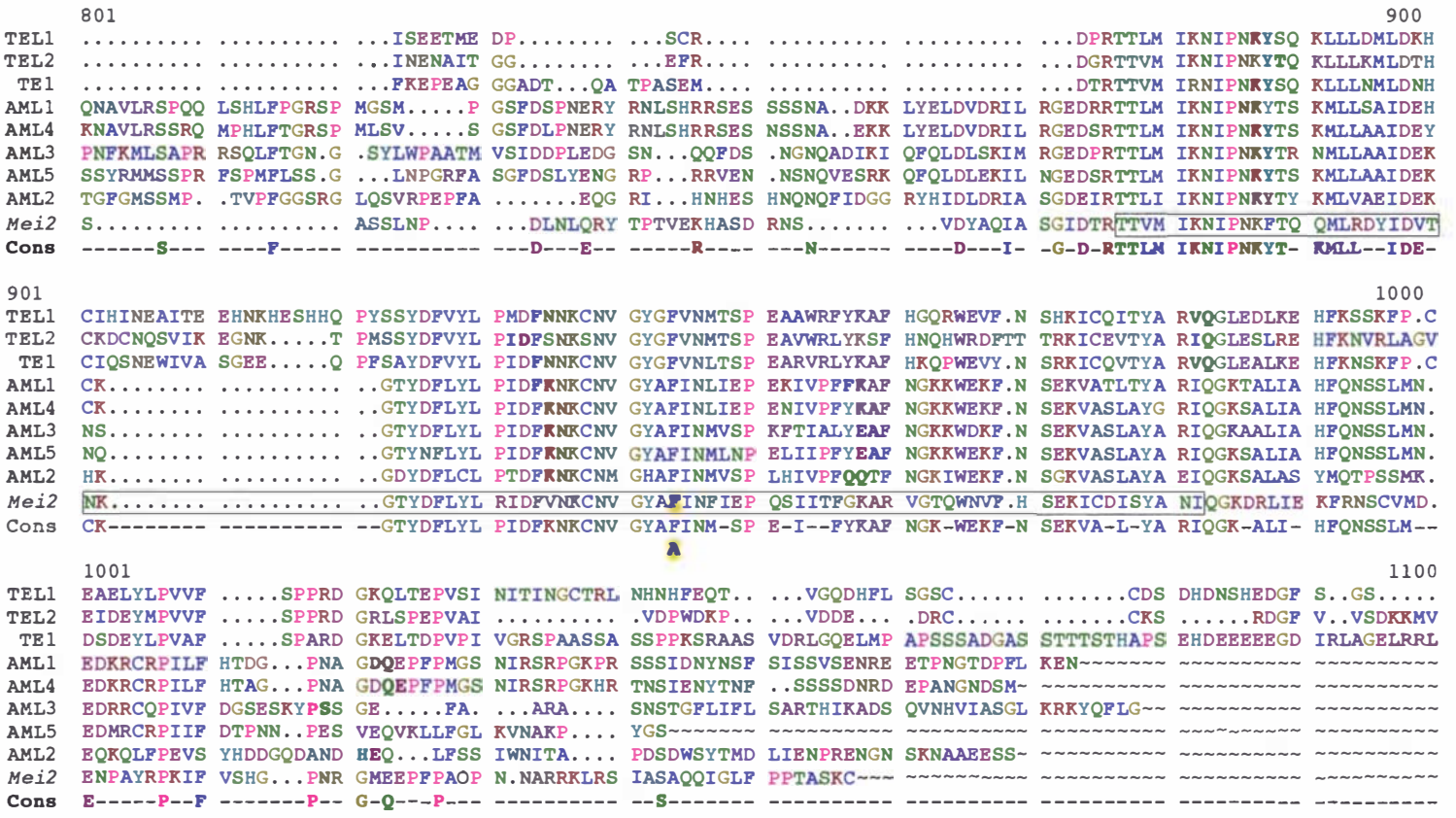
Putative protein sequences of all complete plant *Mei2*-like genes (*TE1, TEL1, TEL2, AML1-5*) and *S. pombe Mei2* were compared by multiple alignment using the GCG program *pileup*. With gap weight 6, gap length weight 1 three regions produced meaningful alignments (Figure 3.3). The first two correspond to the canonical RRMs as previously described (Birney 1993), and the C-terminal region to the third RRM shared between TE1 ad *Mei2* (described in Veit 1998). Within the RRMs as described by Birney (1993) and the Pfam database, RRM1 is 4% identical and 21% conserved, RRM2 is 14% identical and 46% conserved, and RRM3 is 23% identical and 51% conserved.

Figure 3.3. Alignment of all full length *Mei2*-like genes. The peptide sequences for all *Mei2*-like gene currently known or predicted in this thesis were aligned with the GCG program *pileup* (gap weight 6, gap extension 1), and a consensus produced from this alignment with the program *pretty* (consensus plurality 5). Amino acids are colour coded, amino acid groupings are; CAIVLMFW, G, TSNQ, KR, P, ED, YH. The first two RRMs of *Mei2* as defined in the Pfam database entry for the RRM domain, are indicated with a black border, RNP-1 and RNP-2 submotifs are shaded. The third RRM was identified by a structure-guided manual alignment of *Mei2*-like genes with canonical RRM proteins (Figure 3.4). The similarity of *Mei2*-like proteins beyond the RRMs, particularly about the third RRM. TE1-like proteins (TE1, TEL1, TEL2) share some insertions/deletions not present in the AML group (AML1-5). The positions of the two mutations in *S. pombe* MEI2 that decrease (F240L), or eliminate (F644A) RNA binding are indicated with yellow backgrounds (F). The residues that occur either side of the introns are indicated with bold underline (e.g. FKQF).

```
1                                                                                                   100
TEL1  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
TEL2  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 TE1  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
AML1  MKTDMMPENQ GGRDRLSSMP KSSWTSESYQ LKPQSSFSGS HPSGSPNARN TTNG.....S QWESSLFSSS MSDLFSRKLR LQGS.DMLST M...SANTVV
AML4  MKNNPMPE.. GGVDRSSNLP TSSWTSDSYQ LSQQSSLSGA LPSFIPNGRT TTND.....T HWESSLFSSS LSDLFSRKLR LPRS.DKLAF M...SAN...
AML3  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ MASGLDHFS. ..........
AML5  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~MDIPHE AEAGAWGILP GFGRHHHPSS DATLF..SMQ LSDNRDGFSL IDDTAVSRT.
AML2  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~ME LEPNKSLSAD MPSLLSRSSE AFNGGTGYRS SSDLSMFSSS LPTLFHEKLN MTDS.DSWLS FDESSPNLNK
Mei2  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~MI METESPLSIT SPSPSDST.. .......FQV DMEKTMHALP SSLLDSPLLS TNEHYPPKST LLLSGPSPIR
Cons  ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---L------ -----D---- ----------

      101                                                                                           200
TEL1  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~MEDSR LFPF.VGN.L DPRAQEFIPF .......NP. .ISSGFHFPY
TEL2  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~MSVTG PFSH.PTN.L NPTAPAFFPA INQHQNQNPS LIPTRFFLPH
 TE1  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~MG GFPEATGNLL DAAAQEFHPT VCAPYPLQP. .LPQQLYCPH
AML1  THREEEPSES LEEIEAQTIG NLLP.DEDDL FAEVT.GEVG RKSRANTGDE LDEF.....D LFSSVGGMEL DGDIFS.... .......SVS .HRNG.....
AML4  ..REEEPSES LEEMEAQTIG NLLP.DEDDL FAEVV.GEGV HKSRANGGDD LDDC.....D LFSSVGGMEL DGDVFS.... .......SVS .QRDG.....
AML3  ....GGIGNM LDDGDSHPIG NMLPDDEEEL FSGLM.DDLN LSSLPATLDD LEDY.....D LFGSGGGLEL ETDPYD.... .......S.. LN.KGF....
AML5  ....NKFNES ADDFESHSIG NLLP.DEEDL LTGMM.DDLD LGELP....D ADDY.....D LFGSGGGMEL DADFRD.... .......N.. LSMSGP....
AML2  LVIGNSEKDS LEDVEPDALE ILLPEDENEL LPGVI.DELN FTGLPDELDD LEEC.....D VFCTGGGMEL DVESQD.... .......NHA VDASG.....
Mei2  NIQLSATKSS ....ESNSID YL..TDTQNI FPNFVNNENN YQFSTAPLNP IDACRVGERK VFTT.GNVLL SADRQPLSTW QQNISVLSES PPQNGI....
Cons  ---------S ----E---I- -LLP-DE--L ---------- ---------- ---------D LF---GG-EL D-D------- ---------- ----G-----
```


201 300
TEL1 TPPPPQLPPP LPPSSYGLSP TEPRVFTFFN IPPHPMMFSP PPPQPPPPPP RPCFNGVSAA QRLPLPSNTP TRSLSLISVP RDVTESTVRR DLEVYG..DV
TEL2 PPPPP..PPP .PP..... ..PLYFSYFS LP..... ....PPPPPP .... .HLPPTSVTP TRAVMLLQVP ATVTETSLRR DMELFG..EV
TE1 ..PYPAMPVP PPPQ...IAM LQP..... VPPMAMAMA. .....PQP .....GYTLP TTTPVVNGPS SRVVVLGLVP PHAQEADVAQ AMAPFG..AI
AML1 ....ERGGN. ....NSVG.... ELNRGEIP.. .... SRTLLVGNIS SNVEDYELKV LFEQFG..DI
AML4 ....KRGSNV ....STVA.... EHPQGEIL.. .... SRILFVRNVD SSIEDCELGV LFKQFG..DV
AML3 ....SRM GFADSNVDNV .MPQ.NIF.. .QNGVGSIAG EHPYGEHP.. .... SRTLFVRNIN SNVEDSELQA LFEQYG..HI
AML5 ....PRL SLS.SLGGNA .IPQFNI... .PNGAGTVAG EHPYGEHP.. .... SRTLFVRNIN SNVEDSELTA LFEQYG..DI
AML2 ....M QISDRGAANA FVPRKR.... .PNTAGRVSV EHPNGEHP.. .... SRTLFVRNIN SSVEDSELSA LLSRLGRSEV
Mei2 ....QSY ISSSEQAAQA LTRKPSVTGF RSSSLNSNSD DIDIFSHA.. .... SRYLFVTNLP RIVPYATLLE LFSKLG..DV
Cons ---------- --------N- --P------- -P-------- E-P-GE-P-- ---------- ---------- SR-LFV-N-- S-VED-EL-- LFE--G--DV
301 400
TEL1 RGVQ.MERIS EGIVTVHFYD IRDAKRAVRE VCGRHMQQQA RGGSVWSSPS TSSA..... ....RGFVSGR PVWAQFVVPA TSAVPGGCNQ
TEL2 RGVQ.MERAH EGIVIFHFYN LINSQRAFNE IRYRHMQQQE QQQHF....H FTTA..... ....RGLVSGH SLWAHFVFPQ LNAVPEGNNQ
TE1 RSVDACAVAS EGVATVHFFD IRAAELALTC VREQHMRQQS RLGQLYAAAA VAPAWAPAPT PQAWDWPHPN DDGRGLVLGH AVWAHF...A TGA.DDGDNR
AML1 QALHTACKNR .GFIMVSYCD IRAAQ...NA ARALQN.KLL RGTKLDIRY. ...SISKE.N PSQK....DT ....SK
AML4 RALHTAGKNR .GFIMVSYYD IRAAQ...KA ARALHG.RLL RGRKLDIRY. ...SIPKE.N P.KE....NS ....SE
AML3 RTLYTACKQR .GFVMVSYND IRASRA...A MRALQA.KLL KKRKLDIHF. ...SIPKD.N PSEK....DV ....NQ
AML5 RTLYTTCKHR .GFVMISYYD IRSARM...A MRSLQN.KPL RRRKLDIHF. ...SIPKD.N PSEK....DM ....NQ
AML2 CTLHV..KSR .GFVMISYYD IPSCSMPANA VALTRN.TLL RKRTLDIHF. ...SIPKE.N PSEK....DM ....NQ
Mei2 KGIDTSSLST DGICIVAFFD IRQA...IQA AKSLRSQRFF NDRLLYFQFC QRSSIQKMIN QGATIQFLDD ....NE
Cons R-L-T--K-R -GFVMVSYYD IR-A-----A -R-L-----L R-R-LDI--- ---SI-K--N P-------D- ---------- ---------- --------NQ
L
401 500
TEL1 GTLVIFNLDP EV.SSITLRQ IFQVYGP..I KE.LRETPYK KHQRFVEFYD VRDAARAPDR MNGKEIGGKQ VVIEFSRPGG I..... .....
TEL2 GSLVIMNLEP TV.SSSTLRH IFQVYGE..V KQ.VRETPCK REQRFVEFFD VRDAAKALRV MNGKVISGKP MVIQFSRPGG L..... .....
TE1 GSLVVLSPLP GV.SVADLRQ VFQAFGD..L KD.VRESAQR PSHKFVDFFD TRDAARALAE LNGQELFGRR LVVEFTRPSG P..... .....
AML1 GALLVNNLDS SI.SNQELNR LVKSYGE..V KE.IRRTMHD NSQIYIEFFD VRAAAAALGG LNGLEVAGKK LQLVPTYPEG TRYTSQCAAN .DTEGCLPKT
AML4 GALWVNNLDS SI.SNEELHG IFSSYGE..I RE.VRRTMHE NSQVYIEFFD VRKAKVALQG LNGLEVAGRQ LKLAPTCPEG TSFWPQFASD .DGEGGLPKM
AML3 GTLVVFNLAP SV.SNRDLEN IFGVYGE..I KEQIRETPNK RHHKFVEFFD VRSADAALKA LNRTEIAGKR IKLEHSRPGG ARRNMMLQMN PELEQD.DSY
AML5 GTLVVFNLDP SI.SNDDLHG IFGAHGE..I KE.IRETPHK RHHKFVEFYD VRGAEAALKA LNRCEIAGKR IKVEPSRPGG ARRSLMLQLN QDLEND.DLH
AML2 GTLVIFNVDT TV.SNDELLQ LFGAYGE..I RE.IRETPNR RFHRFIEYYD VRDEETALKA LNRSEIGGKC IKLELSRPGG ARRLSVPSQS QDLERT.EVT
Mei2 GQLLLNMQGG SVLSILQLQS ILQTFGPLLI MKPLRS..QN VSQIICEFYD TRDASFALDE LDGRIIHNCC LQV..... ....
Cons G-LVV-NLDP SV-SN--L-- IF--YGE--I KE--RETP-- --Q-FVEFFD VRDA--AL-- LNG-EI-GK- ---E-SRPGG ---------- ---E------

```
      501                                                                                                                600
TEL1  .......... .......... .......... .......... .......... .......... .......... .......... ....KNR..F RS....SRQP
TEL2  .......... .......... .......... .......... .......... .......... .......... .......... ....TKKLFF AS....HFHK
 TE1  .......... .......... .......... .......... .......... .......... .......... .......... ...GPRRRGY AP....HQHR
AML1  SYSNTSSGHI GRHFPGMISS TS.SDGGSMR VIHNS...IG ..SPVNSFIE RHRSLSIPIG FPPSANGISA SKPVGLQEHG HHFDNSNMGI QSMPNLHPHS
AML4  AFNNLSSAHM GRHFPGILAS TS.IDGGSIR GMHNS...VG ..SPMNSFIE RHQSLDVPIG LPPSARVISA SKPVGLQEFG NPFDNSKTGI QSMPNLHPH.
AML3  SYLNHVESPL ASSPIGNW.R NSPID.HPLQ S.FSKSPIFG NLSPTKN..I RYPEFSMKTA SVNNDQEGRR FSHL...DH. LFSSSS...Y NNA.S...HK
AML5  .YLPMIGSPM ANSPPSNWPL NSPVEGSPLQ SVLSRSPVFG .LSPTRN... ..GHLSGLAS ALNSQGPSSK LAPI...GRG QIGSNG...F QQS.S...H.
AML2  NFYNQVGSHV ANSPPGNWPI GSPVKGSP.S HAFTRPHGLG MVRPVNS..D NMPGLASILP AHPSSFHG.. FSPVS.NDQG LLNHSNQTIL NKGLM...HN
Mei2  AYYDAMADSV STSSAS.... .......... .......... .......... ...SLSVPRG F......... ........SG MLNNNSE..W NNSMTMSSNQ
Cons  ---------- ---------- -S-------- ---------G ---P------ ----L----- ---------- ---------G ---------- --------H-

      601                                                                                                                700
TEL1  QLPFQPLRE. .PPILIPPLR RPVSFIKDKN KN........ .......... .....VSPK. ....NGVIAV D....ASMRS LCIIDA.DDN KTRGTRGTES
TEL2  NFIFNNEHHY YPP...PP.. .......... .......... .......... .......PP. ....SRMVKS D....ILMYK .......... ..........
 TE1  .......... .PTAPTPP.R LQATWRPSQP TS........ .......... .....SQPPA SSSSSGSVRA REGVVLLRRS SCKSSAGSDQ SSKGGNAGTS
AML1  FSEYVDNFAN GSPYTSS.AF SEMVSDGSKA NEGFMIHNVR ..GVEGFSGG GI..GS.PMH QSSRR.PINL WSNSNTQQQN PSSGMMWPNS PSHINSIPTQ
AML4  FPDYLDNFAS GSPYKSSTTF SEMVSDGQKA NEGFMMSNVR GVGVDGFNGG VI..GS.PIN QGSHRGNLNL WSNSNSQQHN QSSGMMWPNS PSRVNGVPSQ
AML3  ASTF...... QQPQS..... F.....GSVS SFGSLNS..H PSHVETLSGS EFLWGS.P.. .......... .......... ..SSSAWPV. ....NPFSSN
AML5  ..LF...... QEPKM.DNKY T.....GNLS PSGPLIS..N GGGIETLSGS EFLWGS.PNA R......... ........SE PSSSSVWSTS .STGNPLFST
AML2  IS.Y...... GQPHSLPEHI T.....GGIS NSMRFIA.PH SSGFGTSSDH RYRWGS.PPQ HMNYPG.... .......YTG VSSSS....S .STERPFTV.
Mei2  .......... ETPTAASCAV SRIGSSYGMS NN........ .......... ...FGSVPLG RTE...SSPA WGTSG..YYD VSSTS..PVA PSDRNP..SR
Cons  ---------- --P------- ------G--- ---------- ---------- ----GS-P-- ---------- ---------- -SS-S----- -S--N-----

      701                                                                                                                800
TEL1  ECAETK.... .......SKN VAKWGKKRQM KN........ .......... .......... ......MELS .......... ....QFL... ..........
TEL2  ..QQQK.... .......KKK .....KKYVK KN........ .......... .......... ......LGDP .......... ....YFM... ..........
 TE1  HERKTKGGKI VVAAAAASSS TPTASGKQTQ KG........ .......... .......... ......VGSS GGGSWKGRKS GWEARFL... ..........
AML1  R.PPVTVFSR APPIMVNMAS SPVH..HHI. .GSAPVLNSP FWDRRQAYVA ESLESSGFHI GSHG.SMGIP GSSPSHPMDI GSHKTFS.V. .GGNRMDVNS
AML4  RIPPVTAFSR ASPLMVNMAS SPVH..HHI. .GSAPVLNSP FWDRRQAYVA ESPESSGFHL GSPG.SMGFP GSSPSHPMDF GSHKVFSHV. .GGNRMEANS
AML3  RENHR..F.. ..PYSAQNGS .L.HQLHHI. .GSAP..SG. ....FFPRSP ETSS.....M GSVAFR.GAS G.NMNA.... .......... .QRNLRETSS
AML5  RVDRSVPF.. ..PHQHQNQS RS.HHHFHV. .GSAP..SGV PLEKHFGFVP ESSKDALF.M NTVGLQ.GMS GMGLNGGSFS SK.MANNGII NSGSMAENGF
AML2  ....RHGF.. ..PFAERQAS LLGKYQHHV. .GSAP..SSI ....HF.... .NTQMNCY.T GSPEIPLGFS DMGINRNYNS AHGKANLGVS LPGNSSEQDF
Mei2  QYNSIRYGLD VNPIAPPNSS RLKQRNSDLL NGINPQWS.. ......PFSS NTGKVFDSPT GSLGMRRSLT .VGANASCSN PTNLSFASLT LHDSKADSTL
Cons  -------F-- --P------S -------H-- -GSAP----- ---------- ---------- GS-----G-S G--------- -----F---- ----------
```


801 900
TEL1 ....ISEETME DP..... ...SCR.... ....DPRTTLM IKNIPNKYSQ KLLLDMLDKH
TEL2 ....INENAIT GG..... ...EFR.... ....DGRTTVM IKNIPNKYTQ KLLLKMLDTH
TE1 ....FKEPEAG GGADT...QA TPASEM.... ....DTRTTVM IRNIPNKYSQ KLLLNMLDNH
AML1 QNAVLRSPQQ LSHLFPGRSP MGSM.....P GSFDSPNERY RNLSHRRSES SSSNA..DKK LYELDVDRIL RGEDRRTTLM IKNIPNKYTS KMLLSAIDEH
AML4 KNAVLRSSRQ MPHLFTGRSP MLSV.....S GSFDLPNERY RNLSHRRSES NSSNA..EKK LYELDVDRIL RGEDSRTTLM IKNIPNKYTS KMLLAAIDEY
AML3 PNFKMLSAPR RSQLFTGN.G .SYLWPAATM VSIDDPLEDG SN...QQFDS .NGNQADIKI QFQLDLSKIM RGEDPRTTLM IKNIPNKYTR NMLLAAIDEK
AML5 SSYRMMSSPR FSPMFLSS.G ...LNPGRFA SGFDSLYENG RP...RRVEN .NSNQVESRK QFQLDLEKIL NGEDSRTTLM IKNIPNKYTS KMLLAAIDEK
AML2 TGFGMSSMP. .TVPFGGSRG LQSVRPEPFA .....EQG RI...HNHES HNQNQFIDGG RYHIDLDRIA SGDEIRTTLI IKNIPNKYTY KMLVAEIDEK
Mei2 S..... ..... ASSLNP.... ...DLNLQRY TPTVEKHASD RNS..... ...VDYAQIA SGIDTRTTVM IKNIPNKFTQ QMLRDYIDVT
Cons ------S--- ----F----- ---------- ---D---E-- -----R---- ---N------ ----D---I- -G-D-RTTLM IKNIPNKYT- KMLL--IDE-
901 1000
TEL1 CIHINEAITE EHNKHESHHQ PYSSYDFVYL PMDFNNKCNV GYGFVNMTSP EAAWRFYKAF HGQRWEVF.N SHKICQITYA RVQGLEDLKE HFKSSKFP.C
TEL2 CKDCNQSVIK EGNK.....T PMSSYDFVYL PIDFSNKSNV GYGFVNMTSP EAVWRLYKSF HNQHWRDFTT TRKICEVTYA RIQGLESLRE HFKNVRLAGV
TE1 CIQSNEWIVA SGEE.....Q PFSAYDFVYL PIDFNNKCNV GYGFVNLTSP EARVRLYKAF HKQPWEVY.N SRKICQVTYA RVQGLEALKE HFKNSKFP.C
AML1 CK..... ....GTYDFLYL PIDFKNKCNV GYAFINLIEP EKIVPFFKAF NGKKWEKF.N SEKVATLTYA RIQGKTALIA HFQNSSLMN.
AML4 CK..... ....GTYDFLYL PIDFKNKCNV GYAFINLIEP ENIVPFYKAF NGKKWEKF.N SEKVASLAYG RIQGKSALIA HFQNSSLMN.
AML3 NS..... ....GTYDFLYL PIDFKNKCNV GYAFINMVSP KFTIALYEAF NGKKWDKF.N SEKVASLAYA RIQGKAALIA HFQNSSLMN.
AML5 NQ..... ....GTYNFLYL PIDFKNKCNV GYAFINMLNP ELIIPFYEAF NGKKWEKF.N SEKVASLAYA RIQGKSALIA HFQNSSLMN.
AML2 HK..... ....GDYDFLCL PTDFKNKCNM GHAFINMVSP LHIVPFQQTF NGKIWEKF.N SGKVASLAYA EIQGKSALAS YMQTPSSMK.
Mei2 NK..... ....GTYDFLYL RIDFVNKCNV GYAFINFIEP QSIITFGKAR VGTQWNVF.H SEKICDISYA NIQGKDRLIE KFRNSCVMD.
Cons CK-------- ---------- --GTYDFLYL PIDFKNKCNV GYAFINM-SP E-I--FYKAF NGK-WEKF-N SEKVA-L-YA RIQGK-ALI- HFQNSSLM--
1001 1100
TEL1 EAELYLPVVF .....SPPRD GKQLTEPVSI NITINGCTRL NHNHFEQT.. ...VGQDHFL SGSC..... .....CDS DHDNSHEDGF S..GS.....
TEL2 EIDEYMPVVF .....SPPRD GRLSPEPVAI ..... .VDPWDKP.. ...VDDE... .DRC..... .....CKS .....RDGF V..VSDKKMV
TE1 DSDEYLPVAF .....SPARD GKELTDPVPI VGRSPAASSA SSPPKSRAAS VDRLGQELMP APSSSADGAS STTTSTHAPS EHDEEEEEGD IRLAGELRRL
AML1 EDKRCRPILF HTDG...PNA GDQEPFPMGS NIRSRPGKPR SSSIDNYNSF SISSVSENRE ETPNGTDPFL KEN~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
AML4 EDKRCRPILF HTAG...PNA GDQEPFPMGS NIRSRPGKHR TNSIENYTNF ..SSSSDNRD EPANGNDSM~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
AML3 EDRRCQPIVF DGSESKYPSS GE.....FA. ...ARA.... SNSTGFLIFL SARTHIKADS QVNHVIASGL KRKYQFLG~~ ~~~~~~~~~~ ~~~~~~~~~~
AML5 EDMRCRPIIF DTPNN..PES VEQVKLLFGL KVNAKP.... YGS~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
AML2 EQKQLFPEVS YHDDGQDAND HEQ...LFSS IWNITA.... PDSDWSYTMD LIENPRENGN SKNAAEESS~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
Mei2 ENPAYRPKIF VSHG...PNR GMEEPFPAOP N.NARRKLRS IASAQQIGLF PPTASKC~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
Cons E-----P--F -------P-- G-Q--P--- ---------- --S------- ---------- ---------- ---------- ----------

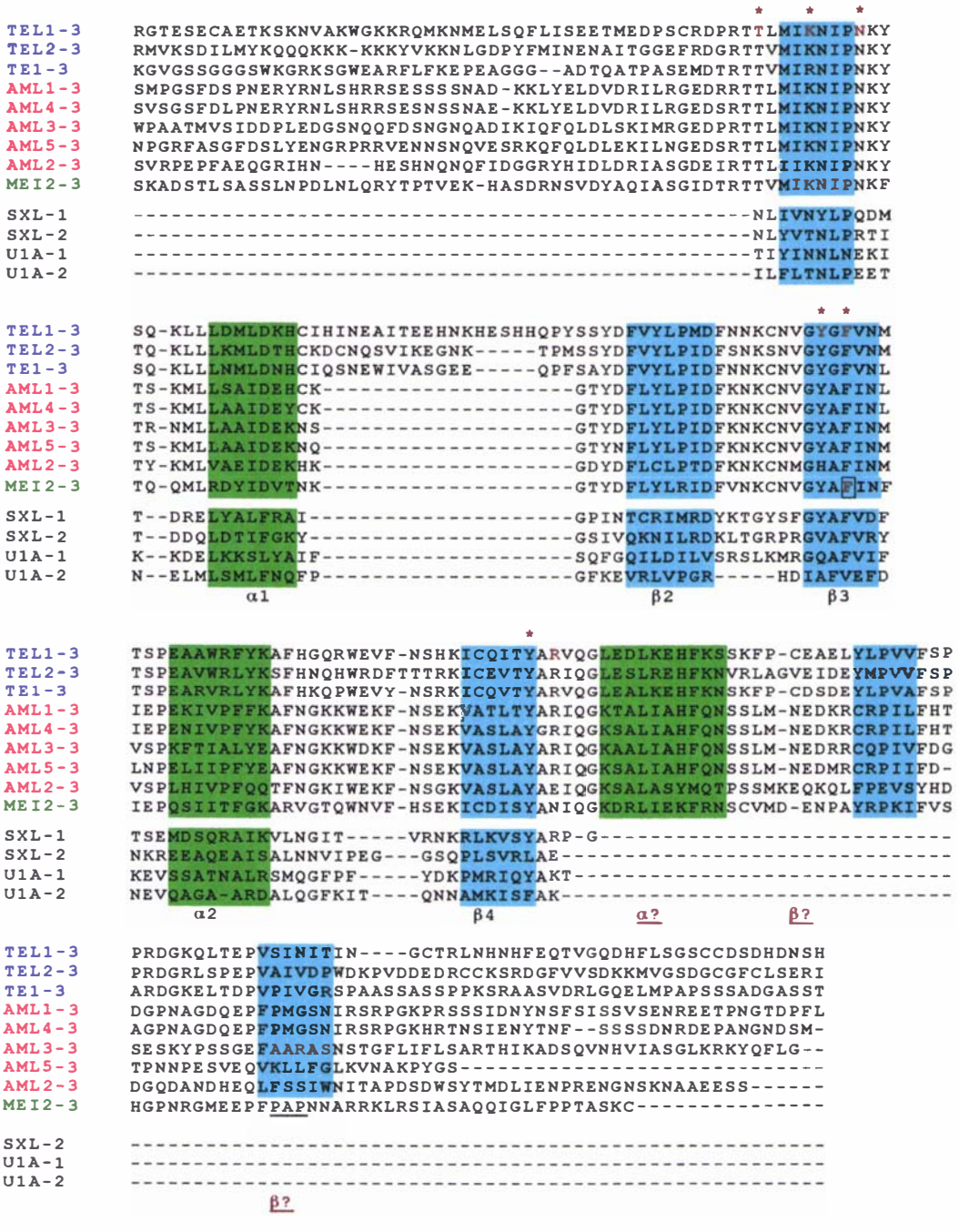
The third RRM of *Mei2*-like genes is predicted to be structurally different from canonical RRM domains

The RRM domain is an ancient structure that is present in all eukaryotes (Birney 1993). A threading comparison of five RRM-domain structures indicates that the RRM fold has been very highly conserved (Figure 1.2, S Moore, unpublished). Since a defining characteristic *Mei2*-like proteins is the high conservation of the third RRM, this region was analysed more closely[9]. Predictions of protein structure were generated by submitting each peptide sequence to the Predict Protein Server (Rost 1996). The output, which consists of predictions of α helix and ß strand forming regions, was used to guide an alignment of *Mei2*-like proteins with the known co-ordinates of α helices and β strands present in *Drosophila* SEX LETHAL (SXL) and human U1A (Figure 3.4). This third domain of *Mei2*-like proteins is very similar in almost every case[10] and is predicted to contain all the elements of canonical RRM proteins and also a C-terminal extension. This region is always predicted to form an α helix and two β strands, structures that are not present in any of the RRM proteins whose structures have been solved (Allain 2000, Conte 2000, Crowder 1999, Deo 1999, Ding 1999, Inoue 1999, Kranz 1999, Nagata 1999, Wang 2001, Xu 1997). *Mei2*-like proteins represent a variant of RRM proteins which may be involved in some novel protein-protein or protein-RNA interactions. The third predicted α helix of *Mei2*-like proteins contains ten residues, 3 1/2 turns of a helix. This helix is uniformly polar on one side, and non polar on another in all *Mei2*-like proteins, and so may be partially imbedded in the protein.

Comparison of the solvent-exposed residues that interact with RNA in the known RRM structures with the structure-based alignments of *Mei2*-like proteins identified some residues that are likely to interact with RNA (Figure 3.4). These residues will be good candidates for mutation studies aimed at characterising the RNA binding activity of plant *Mei2*-like proteins, and generating partial loss of function variants (see Discussion).

[9] This analysis was conducted in collaboration with Stanley Moore.

Figure 3.4. *Mei2*-like proteins contain structural elements not found in other RRM proteins. To assemble this alignment regions of *Mei2*-like proteins that were predicted to form β sheets (blue background M) or α helices (green background L) were aligned with the known SXL and U1A structures. *Mei2*-like proteins all contain an extension that is predicted to form an α helix and two β stands. Residues that are likely to interact with RNA are indicated with an asterix (*), the positions of the F644A point mutation that abolishes RNA binding in MEI2 is indicated (F).



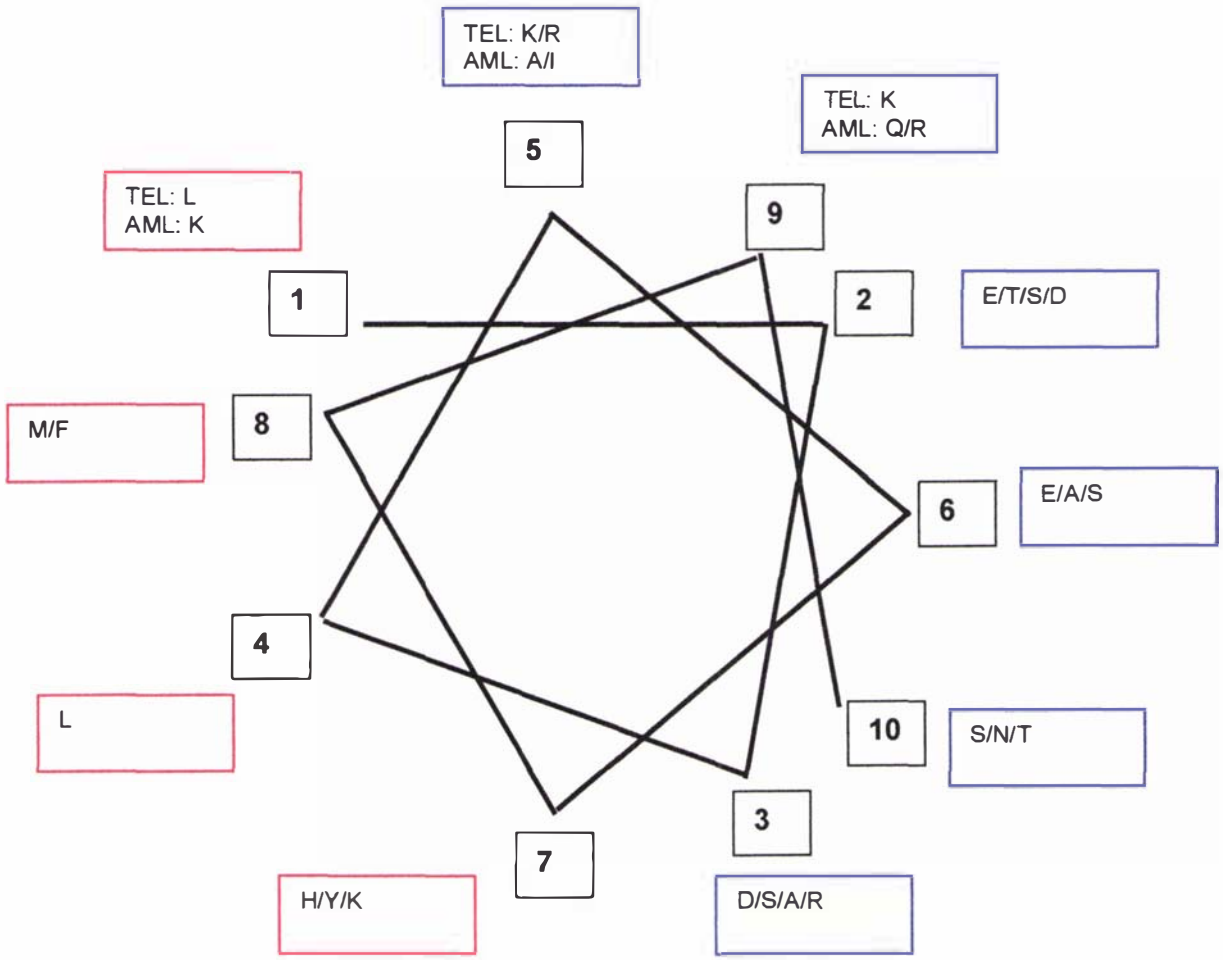


Figure 3.5. The unique C-terminal α helix of RRM3 has polar and non polar sides. A plot of the ten amino acid residues of the predicted α helix. The α carbon chain of α helices diverges by 100° with each residue. This helix is unique to the third RRM of MEI2-like proteins, and has a polar side (red boxes), and a non polar side (blue boxes). Helices of this type are usually exposed at the polar side, and imbedded at the non-polar side. Some residues have slightly different conservation in TEL and AML group proteins.

3.6. The intron positions of plant *Mei2*-like genes are conserved

Comparison of the positions of introns within all compete plant *Mei2*-like genes with respect to the RRMs indicates that there a two groupings of similar intron number and position, the TE1-like genes differ from the AML group (Figure 3.6). All plant *Mei2*-like genes have three introns within the region coding for the third RRM. No significant similarity of nucleotide sequence was detected within these introns (see footnote)[11].

3.7 Comparison of *te1*, *TEL1* and *TEL2* promoters.

In situ expression data indicate that the *TEL1*, *TEL2*, and *te1* transcripts are expressed in . similar, but not identical, domains (Veit 1998; N Alvarez, unpublished data). These sequences were analysed for shared promoter elements using the plant promoter search tool PLACE (Section 3.4). A graph showing the positions of all promoter elements detected by the PLACE search tool (Figure 3.7) allows comparison of promoter element positions without too many assumptions (such as direct conservation of position, equivalent lengths of 5' untranslated regions, overall GC content) that may lose data. The alignment and consensus of all promoter elements produced with the GCG programs *pileup* and *pretty* indicated that there were no regions that were conserved in all three promoters (supplementary file *promoter element alignment*).

The PLACE output is shown graphically in Figure 3.7. The various promoter elements are described in the PLACE website. This data requires further analysis, and promoter element deletion constructs driving the *GUS* gene are being produced in the Veit laborartory at present. There are almost certainly many more promoter elements shown are functional, and careful analysis (beyond the scope of this thesis) combined with functional data will be required. In particular, knowledge of the actual transcription start sites will help determine which regions are promoter *per se*, and which are 5' untranslated regions .

---

[11] Four of the seven *Arabidopsis* Mei2-like genes (57%) have a stop codon in the first position of the conserved intron at the start of the RRM3. However, given the nucleotide frequencies at *Arabidopsis* splice sites, the probability that a codon in this position will be a stop codon is 0.67. Using the GCG programs *pileup* and *pretty* to show a consensus from alignments of this intron did not detect any conserved sequences in these introns (not shown).

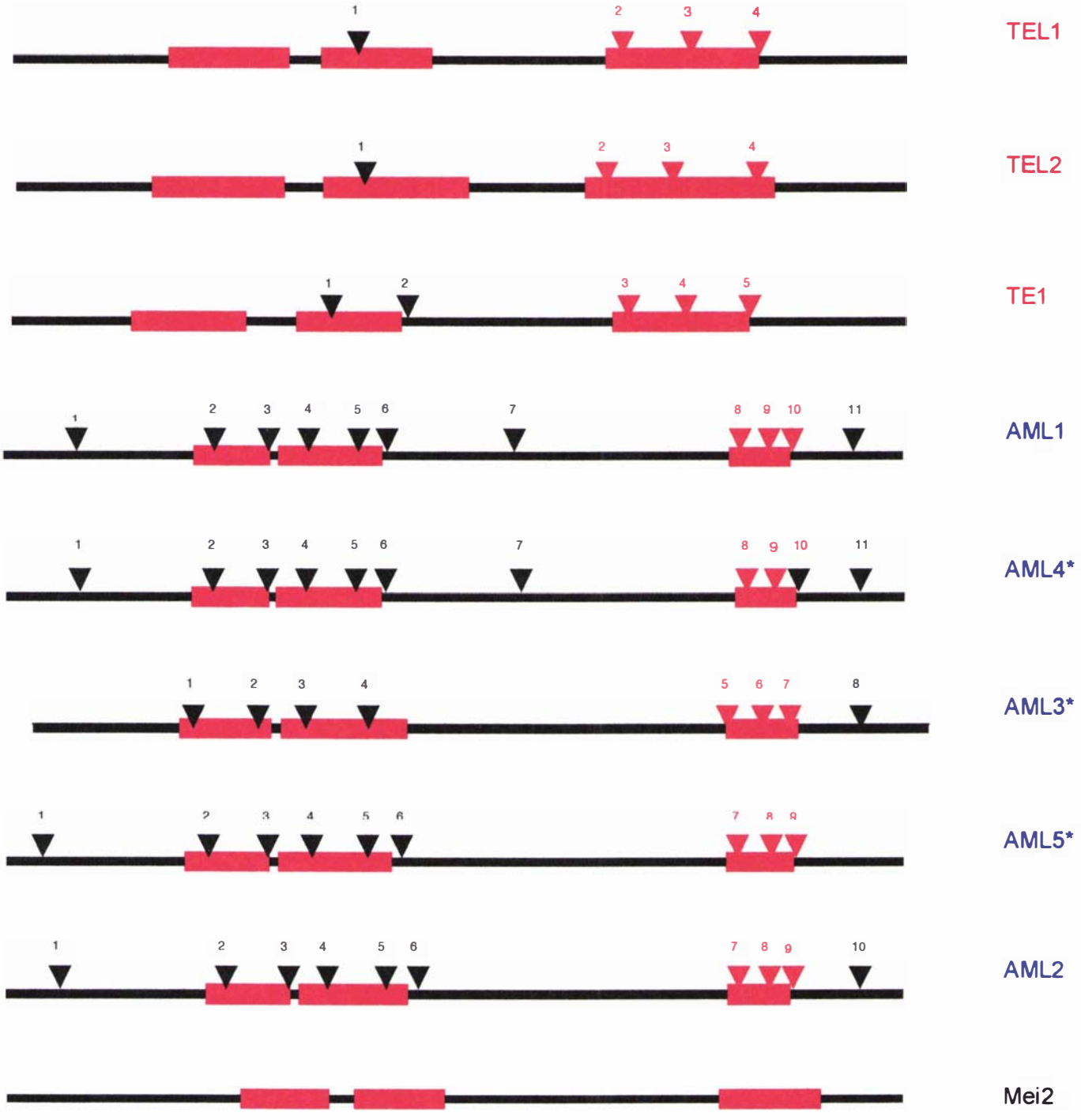


Figure 3.6. Intron positions are conserved in plant Mei2-like genes. The positions of the introns around the third RRM (red triangles) are conserved at the level of amino aid sequence. The lengths of proteins are shown as black lines, with RRMs shown as red blocks, intron positions are indicated with triangles. The TE1-like genes (*te1*, *TEL1*, *TEL2*, in red font) have less introns than the AML group (*AML1-5*, in blue font), *S. pombe Mei2* has no introns. Putative proteins are indicated with an asterix, each predicted protein shares exons structure with at least one cloned Mei2-like gene.

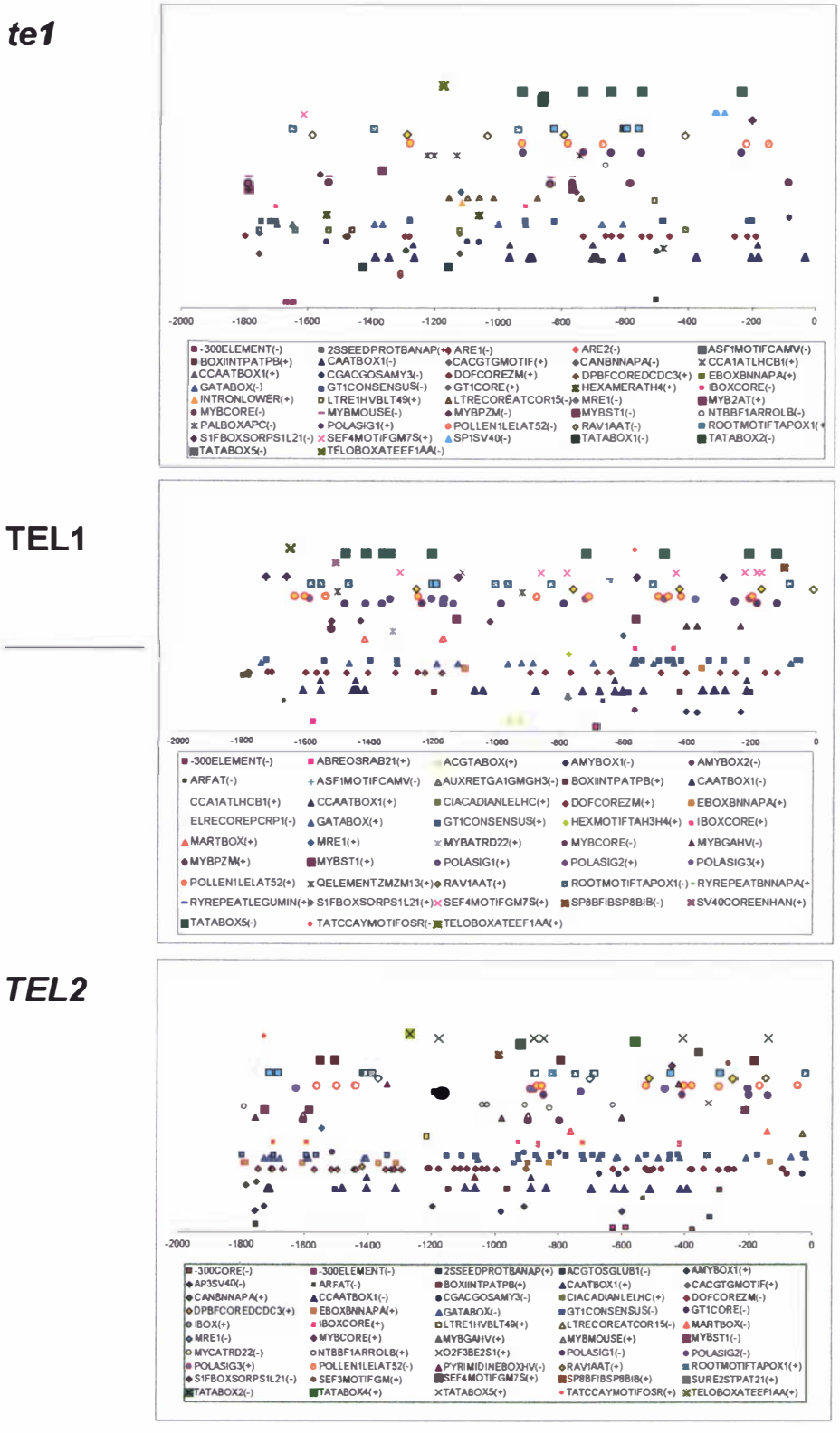


Figure 3.7. Promoter elements present in *te1*, *TEL1 and TEL2* promoter regions. The PLACE server output from equivalent 1799 bp of sequence upstream of the ATG (position zero on graphs) is shown. Clearly, many more promoter elements are predicted by PLACE than are functional. A set of promoter deletion assays are in progress to analyse the functional domains of *TEL* promoters (Trainor, unpublished.)

3.8. A large number of *Mei2*-like genes can be identified in plants.

To identify *Mei2*-like genes from other plant species tBLASTn searches were performed with a consensus[12] of the most highly conserved RRM3 region of the full length *Mei2*-like proteins. Searches against ESTs were the most successful at finding *Mei2*-like genes. The Institute for Genomic Research (TIGR) has compiled ESTs from a variety of sources into searchable, non-redundant complementation groups, called Gene Indices. Gene indices were used in preference to raw genbank ESTs, because they are often longer than ESTs and are closer to a non-redundant gene expression data set. Each gene index was searched by tBLASTn with the RRM3 consensus, and the reading frame was established by eye for any hits with an E value of less than $1x10^{-6}$. Table 3.3 shows the list of genes whose translations produced meaningful alignments with *pileup* (gap weight 6, gap extension weight 1). All other TIGR gene indices (human, mouse, rat, *Drosophila*, Zebrafish, cattle, *C. elegans*, *Trypanosoma brucei*, Leishmania, *Schistomona mansonii*, *Trypanosoma cuzi*, *Plasmodium falciparum, Sacharromyces cerevisiae*, *Brugia malayi, Xenopus laevis, Onchocerca volvulus*), were searched in the same way, only plant indices produced significant hits. An alignment of the plant EST assemblies is shown in supplementary files (file: *gene index alignment*). EST searches with the RRM3 consensus produced a large number of significant hits against various plant species, but never against metazoans.

Genomic fragments of Petunia and rice *Mei2*-like genes have been cloned using degenerate PCR primers (Carmel Gillman, unpublished data). Petunia sequence was sought because there is a T-DNA insertion mutagenesis ('gene machine') project underway (Koes 1995) that could identify a mutant, and rice sequence because it may help to identify a larger region form the rice genome project. tBLASTn searches with these genes indicated that the Petunia gene (referred to here as *PML1*) was most similar to *te1* (E value 10^{-25}), and to *TEL2* (E value 10^{-19}). The Rice gene *Cgrg2*, was very similar to *te1*, and has sequence elements indicating that it is a member of the *te1*-like group rather than the *AML*-like group (not shown).

[12] The consensus used was derived from the alignment shown in Figure 3.2, positions 873 to 1001, with a minimum consensus plurality of 2. The consensus is:
EDXRTTLMIKNIPNKYTQKMLLAAIDEHCKXXNEXIXXEGNKXXXXXQPXGTYDFLYLPIDFKNKCNVGY
AFINMXSPEAIVPFYKAFNGKKWEKFXNSEKVASLXYARIQGKXALIXHFQNSSLMNCE

One putative *Mei2*-like gene of particular interest is a rice genomic clone (accession AC073405), that has all the signatures of a *Mei2*-like gene, and appears to contain sequence of the entire gene. The cDNA, and protein sequence have not yet been predicted. tBLASTn searches indicated that the RRM3 of this rice clone is at ~130,000 nt into accession AC073405.

Table 3.3 Mei2-like genes from plants

Species	Gene[1]	Type	E value	*Tissue source(s)*
Soy	TC21873	TIGR EST assembly	7x10-43	Hypocotyl and plumule of germinating seeds, mature flowers, young cotyledons, whole seedlings without cotyledons
	AI460437	Single EST	1x10-41	Roots
	BE058737	Single EST	$3x10^{-13}$	Immature flowers
	AW705893	Single EST	$5x10^{-06}$	Immature seed coats
Rice	TC30515	TIGR EST assembly	1.5x10-43	Etiolated shoot, callus, panicle at flowering stage, immature seed
	CGRG2	Genomic PCR product	-	nuclear DNA, tissue unknown
	AC073405	Genomic clone	6x10-14	nuclear DNA, tissue unknown
Tomato	TC44969	TIGR EST assembly	1x10-38	Callus, breaker fruit, mature green fruit
	TC44218	TIGR EST assembly	4x10-29	Pre-anthesis buds
	AW621483	Single EST	9x10-19	Roots
	TC36824	TIGR EST assembly	3x10-14	Mixed elicitor, flower 0-3 mm buds
Maize	TC45080	TIGR EST assembly	9x10-44	Roots, anther/pollen
	TC37862	TIGR EST assembly	3x10-33	Roots, anther/pollen
	TC37987	TIGR EST assembly	7x10-26	Leaf primordia, tassel
	TC39811	TIGR EST assembly	1x10-08	Glume, early embryo
Potato	Nil			
Medicago	AL386617	Single EST	4.6x10-43	Roots
	AW257211	Single EST	9.2x10-36	Roots
	TC7081	TIGR EST assembly	1.4x10-23	Roots
	TC15332	TIGR EST assembly	8.9x10-20	Roots
Sorghum	TC606	TIGR EST assembly	4.7x10-14	Anthracnose-infected leaves
Wheat	BE415912	Single EST	6.5x10-44	Young spikelets
	BE405924	Single EST	1.6x10-17	Roots
Petunia	*PML1*	Genomic PCR product	-	nuclear DNA, tissue unknown

[1] Gene names are either TIGR gene index names (for EST assemblies), names of degenerate PCR clones(*), or genbank accession numbers. A prediction for the exon positions of *Cgrg2* was made with Genscan, using the maize parameters.

3.9. *Mei2*-like genes from taxa other than plants

The most sensitive BLAST search for plant sequences used the RRM3 consensus as a query against ESTs. Using these criteria to search against fungi, animals, and all ESTs found only one significant hit other than those of plants and *Schizosaccharomyces pombe*: an EST from *Pneumocystis carinii* with an E value of $6x10^{-19}$ (accession AW332476). *P. carinii* is an ascomycete fungus that is related to *Schizosaccharomyces pombe.* An alignment of this sequence with all other complete *Mei2*-like protein sequences indicated that this is certainly a *Mei2*-like gene (supplementary file: *Pneumocystis alignment*). A tBLASTn search with the RRM3 consensus showed one relatively insignificant hit from the incomplete *Neurospora* genome project database (E value 0.001) (see footnote [13]).

3.10. Phylogenetic analysis *Mei2*-like genes

3.10a. Phylogenetic analysis of complete *Mei2*-like genes.

The alignment of all complete *Mei2*-like protein sequences (TE1, TEL1-2, AML1-5) was adjusted manually for mis-aligned regions, and used to infer the relationships between these genes using maximum parsimony, maximum likelihood and neighbour joining methods with the package PAUP*. Gaps, or regions that did not contain a convincing alignment of all genes, were excluded from the analysis. Standard-weighted parsimony trees indicated strong support for three main clades,

- the *terminal ear*-like group (*te1*, *TEL1*, *TEL2*) (TEL group),
- the *AML1*-like group (*AML1-5*) (AML group)
- *Mei2*

With a Goloboff correction of 2, which down weights less conservative sites that may have been saturated, these clades remained unchanged. *Mei2* was consistently positioned as an outgroup. When functional groups S/T and I/L/V were coded together, which could correct for non-phylogenetic bias, such as GC content, all the clades remained (Figure 3.8).

[13] *Neurospora crassa* sequence contig 1.1011.

These groups are consistent with the two types of exon structure (Figure 3.6), and the presence of several insertions/deletions in TE1 and TEL1/2 protein sequences (Figure 3.), even when such insertions are not included in data used to build the tree.

3.10b. Phylogenetic analysis of incomplete *Mei2*-like gene sequences and ESTs.

The protein sequences from all putative and cloned *Mei2*-like genes including ESTs were aligned manually. For many ESTs this includes only the region around the third RRM. Those ESTs that were very short, and would limit the data set available for tree-building were excluded. Using this larger set of *Mei2*-like protein sequences, including TIGR gene index data, and the *Pneumocystis* gene (PcAW332476) confirmed the TEL and AML clades. All ESTs clustered consistently with *Arabidopsis* AML group genes (Figure 3.9). A common practise in developmental biology is to tentatively identify orthologous proteins (those that have the same function) by close sequence similarity. No clear maize, rice, or wheat orthologs of the *te1* gene were identified in this way.

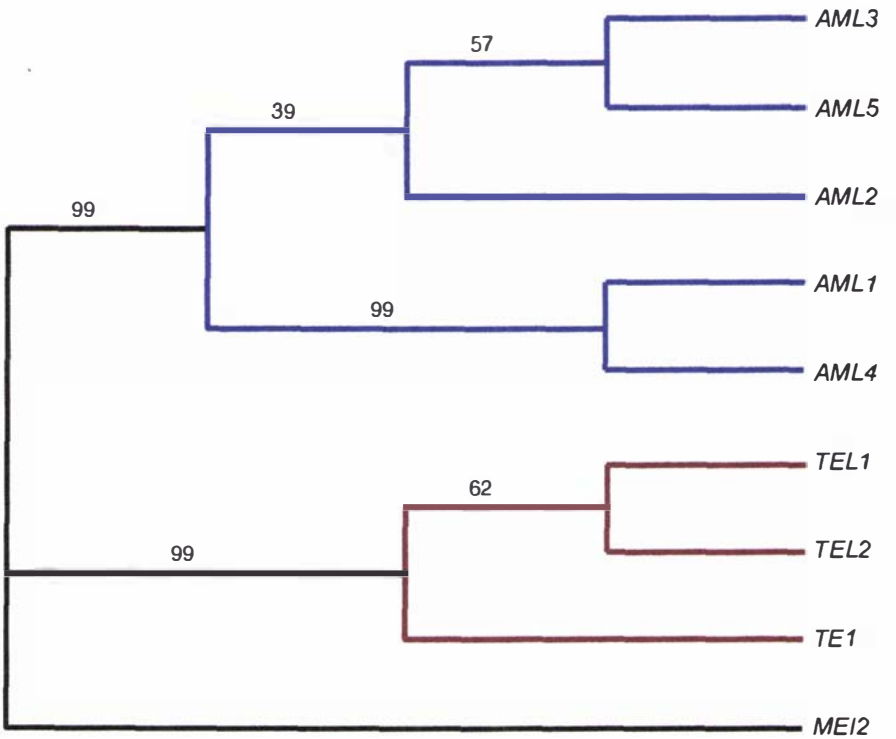


Figure 3.8. Three groups of Mei2-like genes have strong support in maximum parsimony trees. A maximum parsimony tree showing the relationships of all complete Mei2-like genes. The TEL group (TE1, TEL1, TEL2), and the AML group are from plants, MEI2 is from Schizosaccharomyces pombe. AML1 and AML4 are very similar and always group together. In this analysis the functional groups S=T, I=V=L were coded together, and a Goloboff correction of 2 was used to downweight rapidly changing sites in the alignment. Numbers indicate support for clades from 1000 bootstrap replicates.

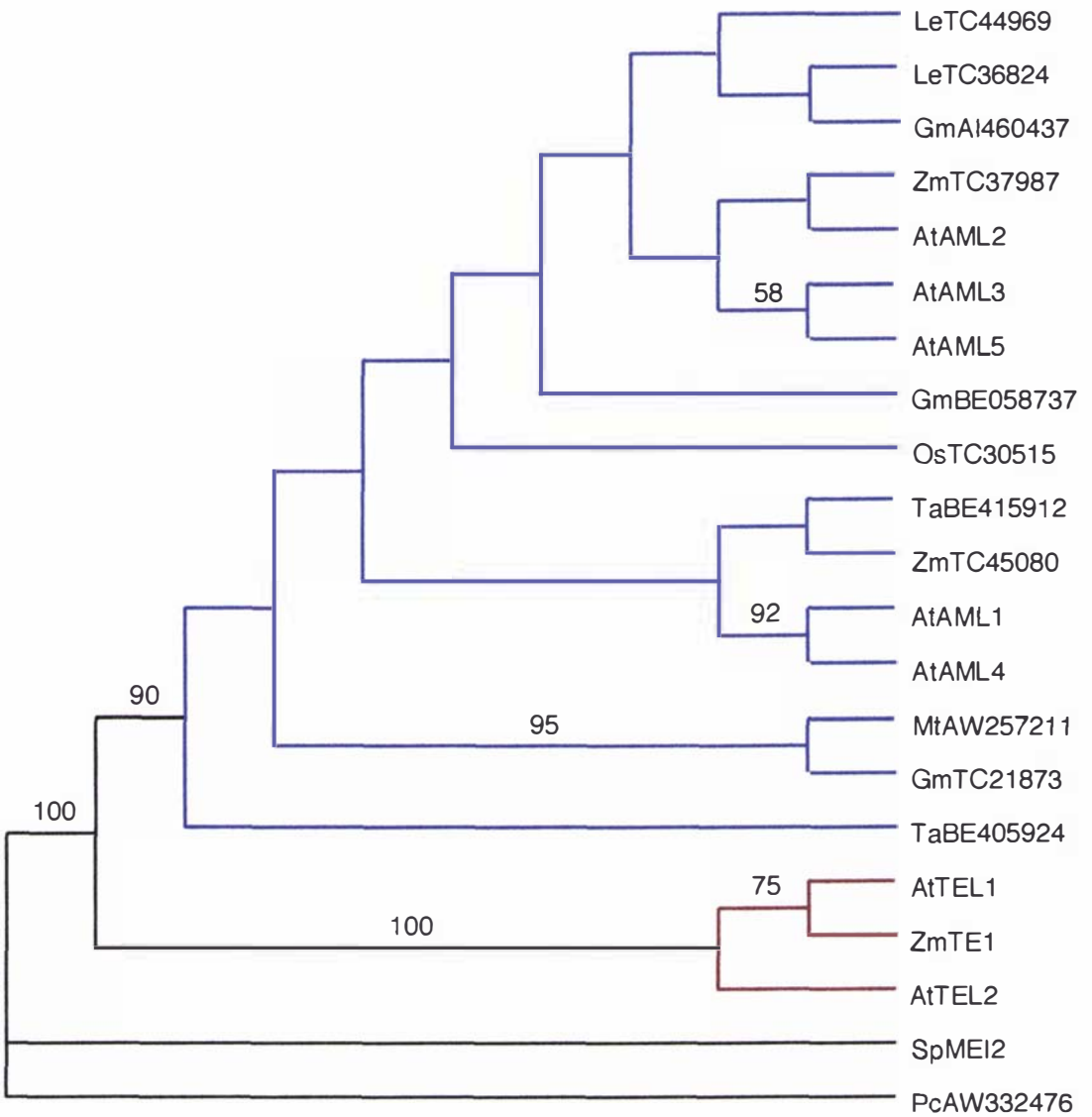


Figure 3.9. All *Mei2*-like ESTs identified group with the AML type. Trees were built from an alignment of all ESTs that were of sufficient length, and adjusted manually. The data set included 218 characters, gaps were included, coded as no data. Three main clades were consistent regardless of tree building parameters, the Ascomycota group (S. pombe *Mei2* and the *P. carinii* EST PcAW332476), the AML group (blue branches), and the TEL group (red branches). This particular tree was constructed with parsimony, bootstrap values from 1000 replicates that were above 50% are indicated. Each protein is given a prefix indicating the species source; Le *Lycopersicum esculentum*, Gm *Glycine max*, Zm *Zea mays*, At *Arabidopsis thaliana*, Os *Oryza sativa*, Ta *Triticum aestivum*, Sp *Schizosaccharomyces pombe*, Pc *Pneumocystis carinii.*

DISCUSSION (3)

3.11. Function and biochemistry of MEI2 with respect to plant *Mei2*-like genes

The *S. pombe Mei2* gene has been intensively studied (Watanabe 1988; Watanabe 1994; Watanabe 1997; Yamashita 1998; Shinozaki-Yabana 2000), its gene product is the only *Mei2*-like protein for which there is any biochemical data. Because there has been considerable evolutionary time since fungal and plant *Mei2*-like gene diverged and the genes function in very different biological contexts, propositions about plant *Mei2*-like gene function and activity based on *Mei2* must be analysed very carefully. A discussion of the current understanding of *Mei2* function, and the evidence that each particular aspect of function is applicable to plant *Mei2*-like genes follows.

3.11a. The biological context of *Mei2* and plant *Mei2*-like genes

The *Schizosaccharomyces* pombe *Mei2* gene product is crucial for commitment to meiosis, but is not required for normal mitotic growth. The gene is expressed at a low level in mitotic cells and higher levels of transcription are induced by the STE11 transcription factor in conditions of nutrient starvation in response to low intracellular levels of cAMP (Watanabe 1988). Investigation of a temperature sensitive allele, *mei2-33*, indicated that the gene is required for two stages of meiosis (Watanabe 1994). A diploid strain carrying *mei2-33* was meiosis-competent at 25°C, and deficient at 37°C. At the semi-restrictive temperature of 31.5°C the *mei2-33* strain was able to perform premeiotic DNA synthesis, but was unable to carry out subsequent meiotic divisions (meiosis I). Strains lacking *Mei2* entirely (*mei2Δ*) do not carry out premeiotic DNA synthesis or meiosis I.

At present there is no strong evidence to suggest that any particular plant *Mei2*-like gene regulates the commitment to meiosis. Although the maize *te1* transcript is expressed in ears and tassels, (Section 2.12), all the mutant *te1* alleles produce viable gametes, including plants with partial transformation of tassels into ears. The *Arabidopsis AML1* gene was cloned by its ability to complement an *S. pombe* steroid

receptor mutant[14] that is meiosis deficient and acts upstream of *mei2* (Hirayama 1997). Although expression of the *AML1* cDNA from the constitutive *S. pombe adh* promoter allowed steroid receptor mutant to sporulate with low efficiency, it could not rescue *mei2Δ* strains. There is currently no supporting evidence for *AML1* functioning in meiosis in *Arabidopsis*. Of those *Arabidopsis* genes whose expression domains are known by *in situ* hybridisation (*AML1*, *TEL1* and *TEL2*), only *TEL1* is expressed in meiotic tissues (locules enclosing the pollen sac); preliminary genetic evidence suggests, however, that *TEL1* T-DNA knockouts produce viable pollen (N Alvarez, unpublished data).

3.11b. RNA dependent localisation

MEI2 protein function is dependant on an untranslated RNA species, meiRNA. The meiRNA gene (*sme2*) was identified as a high copy number suppresser of the *mei2-33* mutant, subsequent deletion of the *sme2* locus resulted in strains that could not sporulate (Watanabe 1994). The 0.5 kb *sme2* transcript was induced by nitrogen starvation, was polyadenylated, and contained several short open reading frames (ORFs). The *sme2* transcript appears to act as an RNA since disruption any of the ORFs in the *sme2* gene did not affect meiosis, and the MEI2 protein was shown to bind the meiRNA *in vitro* and to form a complex *in vivo* (Watanabe 1994).

The meiRNA is required for the localisation of MEI2 to a specific point in the nucleus during meiosis (Yamashita 1988). This '*Mei2* dot' was a small point of MEI2 accumulation that formed in a position between the centromeres and the spindle pole body after the onset of meiosis. GFP:MEI2 fusion constructs remained predominantly cytoplasmic in strains lacking the meiRNA. It appears that the meiRNA-MEI2 complex itself is required for correct localisation since GFP:MEI2-F644A derivatives, carrying the mutation in RRM3 that abolished meiRNA binding *in vitro*, were not correctly localised. Also, fluorescent *in situ* hybridisation studies showed the meiRNA to be co-localised with the MEI2 protein in the *Mei2* dot. The localisation of MEI2 by association with the meiRNA is developmentally controlled, and very precise; the *Mei2* dot is much smaller than the nucleus, and does not appear to correspond to the nucleolus (Yamashita 1998). When GFP:MEI2:NLS fusions were localised to the nucleus they were distributed more evenly throughout the nucleus. The nuclear import apparatus that produces the RNA dependant localisation of MEI2 appears to

[14] *AML1* complimented a double mutant for *S. pombe map3* and *mam2*, two types of pheromone

have been conserved throughout evolution, since the MEI2 protein shows meiRNA-dependant localisation in mammalian cells (Yamashita 1998). This is surprising since no *Mei2*-like genes have been found in metazoans, and could be interpreted to mean that the MEI2-meiRNA complex contains one example of this type of nuclear localisation 'signal', others that cannot be recognised by sequence similarity may still function in mammalian cells.

While there is indirect evidence that plant *Mei2*-like gene bind RNA (see below), there is no information at present about the RNA target(s). It seems likely that *Arabidopsis* AML1 was able to bind the *S. pombe* meiRNA in complementation experiments (Hirayama 1997), but no published studies have shown AML1 binding to *Arabidopsis* RNA. FASTA searches of *Arabidopsis* ESTs with the meiRNA did not identify sequences in *Arabidopsis* that were very similar to the meiRNA (see footnote[15]). Such a search is by no means conclusive, if *Arabidopsis Mei2*-like proteins do have a meiRNA analogue, the RNA may be very difficult to locate by sequence. Since the MEI2 protein shows meiRNA-dependant localisation in mammalian cells it seems likely that this mechanism will also occur in plant cells. This hypothesis is currently being tested in the Veit laboratory by transient expression of MEI2 and meiRNA in plant cells.

Both maize *te1* and *Arabidopsis TEL2* were expressed in *Arabidopsis* with nuclear localisation tags (this study). In both cases there was no evidence that nuclear localisation is required for function (Chapter 5, see discussion Section 5.13). Because of the possible redundancy of the *Arabidopsis* genes TEL1 and TEL2, and the fact that *te1* is not a native *Arabidopsis* gene, these experiments are not entirely conclusive. It remains to be seen if any of the plant *Mei2*-like proteins are nuclear localised without synthetic nuclear localisation tags. Transgenic *Arabidopsis* plants expressing GFP:TEL2 fusions were being produced in the Veit laboratory while this thesis was being written.

---

receptors, the M-factor and the P-factor receptors (Hirayama 1997).

[15] The best FASTA search hits with meiRNA against *Arabidopsis* ESTs were BE039135 (E value 0.22) which is similar to homogentisate 1:2-dioxygenase mRNA sequence, and to two other ESTs wih unknown function, AA007118 (E value 0.68) and AA007118 (E value 0.7).

3.11c. RNA binding activity of MEI2 and plant *Mei2*-like genes

One of the most intriguing features of MEI2 is that the RNA binding activities appear to be required at several stages during meiosis, and that there is probably more than one RNA target. Analysis of the temperature sensitive *mei2-33* allele indicated that MEI2 is required for premeiotic DNA synthesis (Watanabe 1994). This allele performs premeiotic DNA synthesis but not meiosis I at a semirestrictive temperature and is the result of a point mutation that converts phenylalanine 240 to leucine (F240L). This mutation, which is within the RNP-1 octamer of the N-terminal RRM (RRM1), reduced, but did not abolish, the ability of MEI2 to bind the meiRNA *in vitro* (Watanabe 1994). It seems that the RNA binding of the C-terminal RRM (RRM3) is required for premeiotic DNA synthesis as well. A site directed mutation in a conserved residue of the RRM3 (F644A) completely abolished meiRNA binding *in vitro* was not able to complement the meiosis deficiency of *mei2Δ* strains. Strains with this mutation are not able to undergo premeiotic DNA synthesis. Taken together, this evidence suggests that both RRM1 and RRM3 are required for MEI2 to promote premeiotic DNA synthesis.

Secondly, GFP fusion with the MEI2 protein carrying a mutation in RRM3 that abolished RNA binding (GFP:MEI2-F644A) were not localised to the *Mei2* dot, indicating that meiRNA-MEI2 complex transport is another activity of MEI2 that requires RNA binding, separation of these two phenotypes in the *mei2-33* mutant indicates that these two functions require different conditions, perhaps only more stringent binding of the meiRNA[16].

Finally, the RNA binding ability of MEI2 is required in the nucleus after meiRNA-dependant nuclear localisation. The MEI2 protein is not correctly localised in meiRNA deficient (*sme2Δ*) strains, but when the MEI2 protein is localised to the nucleus by fusing it to the SV40 nuclear localisation signal (MEI2:NLS) it promotes meiosis in the absence of meiRNA. However, when MEI2-F644A protein, which has a mutation in a critical residue for RNA binding, is nuclear localised by the SV40 NLS it fails to promote meiosis (Yamashita 1988). Hence, RNA binding in the nucleus appears to be required, but the meiRNA target appears to be dispensable. It seems that some other RNA is required, although the meiRNA does co-localise with the MEI2 protein the *Mei2* dot in wild type cells. At present there is no information about the other RNA

target(s) of MEI2 in the nucleus. While the position of the *Mei2* dot is very precise, and constant with respect to the centromeres and the spindle pole body, it has not been conclusively associated with any particular RNA/protein complex.

From the conservation of RRMs in plant *Mei2*-like genes, it is clear that they bind RNA. Two mutations in *S. pombe* MEI2 have been shown to affect RNA binding. One in the third RRM (F644A), abolishes RNA binding, and this position is identical in all *Mei2*-like genes (Figure 3.5). Another mutation in RRM1 (F240L) decreased the affinity of MEI2 for the meiRNA, and tyrosine is the consensus for this position in *Mei2*-like genes. Studies of deletion derivatives of MEI2 indicated that only RRM3 is required for function (Watanabe 1997). This region is the most highly conserved in *Mei2*-like genes, with several regions of identity that extends beyond the canonical RRM structure, so indications are that the this C-terminal motif is crucial for function.

3.11d. Phosphorylation of *Mei2*-like proteins

The activity of the MEI2 protein is negatively regulated by phosphorylation. When *mei2* was expressed in *Pat1Δ* strains they initiated meiosis, but this activity was repressed in *Pat1*$^{+}$ strains. The PAT1 serine/threonine kinase was shown to phosphorylate MEI2 *in vitro*, and to form a complex with PAT1 *in vivo* (Watanabe 1997). The two residues of MEI2 that were primarily phosphorylated Ser438 and Thr527, which are between RRM2 and RRM3, were identifed by tryptic phosphopeptide mapping of *in vivo* phosphorylated MEi2, and confirmed by production of a protein containing alanine at these positions (MEI2-SATA) (Watanabe 1997). Interestingly ultraviolet crosslinking experiments indicated that phosphorylation does not significantly affect the RNA-binding ability of MEI2. The *Pat1* kinase gene is inhibited by direct association with MEI3, which acts as a pseudosubstrate. The combination of upregulated *Mei3* expression, *Mei2* and *sme2* (meiRNA) expression all contribute to the level of unphosphorylated, and localised MEI2.

At present, there is no evidence that plant *Mei2*-like genes are phosphorylated. The amino acid sequence around the residues that are phosphorylated in MEI2, Ser438 and Thre527 is not highly conserved in plant *Mei2*-like genes (Figure 3.3). NetPhos (Blom 1999), a neural network-based method for predicting potential phosphorylation sites in protein sequences, predicts a large number of serine/threonine kinase sites in

[16] The protein product of the *mei2-33* temperature senstive allele showed reduced binding to the

TE1, TEL1 and TEL2 protein sequences (28, 20 and 7 sites respectively) (data not shown). Without biochemical studies, or a more detailed understanding of which residues are in conserved positions on the surface of the proteins it is difficult to interpret this data.

3.11e. Interaction with MIP1 WD repeat protein

The *S. pombe* Mip1 gene was identified as interacting with *Mei2* in a screen for high copy number suppressors of the MEI2 derivative MEI2-SATA, which cannot be phosphorylated by PAT1 kinase (Shinozaki-Yabana 2000). Mip1 encodes a WD-repeat protein, which interacts in the cytoplasm with MEI2 protein, only in unphosphorylated form. Genetic evidence indicated that Mip1 is required for MEI2 function. Mip1 also appeared to be involved with conjugation and to associate with other proteins *in vivo*.

Proteins similar to Mip1 have been found by genome projects in other eukaryotes including *Arabidopsis*[17], but their functions have not been studied. The authors note that since *S. pombe* MIP1 is required for mitotic as well as meiotic growth, and interacts transiently with MEI2, it may act as a kind of molecular chaperone.

3.12. Evolution and derivation of *Mei2*-like genes

Phylogenetic analysis of *Mei2*-like genes indicated that there were three main groups of genes, *terminal ear1*-like (TEL group), *AML1*-like (AML group) and the Ascomycota group from *Schizosaccharomyces pombe* and *Pneumocystis carinii* (Section 3.10). Neighbour joining trees indicated that the distances between these groups were about equal, that is, both AML and TEL type genes are equally distant from the ascomycete genes (not shown). The two plant groups appear to have undergone a considerable degree of evolutionary change, and may function via different biochemical mechanisms, and in different biological contexts. It is clear also that some recent divergences in plant *Mei2*-like genes have occurred, most notably AML1/AML4, which always group together.

---

meiRNA than the wild type protein at 35°C, but affinity was unaltered at 25°C (Watanabe 1994).

[17] The *Arabidopsis* Mip1-like protein accession is T16011.22.

The question of whether *Arabidopsis TEL1* or *TEL2*, or both are functional orthologs of the maize *te1* gene has not been answered by this analysis. Neither the Kishino-Hasegawa test (Kishino 1989), nor the Templeton (Wilcoxon signed-ranks) and winning-sites test (Templeton 1983) indicated statistical support for any particular grouping of *te1, TEL1* or *TEL2* (not shown).

3.13. Functions of plant *Mei2*-like genes

The mutant phenotype, and *in situ* expression pattern of *te1* led to specific models of gene function (Veit 1998). The accumulation of *te1* gene products were thought to influence the precise positioning of leaf primordia, and on a molecular level, to act as a molecular switch between leaf and internode fates. The semicircular ring expression pattern shown in maize SAMs has now been found in embryo tissue (N Alvarez, unpublished data). The *te1* transcript has been detected by RT-PCR in early tassel and ear tissues at approximately the same level as SAM tissue (this study, Section 2.12). It is conceivable that *te1* function is very similar in these tissues.

Several sources of data suggest that TEL1 and TEL2 perform redundant functions in *Arabidopsis* that overlap with the function of *te1 in* maize. Homozygous T-DNA insertion mutants of *TEL1* show no clear phenotype (Carmel Gilman, unpublished data), compared to the many *te1* mutant alleles that show a clear mutant phenotype. Statistical tests show that, by sequence, both TEL1 and TEL2 are equally similar to *te1* (Section 3.12). The *in situ* expression domains of *TEL1* and *TEL2* provide more evidence that these *Arabidopsis* TEL group genes have slightly different functions, but that they may overlap (N Alvarez, unpublished data). *TEL2* appears most similar to *te1* in that it is expressed in regions of the SAM that are not recruited into leaf primordia, in the case of *Arabidopsis* this is an the central zone of vegetative and inflorescence SAMs, *TEL1* is expressed throughout vegetative SAMs. *TEL1* is also expressed in the heart shaped embryo SAM, cotyledons and root meristem. Because of the apparent overlap in function, it is likely that *TEL1/TEL2* double mutants will be required before any mutant phenotype is clear.

The functions of AML group genes is still less clear. These genes have been isolated from a variety of plant tissues (see Tables 3.2 and 3.3), the single published analysis of *AML1* (Hirayama 1997) indicated that the transcript was present in approximately equal quantities in siliques, roots, leaves shoots and flowers.

3.14. Concluding remarks

This study has identified a new family of RNA binding proteins in plants. Our understanding of plant *Mei2*-like gene function is just beginning. While *in situ* data and mutant analysis indicate that the *TEL* group gene function in a developmental context, even the broad functions of the *AML* group genes are not known. While there are sequence similarities between MEI2 and the plant *Mei2*-like proteins, it seems that plant genes have been recruited to a variety of developmental functions

CHAPTER 4. EXPRESSION OF THE TE1 PROTEIN IN HETEROLOGOUS HOSTS AND PRODUCTION OF ANTIBODIES

4.1 ABSTRACT

Three *E. coli* expression systems and a *Pichia pastoris* system were used in attempts to produce sufficient quantities of purified full length TE1 protein for biochemical studies. Only the amino-terminal half of the protein was efficiently expressed. Polyclonal antibodies raised against this amino terminal portion produced as a fusion with the maltose binding protein (MBP:TE $^{1-286}$) were shown to have activity against both the MBP and TE $^{1-286}$ peptides. These antibodies were affinity purified against the MBP:TE $^{1-286}$ fusion protein, but this sera showed cross reactivity to other protein from maize extracts that precluded its use for experiments requiring specific detection of the TE1 protein. Since the C-terminal RRM, which could not be produced in significant quantities, is probably a requirement for full protein function *in vitro* protein studies were not attempted. Possible reasons for the difficulties in expressing the protein are discussed, including the effects of rare codons upon translation fidelity.

METHODS (4)

4.2 General protein methods

4.2a. Bradford assay estimation of protein concentration

Estimation of total protein concentrations were performed in microtitre plates in a total volume of 200 µl per well. 40 µl of Bradford reagent (Biorad) was mixed with an appropriate dilution of the protein extract in 160 µl volume. The reagent was allowed to react for 5 minutes at room temperature, and absorbance read at 595 nm on an Anthos HTII microtitre plate reader (Anthos). Values were compared to a Bovine Serum Albumin (BSA) standard curve prepared with 0.01, 0.02, 0.03, 0.04 and 0.05 mg/ml BSA.

4.2b. SDS polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was carried out as described in Sambrook (1989). Typically 10% polyacrylamide resolution gels (10% acrylamide mix, 375 mM Tris-HCl pH 8.8, 0.1% SDS, 0.1% ammonium persulphate, $4x10^{-4}$% TEMED) were prepared with 5% stacking gels (5% acrylamide mix, 125 mM Tris-HCl pH 6.8, 0.1% SDS, 0.1% ammonium persulphate, $1x10^{-3}$% TEMED). Acrylamide mix consists of 29% acrylamide, and 1% *N, N'*-methylenebisacylamide. Gels were poured in a vertical electrophoresis apparatus (Biorad Mini-Protean® II Cell).

Protein extracts were boiled for 5 minutes in SDS loading buffer (50 mM Tris pH 6.8, 2% SDS, 0.2% bromophenol blue, 20% (v/v) glycerol, 70 mM 2-ME), then centrifuged for 5 minutes at maximum speed at room temperature in a benchtop microfuge. Equal volumes of samples were loaded, typically 10 µl. Molecular weight markers, either Biorad Broad Range Molecular Weight Markers, or Pharmacia Low Molecular Weight Calibration Kit, were always run in the left lane of the gel. Biorad Broad Range Biotinylated Molecular Weight Markers were used if the gel was to be used for western blotting (Sections 4.2c and 4.2d).

Proteins were electrophoresed in Tris-glycine electrophoresis buffer (25 mM Tris-HCl, 250 mM glycine pH 8.3, 0.1% SDS) at 100 V until the loading dye front was close to the

bottom of the gel. Gels were either stained for 30 minutes with Coomassie stain (5 g of Coomassie Brilliant Blue in 90 ml of methanol:water (1:1), and 10 ml of glacial acetic acid) and destained in destain solution (90 ml of methanol:water (1:1), and 10 ml of glacial acetic acid), or transferred to PVDF for western blotting (Sections 4.2c and 4.2d). When gels were to be blotted, 10ng of MBP:TE1$^{1-286}$ was loaded to indicate the detection limit of the procedure, and broad range biotinylated molecular weight markers (Biorad), or Kaleidoscope pre-stained markers (Biorad) were run in the leftmost lane of the gel.

4.2c. Western blotting, luminescent substrate

A variety of primary antibodies were used for western blotting (see specific blot figure legends for actual dilutions of antibodies). In general, unpurified sera was bound at 1:200 to 1:800 dilution, while affinity purified sera was bound at 1:6 000 dilution. The appropriate dilution of unpurified sera was determined by diluting antibodies 100-fold less for western blotting than the optimum signal-to-noise ratio for ELISA. The optimum dilution for the affinity purified anti-MBP:TE1$^{1-286}$ sera was determined empirically, with the expectation that the highest effective dilution would produce a single band. To detect MBP portions of fusion proteins, an anti-MBP antibody (NEB) was bound at a dilution of 1:10 000.

Protein extracts were electrophoresed in 10% SDS-PAGE as described in Section 4.2b. A blotting stack was set up between the electrodes of a Biorad electro-blotting apparatus and protein was transferred to the PVDF membrane (Millipore) in electroblot buffer (200 mM glycine, 25 mM Tris-HCl, 10% methanol, 0.03% SDS) with 200 mA of current overnight at 4°C. Blotting stacks for protein transfer to PVDF were, from the negative electrode upwards; a coarse sponge approximately 1 cm wider and longer than the gel, a sheet of Whatman 3MM blotting paper of the same dimensions, the polyacrylamide gel, a PVDF membrane of the same dimensions that had been wet with methanol, a sheet of Whatman 3MM blotting paper, another coarse sponge.

To suppress non-specific antibody binding, membranes were incubated in 1.0% I-Block (Tropix), 0.5 % Tween 20 (BDH) in phosphate buffered saline (PBS: 270 mM NaCl, 400 mM Na_2HPO_4, 30 mM KH_2PO_4, pH 7.4) for 1 hour at room temperature. A dilution of the

primary antibody was bound in 1.0% I-Block, 0.5 % Tween 20 in PBS Affinity purified anti- MBP:TE1$^{1-286}$primary antibody was bound at a dilution of 1:6 000.

After binding of primary antibody, membranes were washed 3 times for 15 minutes in 0.2% I-Block, 0.5 % Tween 20 in PBS. Secondary antibodies horseradish peroxidase (HRP) conjugated goat anti-rabbit IgG (Sigma) (1:20 000 dilution) and avidin-HRP conjugate (Biorad) (1:6 000) were bound in 1.0% I-Block, 0.5 % Tween 20 in PBS for an hour at room temperature. Avidin-HRP was included to detect the biotinylated molecular weight markers. Membranes were then washed 3 times in 0.2% I-Block, 0.5 % Tween 20 in PBS as above.

Membranes were washed briefly in PBS, prior to substrate application. Substrates were applied according to the ECL western blotting protocol (Amersham), then membranes were exposed to film.

4.2d. Western blotting, colorimetric substrate

The protocol was identical to the luminescent substrate protocol except that: a) secondary antibodies were 1:10 000 goat anti-rabbit alkaline phosphatase (Sigma), and 1:3 000 avidin alkaline phosphatase conjugate (Biorad), and b) the substrate was NBT/BCIP (see below), which precipitates directly onto the PVDF membrane. This method has the disadvantage that only one development level can be visualised, whereas the luminescent substrate allows multiple different exposures to film.

NBT/BCIP substrate was prepared as follows: 1 ml of 10 mg/ml x-phosphate/5-bromo-chloro-indolyl-phosphate (BCIP) in dimethylsulphoxide (DMSO), was mixed with 100 ml of 0.2 mg/ml 4 Nitroblue tetrazolium chloride (NBT) in 150 mM Tris pH 9.7 and 800 ml of 1M $MgCl_2$. This substrate was stored in the dark until use.

4.2e. Enzyme linked immuno-sorbent assay (ELISA)

Microtitre plates were coated with 1 mg/ml MBP:TE1$^{1-286}$ antigen in coating buffer (15 mM Na_2CO_3, 35 mM $NaHCO_3$) overnight at 4°C. Plates were washed 3 times with PBS-Tween (0.5% Tween 20 in PBS), then blocked with 100 ml of 0.5% milk powder solution in PBS for 1 hour at room temperature. Unbound blocking solution was removed with by washing 3 times with PBS-Tween.

Anti-MBP:TE1$^{1-286}$ antibody sera was diluted in PBS to 1:100, 1:1 000, 1:5 000, 1:10 000, 1:20 000, 1:50 000, 1:100 000, and 90 µl of each dilution was incubated in microtitre plate wells at room temperature for an hour, then 37°C for 45 min. Plates were washed 3 times with PBS-Tween, and incubated with 90 µl per well of a 1:10 000 dilution of goat anti-rabbit Alkaline phosphatase (Sigma) for 45 min at 37°C. Plates were washed 3 times in PBS-Tween, and substrate (1mg/ml *p*-Nitrophenyl phosphate disodium (pNPP), 3 mM $MgCl_2$ in coating buffer) was added, and incubated at room temperature until colour development was visible. Substrate development was quantified by reading absorbance on an Anthos HTII microtitre plate reader at 405 nm.

4.3 Production of maltose binding protein fusion proteins

4.3a. Construction of maltose binding protein fusion vectors

The maltose binding protein (MBP) expression vectors (pMAL-P2, pMAL-C2) (NEB) are designed to produce translational fusions of the *MalE* gene encoding the MBP and the gene of interest. Two vectors were available, pMAL-P2 includes the *MalE* signal sequence, which targets MBP fusion proteins to the periplasm, pMAL-C2 lacks the signal sequence.

The entire putative open reading frame of the *te1* cDNA cloned by Bruce Veit (pBV432) (Veit 1998) was cloned into pMAL-C2 and pMAL-P2 as an *Eco*RI fragment to create pDJ01 and pDJ02 respectively. To produce a fusion of the MBP with an N-terminal portion of the TE1 protein the 877 bp *Eco*RI -SmaI fragment derived from pBV432 was cloned into pMAL-C2 and pMAL-P2 to create pDJ03 and pDJ04 respectively. This region of the *te1* cDNA encodes the N-terminal 286 amino acid residues of the TE1 protein, and includes the first two RRMs.

4.3b. Small scale induction of *E. coli* cultures harbouring expression vectors

Small scale inductions of cultures were conducted as follows unless specified otherwise. Constructs were transformed into *E. coli* BL21 or TB-1 cells by heat shock (Section 2.2e), and re-streaked for single colonies to seed glycerol stocks or cultures. Starter cultures of 4 ml of LB Amp^{100} were inoculated with a single colony and grown overnight at 37°C. This overnight culture was used to inoculate 50 ml of LB Amp^{100} in a 250 ml conical flask. This culture was grown at 37°C, shaking at 200 rpm. Relative cell density was estimated by measuring the absorbance of cultures at 600nm. Log phase cultures were induced by adding IPTG to a final concentration of 0.6 mM.

Cells from 1 ml of culture were harvested by centrifugation for 10 minutes at 20 000 xg at 4°C in a benchtop microfuge immediately prior to and at various time points after the addition of IPTG. Cell pellets were resuspended in 2x Laemlii buffer (1x is 50 mM Tris pH 6.8, 2% SDS, 0.2% bromophenol blue, 20% (v/v) glycerol, 70 mM 2-ME), boiled for 5 minutes and centrifuged at 14 000 rpm for 5 minutes at room temperature. So that

approximately equal quantities of protein were loaded, cell pellets were resuspended in 100 μl of 2x Laemlii buffer for each 1.0 OD_{600} of the culture (see footnote[1]).

Equal volumes (10 μl) of the protein extracts were electrophoresed in 10% polyacrylamide as described in Section 4.2b. Initially induction trials were examined by staining gels with Coomassie. Once polyclonal sera to the MBP:$TE1^{1-286}$ fusion protein had been produced inductions were assayed by western blot as described in Sections 4.2c and 4.2d.

4.3c. Optimisation of MBP:$TE1^{1-286}$ expression conditions

Small scale cultures were induced as described for small scale inductions (above), except for the variations in the protocols described. Cultures of *E. coli* DH5α, BL21 and TB-1 strain carrying either pDJ01 vector (pMAL-C2:*te1*), or pDJ02 vector (pMAL-P2:*te1*) were induced at 30°C with 1 mM IPTG. Cultures of *E. coli* BL21 harbouring either pDJ01, pDJ02, or pDJ03 (pMAL-C2:*te1-5'*) were induced at 30°C and 37°C with 0.6 mM IPTG. In all cases cultures were induced by adding IPTG at log phase and grown for a further one, two or 3 hours. Inductions were assayed for expression by SDS-PAGE comparing induced cultures to cultures induced with various conditions. A culture harbouring pMAL-C2 with no cDNA insert which produces the maltose binding protein was always included as a positive control for the induction protocol.

4.3d. Large scale inductions and affinity purification of MBP:TE fusion proteins

Overnight 'starter cultures' of *E. coli* BL21 harboring pDJ03 (pMALC2:*te-5'*) in 1 ml LB media were used to inoculate 250 ml of LB Amp^{100}. These cultures were grown at 30°C until mid log phase, then induced with 1 mM IPTG, and incubated for a further 2 hours at 30°C. Cultures were chilled on ice, and cells harvested by centrifugation at 4 300 xg for 20 minutes at 4°C. All further manipulations were either in tubes on ice, or in a 6°C cold room. Cells were resuspended in 12.5 ml of column buffer (20 mM Tris-HCl pH 7.4, 200 mM NaCl, 1 mM EDTA, 10 mM 2-ME) and frozen at -20°C. Cell suspensions were then thawed on ice, and 1 ml of 1 mg/ml lysozyme in 10 mM Tris pH 8 was added. Cell suspensions were sonicated 4 times for 30 second pulses (separated

[1] For example, culture with OD_{600} of 0.5 would be resuspended in 50 μl of 2x Laemlii buffer

by 30 second breaks) at amplitude 12, at 23 kHz, on an Soniprep 150 sonicator (MSE Scientific Instruments).

Cell debris was removed by centrifugation at 9 000 xg for 30 min. at 4°C. The supernatant was decanted and diluted 1:2 with column buffer, and loaded onto an amylose resin (NEB) column at a flow rate of 10 x [diameter of column in cm]^2 ml/hr. The column was washed with 10 column volumes of column buffer. Bound fusion protein was eluted with column buffer containing 10 mM maltose and collected in 0.5 ml fractions.

Fractions were assayed for protein content by Bradford assays. Those fractions that contained significant quantities of total protein were examined by SDS-PAGE and Coomassie staining. Fractions containing significant quantities of fusion protein were pooled and concentrated using Nanosep 10K or 30K microconcentrators (Pall Filtron).

4.4. Production of anti MBP:TE1$^{1-286}$ antibody

4.4a Inoculation of rabbits and collection of sera

Approximately 4 ml of blood was taken from two New Zealand White rabbits (referred to hereafter as rabbit A and rabbit B) for pre-immune sera. To separate sera from blood cells, blood was incubated overnight at 4°C, then centrifuged at 500 xg at 4°C for 30 min to pellet blood cells. The sera supernatant was aliquoted and stored at -20°C. After collection of pre-immune sera, both rabbits were injected with 500 μg of MBP:TE1$^{1-286}$ fusion protein mixed with Freunds complete adjuvant (Sigma).

Booster injections of at least 100 μg of MBP:TE1$^{1-286}$ protein with Freunds incomplete adjuvant (Sigma) were given after 21, 35 and 60 days after the first inoculation. Small (5-10 ml) sample collections of immune sera were collected 7-10 days after the second, and all subsequent injections with MBP:TE1$^{1-286}$ fusion protein. Sera was collected as above, aliquoted, and stored at -20°C. The titre of immune sera collections were analysed via enzyme linked immuno-sorbent assays (ELISA, see below).

After it had been established by western blot that rabbit A had activity against the TE1 portion of the MBP:TE1$^{1-286}$ fusion protein (see results) further inoculations with at least

100 μg of MBP:TE1$^{1-286}$ fusion protein mixed with Freunds incomplete adjuvant were given to this rabbit. Blood collections of 30-50 ml were taken 7-10 days after injections, and sera was seperated from blood cells as above and stored at -20°C.

4.4b. Affinity purification of polyclonal antibody

E. coli BL21 cultures harboring pDJ03 (pMAL-C2:*te1-5'*) were induced, and the fusion protein purified as described in Section 4.3d. 4 L of culture produced 4 mg of protein, which was concentrated from amylose column elution fractions with Nanosep 10K microconcentrators (Pall Filtron), and stored in 850 μl of PBS.

Collections of sera to be affinity purified (collections 7, 8 and 9 of rabbit A) were thawed on ice and precipitated by gradual addition of saturated NH_4SO_4 up to 50% by volume. The precipitate was collected by centrifugation at 12 000 rpm at 4°C, and resuspended 5 ml of PBS, sealed in dialysis tubing (Size 5, Medicell International Ltd.), and dialysed 4 times against 2 L of PBS at 4°C. This sera contained 24.5 mg/ml protein.

A total of 3.8 mg of MBP:TE1$^{1-286}$ fusion protein was bound to a HiTrap® NHS activated affinity column (Pharmacia) as recommended by the manufacturers except that the loaded column was stored in PBS containing 0.1% NaN_3 at 4°C.

Ammonium sulphate precipitated sera was centrifuged at 14 000 rpm in a benchtop microfuge for 15 minutes at 4°C to pellet any IgG aggregates prior to loading on the column. All column buffers were degassed with a vacuum trap before use. The affinity column with covalently bound MBP:TE1$^{1-286}$ protein was washed with 10 ml of PBS containing 0.1% Tween 20, then sera was circulated through the column with a peristaltic pump at 0.5 ml/minute for 24 hours at 4°C. Bound antibodies were eluted with 0.2 M glycine-HCl pH 2.5, fractions of 1 ml were collected directly into 0.2 ml of 7% Tris (see footnote[2]). Fractions containing significant quantities of protein (determined by Bradford assay) were pooled, and concentrated with Nanosep 10K microconcentrators (Pall Filtron) to a final volume of 3 ml at 4.4 mg/ml protein concentration. This affinity purified sera was aliquoted and stored at -20°C. Aliquotes in current use were mixed with an equal volume of sterile glycerol and stored at -20°C.

4.5 Histidine-tagged construct with ArgU transfer RNA gene

4.5a. Construction of histidine tagging vector

The pPROEX vector (Gibco BRL) produces proteins with an N-terminal six histidine tag upon induction with IPTG. Proteins with this tag can be affinity purified with nickel-nitrilotriacetic acid columns. The entire open reading frame of the *te1* cDNA pBV432 was cloned into pPROEX-1 as an *Xba*I - *Kpn*I fragment to create pDJ05.

4.5b. Subcloning the ArgU tRNA gene by PCR

The presence of closely spaced AGG codons have been associated with low yields of heterologous protein production in *E. coli*, probably due to the low abundance of the *argU* tRNA gene product (see Discussion). The *te1* reading frame contains eight AGG codons after RRM2 (see Figure 4.12). We suspected that these codons were precluding efficient translation of the *te1* transcript. In an effort to circumvent this problem, the native *E. coli argU* gene was cloned from the pSBET vector using a PCR strategy.

Primers DJ1 and DJ2 (see Appendix 2) were designed to amplify a region of the *argU* gene from -44 to (where the transcription start site is +1) to +230bp. An adapter that included both NdeI (CATATG) and *Nsi*I (ATGCAT) sites was included on each primer. A PCR product of the *E. coli argU* gene was produced from 10 ng of the vector pSBET with using a standard PCR reaction (Section 2.3c) except that final DJ1 and DJ2 primer concentrations in the reaction were 1 mM. The thermocycling program DJ1 was used (Appendix 1). The *argU* PCR product was examined for quantity and homogeneity by gel electrophoresis, and then digested with *Nsi*I.

To remove the small end fragments that were liberated by *Nsi*I dgiestion the digest was purified with a Concert PCR Extraction Kit (Gibco BRL). The purified PCR product was then cloned into the *Nsi*I site of pDJ05 to create pDJ06. In summary, pDJ06 consists of the pPROEX expression vector with the *te1* cDNA cloned in frame with the six histidine tag, and the ArgU tRNA gene cloned into the *Nsi*I site.

---

[2] This proportion of 0.2 M glycine-HCl to 7% Tris had been empirically determined to restore pH to 7.0.

4.5c. Expression trials with histidine tagging vector

Expression trials were performed with pDJ06 using small scale inductions, essentially as described in Section 4.3b. Several variations of conditions were examined. *E. coli* strains DH5α, BL21, and TB-1 were induced at 37°C in LB media, taking cell samples one, two and 3 hours after induction with 0.6 mM IPTG. Each of these strains was also induced in SOC media (2% tryptone, 0.5% yeast extract, 0.5% NaCl, 2.5 mM KCL, 10 mM $MgCl_2$) and Terrific Broth (1.2% tryptone, 2.4% yeast extract, 0.4% (v/v) glycerol, 17 mM KH_2PO_4, 72 mM K_2HPO_4, 20 mM glucose) at 37°C with 1 mM IPTG.
Media were prepared as described in Sambrook (1989).

A vector consisting of the *Trifolium repens* ACC oxidase (*TRAC02*) cDNA in pPROEX-1, which had been shown to express high levels of histidine-tagged TRACO2 protein (D Hunter, pers. comm.), was used as a positive control for the induction protocol. A culture of *E. coli* TB-1 cells harbouring pPROEX:*TRACO2* was always grown and induced with the same conditions as pDJ06 cultures.

Protein extracts from induced and uninduced cultures were compared by Coomassie stained SDS-PAGE gels (Section 4.2b), and western blots using the affinity purified anti-MBP:$TE1^{1-286}$ antibody at 1:6 000 dilution (Section 4.2c).

4.5d. Purification of histidine-tagged TE1 protein

Histidine-tagged proteins have a high affinity for nickel-nitrilotriacetic acid columns, which allows affinity purification. Attempts were made to affinity purify the small quantities of 6H:TE1 that were detected with western blots. 100 ml of BL21 culture harboring pDJ06 vector was induced as for previous large scale inductions (Section 4.3d), except that the culture was grown at 37°C rather than 30°C. Cells were harvested, resuspended in lysis buffer (50 mM Tris-HCl pH 8.5, 10 mM 2-ME, 1 mM phenylmethylsulfonyl Fluoride (PMSF)), and frozen at -20°C. Cell suspensions were thawed on ice and sonicated as described in Section 4.3d. Cell debris were removed by centrifugation at 9 000 xg for 30 minutes at 4°C.

All further manipulations were performed at 6°C. A 1.5ml nickel-nitrilotriacetic acid column (Ni-NTA resin, BRL) was prepared by lining the bottom of a 5ml syringe with sterile glass fibre, adding 1.5 ml of Ni-NTA resin, and centrifuging for 1 500 xg for 5

minutes. The column was equilibrated with 2 ml of buffer A (20 mM Tris-HCl pH 8.5, 100 mM KCl, 20 mM imidazole, 10 mM 2-ME, 10% (v/v) glycerol). The cell extract was loaded onto the column at a flow rate of 0.5 ml/min. The column was washed with 15 ml of buffer A, 3 ml of buffer B (20 mM Tris-HCl pH 8.5, 1 M KCl, 10 mM 2-ME, 10% (v/v) glycerol), and then 3 ml of buffer A. Bound protein was eluted with 5.5 ml of buffer C (20 mM Tris-HCl pH 8.5, 100 mM KCl, 100 mM imidazole 10 mM 2-ME, 10% (v/v) glycerol) in 0.5 ml fractions. Fractions were concentrated by a factor of four (400 μl of elution buffer concentrated to 100 μl) with Nanosep 10K microconcentrators (Pall Filtron), and the total protein content estimated by Bradford assay. All fractions were examined by SDS-PAGE and Coomassie staining, and by western blotting using the affinity purified anti-MBP:TE1$^{1-286}$ antibody at 1:6 000 dilution (Section 4.2c). Protein samples containing imidazole were heated to 37°C rather than boiled before electrophoresis to prevent lysis of protein by imidazole.

4.6 Transformation of *Pichia pastoris* with integrating expression constructs

4.6a. Overview of *Pichia* expression system

Pichia pastoris is a methyltropic yeast, capable of utilising methanol as a carbon source. The *Pichia* expression vector pPIC9K (Invitrogen) is an integrating vector that allows *his4* histidine auxotroph strains to grow on media lacking histidine, and also confers resistance to the antibiotic G418 (Geneticin, Gibco BRL) in a dose dependant manner. The genes cloned into the pPIC9K vector are under control of the native *alcohol oxidase 1* (*AOX1*) promoter, which is induced by methanol to very high levels (Koutz 1989; Tschopp 1987), and repressed by glucose. The vector produces an N-terminal fusion of the *Pichia* α factor secretion signal (αFS), with the cloned gene which results in the fusion protein being secreted into the culture media. Two *his4* strains of *Pichia pastoris* were available for transformation with vector pPIC. Strain KM71 utilises methanol as a carbon source slowly due to a mutant *AOX1* locus, and GS115 which is wild type for methanol utilisation.

The entire coding region of the *te1* cDNA was cloned into pPIC9K and the resulting vector (pDJ07) was transformed into the strain KM71 by electroporation. Transformants

were initially selected on media lacking histidine, and then by resistance to increasing levels of G418. Expression trials were performed and assayed by SDS-PAGE and western blotting.

4.6b. Production of integrating expression vector in *E. coli*

An *Eco*RI fragment of pBV432, encoding the complete open reading frame of the *te1* cDNA, was cloned into the vector pPIC9K (Invitrogen) as described in Section 2.3e. This vector will produce a translational fusion of the *te1* open reading frame, and the open reading frame of the α factor secretion signal peptide (αFS) contained in the pPIC9K vector. The construct is referred to as pDJ07 (pPIC9K/*te1*), and was designed to secrete the fusion protein αFS:TE1 into the growth medium.

4.6c. Preparation of electro-competant *Pichia* cells

To remove any traces of SDS or other detergents that may compromise electroporation efficiency, all glassware and plastics used to prepare *Pichia* cultures for electroporation were washed thoroughly with hot water, rinsed in milli-Q water, and autoclaved filled with milli-Q water. This water was then discarded. Cultures for electroporation were prepared as follows. 10 ml of YPD media (1% yeast extract, 2% peptone, 2% dextrose) in a 50 ml flask was inoculated with a single colony from a YPD plate (YPD media with 2% agar), and grown for 24 hours at 30°C, shaking at 200 rpm. Two 1 L flasks containing 500 ml of YPD media were inoculated, one with 50 µl of this 10 ml culture, the other with 100 µl of this 10 ml culture. These cultures were grown at 30°C, shaking at 200 rpm until the OD_{600} readings were approximately 1.3 (11 hours, and 12.5 hours respectively). Cells were harvested by centrifugation at 1 500 x g at 4°C for 5 minutes, and resuspended in 500 ml of sterile ice cold milli-Q water. Cells were harvested, and resuspended 3 more times in the following solutions; 250 ml of ice cold sterile Milli-Q water, 10 ml of ice cold 1 M D-sorbitol (Sigma), and finally 1 ml of ice cold 1M sorbitol. A small proportion of the electrocompetant cells were used immediately (see below), and the remainder were aliquoted, frozen in liquid nitrogen, and stored at -80°C.

4.6d. Electroporation of *Pichia* cells

Vectors pPIC9K and pDJ07 were prepared by alkaline lysis and purified by phenol/chloroform extraction until the interface was clear. The vectors were linearised with *Xcm*I, precipitated with 1/10th volumes of 7.5 M sodium acetate and 2.5 volumes of

ethanol, and resuspended in sterile milli-Q water to a final concentration of approximately 0.5 mg/ml. For both pPIC9K and pDJ07 vectors, 100 ng, 1 μg and 2 μg of *Xcm*I-linearised vector was added to 100 μl of *Pichia* cells prepared for electroporation (see previous section). Cell-DNA mixtures were transferred to ice cold 2 ml electroporation cuvettes (EquiBio Ltd) with a 2 mm electrode gap, and incubated on ice for 5 minutes. Cuvettes were pulsed in a Biorad GenePulser electroporation device with the parameters recommended by Biorad for *Sacchararomyces cerevisiae* (1.5kV, 25μF, 200W). The cells were immediately resuspended in 1 ml of ice cold 1 M sorbitol, and then transferred to a sterile microfuge tube. 200 μl of these cell suspensions were plated onto RDB plates (1M sorbitol, 2% dextrose, 1.34% yeast nitrogen base without amino acids (Gibco BRL), 1% (v/v) glycerol, $4x10^{-5}$% biotin, 2% agar), and grown for two days at 30°C to select for histidine prototrophy. Colonies were streaked for single colonies on RDB plates, allowed to grow for two days at 30°C, then single colonies were used to inoculate 2 ml YPD cultures for glycerol stocks. Glycerol stocks were produced exactly as for *E. coli* (Section 2.2a) except that 0.8 ml of 48 hour YPD culture[3] was mixed with 0.2 ml of sterile glycerol. Initially 40 glycerol stocks were made from 40 single colonies.

4.6e. Selection of transformants with increasing numbers of integrations

All of approximately 4 000 colonies from the plate transformed with 2 μg pDJ07 were resusupened in 1 ml of sterile water by plating the water directly onto the RDB transformation plate and swirling gently to resuspend the cells. This cell suspension was drawn from the surface of the plate with a 1 ml autopipette, and the cell density was estimated by measuring its absorbance at 600 nm. 10 μl (approximately 10^5 cells) of this cell suspension was plated onto YPD plates containing 0.25, 0.5, 1.0, 1.5, 2.0 and 3.0 mg/ml G418 sulphate antibiotic, and grown for 3 days at 30°C. G418 resistance levels were confirmed by re-streaking each colony for single colonies twice on YPD plates containing the same concentration of G418 that they were initially selected on. Lines that grew consistently were chosen for more extensive G418-resistance tests. Each of these chosen lines were streaked for single colonies on YPD plates containing 0.25, 1.0 and 4.0 mg/ml G418. Glycerol stocks were made from eight lines representing different levels of resistance to G418 (see Table 2.1).

[3] Since linearised pPIC vectors integrate into the genome, no selection is required after

4.6f. Extraction of *Pichia* genomic DNA

Transformed and parent (KM71) strains of *Pichia* were streaked for single colonies onto YPD plates and grown for 48 hours at 30°C. A single colony of each line was used to inoculate 10 ml of YPD media. Cultures were grown overnight at 30°C, shaking at 200rpm. Cells were harvested by centrifugation at 1 500 xg for 10 minutes at room temperature. Cell pellets were resuspended in 10 ml of sterile water, harvested by centrifugation at 1 500 xg, then resuspended in 2 ml of freshly prepared SCED (1 M sorbitol, 10 mM sodium citrate pH 7.5, 10 mM EDTA, 10 mM DTT).

The SCED cell suspensions were incubated with 4 µl of 4.5 mg/ml Lyticase (Sigma) at 37°C for 1 hour, then 2 ml of 1% SDS was gently mixed with the suspensions, and tubes were placed on ice for 5 minutes. Finally 1.5 ml of 5 M potassium acetate pH 5.9 was gently mixed with the cell suspensions, and they were centrifuged at 10 000 xg for 10 minutes at 4°C. The supernatant was transferred to a new tube, mixed with 2 volumes of ethanol, and incubated at room temperature for 15 minutes. The precipitate was collected by centrifugation at 10 000 xg for 20 minutes at 4°C. The pellet was gently resuspended in 0.7 ml of TE 50/5 (50 mM Tris pH 8, 10 mM EDTA) and extracted with an equal volume of phenol:chloroform (1:1), and then and equal volume of chloroform:isoamyl alcohol (24:1). The nucleic acids were precipitated overnight at -20°C with 1/2 volume of 7.5 M ammonium acetate pH 7.5 and 2 volumes of ethanol, then washed with 80% ethanol, air dried and resuspended in 48 µl of TE 50/5 with 2 µl of 10 mg/ml RNase A.

4.6g. Southern blotting

Pichia genomic DNA samples were quantified by spectrophotometry. *Eco*RI and *Bgl* II digests containing 300 ng of DNA were electrophoresed on a 0.8% TAE gel as described in Section 2.13c. Transfer of the DNA to a nylon membrane and fixing by UV light irradiation was performed exactly as described in Section 2.13c. The membrane was pre-hybridised in 5 x Denhardts solution (0.1% Ficoll® (Ficoll® 400, Pharmacia), 0.1% polyvinylpyrrolidone (Sigma), 0.1% BSA (Fraction V, Roche)), 5 x SSC, 1% DIG blocking reagent (Roche), 0.1% N-lauryl sarcosine, 0.02% SDS at 65°C for an hour.

---

transformation to maintain the vector.

A DIG-labelled probe comprising the 5' end of the *te1* cDNA was produced by PCR using the primers[4] SW1 and TE1 with the thermocycling program DJ50x40 as described in section 2.9e. The probe was assayed by gel electrophoresis for quantity and homogeneity before use. Hybridisation and washes were performed at 65°C as described in Section 2.9e, except for the hybridisation buffer (described above). Stringency washes were performed in 0.1 x SSC, 0.1% SDS at 65°C.

Blocking, anti-DIG AP antibody binding, and washes were performed as recommended by Roche. Antibodies were bound at 1:20 000 dilution. The CDP-Star substrate was used.

4.6h. Induction trials with transformed *Pichia pastoris*

The genes cloned into the Pichia vector pPIC9K are under control of the native *alcohol oxidase 1* (*AOX1*) promoter, which is induced by methanol to very high levels (Koutz 1989; Tschopp 1987), and repressed by glucose. The vector pDJ07 (pPIC9K:*te1*) was transformed into the *Pichia* strain KM71, which utilises methanol as a carbon source slowly due to a mutant *AOX1* locus. These lines were grown initially in YPD, which allows vigorous growth (but represses genes under control of the *AOX1* promoter), then harvested and resuspended in a minimal media containing methanol and lacking glucose.

For all inductions of *Pichia* strains the untransformed parent strain KM71 was included as a negative control. A line that had been transformed with same *Pichia* vector (pPIC9K) carrying the cDNA for the *Candida albicans* gene *BGL2* was included as a positive control. The BGL2 line had been shown to produce high levels of the BGL2 protein, which is secreted into the media (R Ramsay, pers. comm).

Pichia glycerol stocks were streaked for single colonies onto YPD plates containing an appropriate concentration of G418 antibiotic, as determined by *in vivo* screening for G418 resistance (Section 4.6e). These plates were incubated at 30°C for 48 hours. 10

[4] These primers span the region between 106 to 1127 bp of the *te1* cDNA.

ml of YPD media was inoculated with a single colony from these plates, and incubated at 30°C for 48 hours, shaking at 200rpm. A 1 ml sample of cultures in an un-induced state was taken at this stage; the cells from this 1 ml of culture were harvested by centrifugation at maximum speed for 3 minutes at room temperature and the supernatants removed by aspiration. Cell pellets and supernatants were frozen separately in liquid nitrogen and stored at -80°C. Cells were harvested from the remainder of the culture by centrifugation at 1 500 xg for 5 minutes at room temperature, the supernatant drained off, and the cell pellet resuspended in 1 ml of buffered methanol-complex medium (BMMY; 1% yeast extract, 2% peptone, 100 mM potassium phosphate pH 6.0, 1.34% yeast nitrogen base without ammonium sulphate and without amino acids (Gibco BRL), $4x10^{-5}$% biotin, 0.5% methanol). BMMY cultures were grown for 2-5 days at 30°C shaking at 200rpm. To account for evaporation of methanol, 100% methanol to a final concentration of 0.5% was added every 24 hours. After 48 hours growth in BMMY medium samples of cultures were taken (volumes varied between inductions), cells harvested by centrifugation in a benchtop centrifuge, at maximum speed for 3 minutes at room temperature, and both cell pellets and supernatants were frozen separately in liquid nitrogen, and stored at -80°C.

Both cell pellets and supernatants were assayed for αFS:TE1 protein expression by SDS-PAGE and western blotting. Cell pellets were resuspended in 2x Laemlli buffer, boiled for 5 minutes then centrifuged for 5 minutes at room temperature to extract protein. In some cases supernatants were concentrated using Nanosep 10K microconcentrators (Pall Filtron). See results for quantities loaded.

RESULTS (4)

Table 4.1 describes the expression vectors produced in attempts to produce the TE1 protein. Diagrams of these vectors are shown in Figure 4.1.

Table 4. 1 . Protein expression vectors

Vector	Construction	Expected protein product	Expected protein molecular weight
pDJ01	pMAL-C2 containing entire ORF of *te1* cDNA (pBV432). Cloned as *Eco*RI *Xba*I fragment	MBP:TE1 cytoplasmic MBP	113 kD
pDJ02	pMAL-P2 containing entire *te1* cDNA. Cloned as pBV432 *Eco*RI *Xba*I fragment	MBP:TE1 periplasmic MBP	113 kD
pDJ03	pMAL-C2 with 5' end of *te1* cDNA. Cloned as *Eco*RI *Xba*I fragment of pBV426 (an *Eco*RI SmaI subclone of pBV432)	MBP:TE1$^{1-286}$ cytoplasmic MBP	79 kD includes 42.7 kD MBP and 37 kD TE1$^{1-286}$
pDJ04	pMAL-P2 with 5' end of *te1* cDNA Cloned as *Eco*RI *Xba*I fragment of pBV426	MBP:TE1$^{1-286}$ periplasmic MBP	79 kD
pDJ05	pPROEX containing entire *te1* cDNA cloned as a *Xba*I *Kpn*I fragment of pBV432	6H:TE1	71 kD
pDJ06	derivative of pDJ05, with ArgU PCR product cloned into *Nsi*I site	6H:TE1	71 kD
pDJ07	pPIC9KK containing entire *te1* cDNA cloned as *Eco*RI fragment of pBV432	αFS:TE1	70 kD
pDJ12	pSBET containing *Bam*HI fragment of *te1* cDNA from pBV432	T710$^{1-13}$:TE1$^{1-577}$*	63 kD

*Cloning into the *Bam*HI site of pSBET produces a fusion of the N-terminal 13 amino acids of T7 phage protein 10 with the N-terminal 577 amino acids of the TE1 protein.

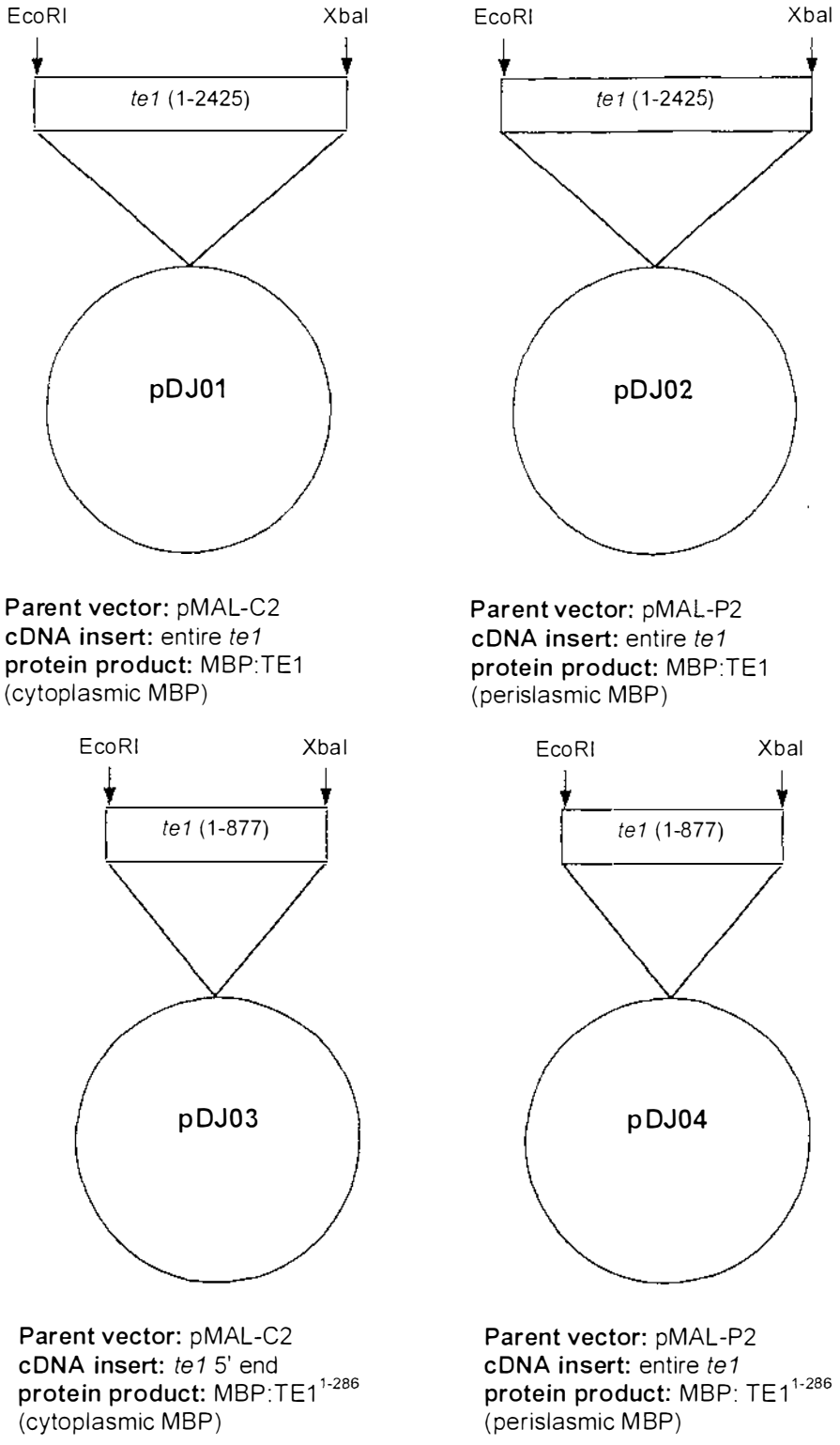
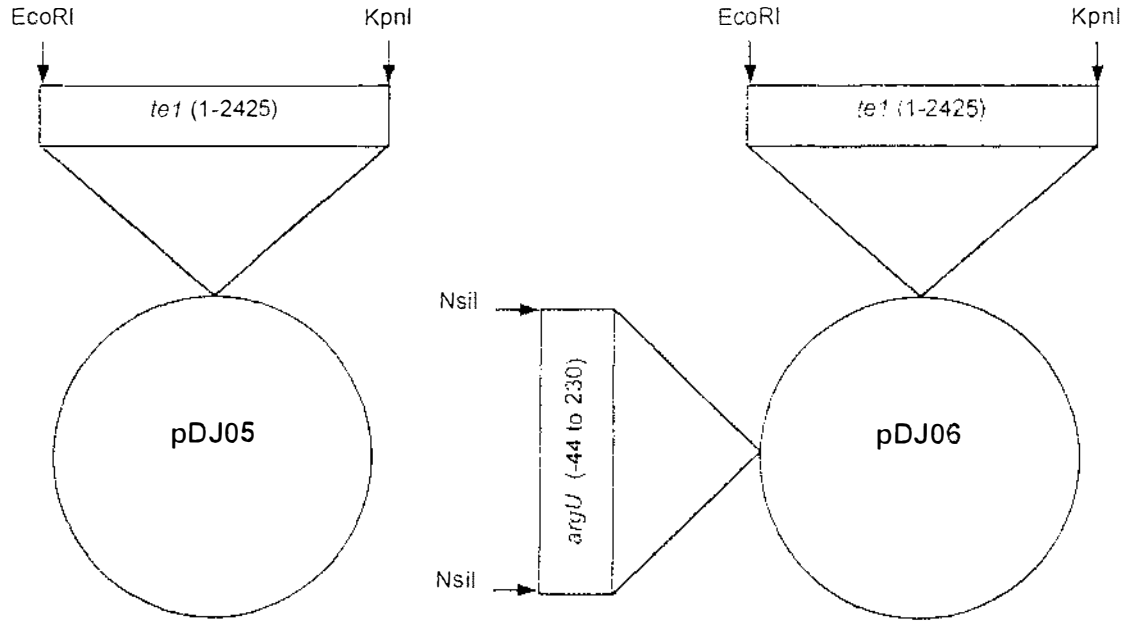
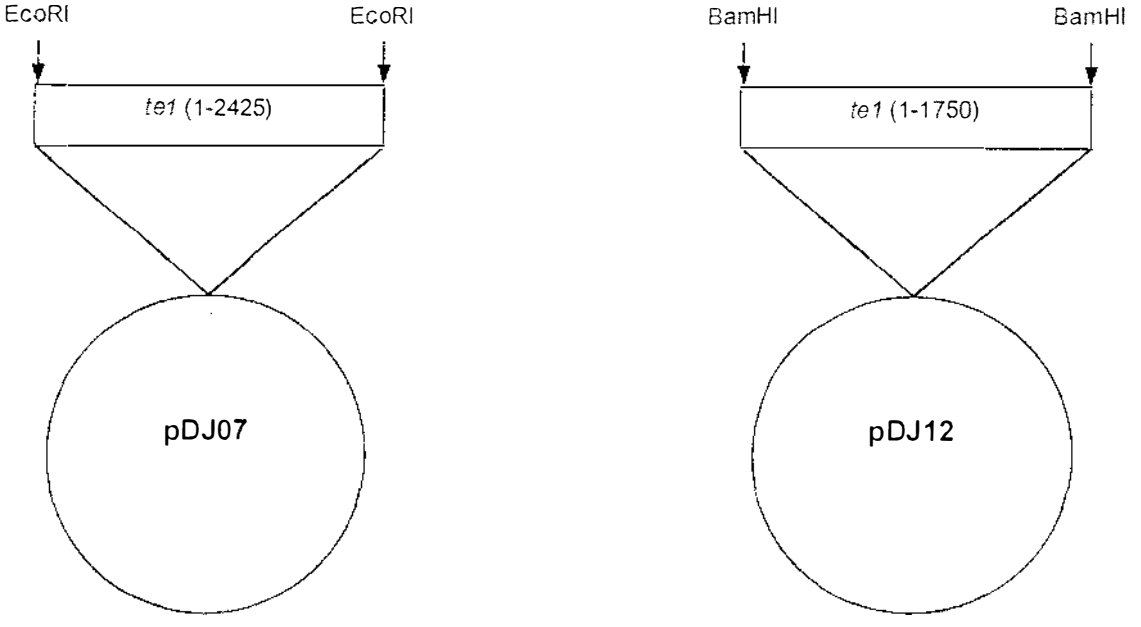


Figure 4.1a. Protein expression vectors: MBP fusion vectors. Onnly the essential features of vector construction are indicated here. Other details of the pMAL series can be found at www.neb.com. The region of the *te1* cDNA that is included are indicated in nucleotides, the restriction sites that were used to clone these fragments from the various *te1* cDNA clones are indicated. Table 4.1 includes a more precise description of the construction of these vectors.



Parent vector: pPROEX
cDNA insert: entire *te1*
protein product: 6H:TE1

Parent vector: pDJ05 (pPROEX)
cDNA insert: entire *te1*
protein product: 6H:TE1



Parent vector: pPIC9K
cDNA insert: entire *te1*
protein product: αFS:TE1

Parent vector: pSBET
cDNA insert: *te1* BamHI fargment
protein product: $T710^{1-13}$:$TE1^{1-577}$

Figure 4.1b. Protein expression vectors: other constructs. As with Figure 4.1a other details of vector construction can be found in Table 4.1. Details of the pPROEX vector www.lifetech.com, and www.invitrogen.com for pPIC9K and Schenk (1995) for details of the parent vector pSBET.

4.7 Only the amino-terminal half of the TE1 protein was produced as MBP fusions

Pilot studies by Bruce Veit with a vector identical to pDJ01 showed no detectable MBP:TE1 fusion protein expression when assayed by SDS-PAGE (data not shown). Small scale inductions of pDJ01, pDJ02 (pMAL-C2:*te1*, and pMAL-P2:*te1* respectively) at 37°C in BL21 cells did not show detectable inductions in my experiments either (data not shown), although pDJ03 (pMAL-C2:*te1-5'*), produced a clear band of the expected 79 kD molecular weight in these conditions (data not shown). Similarly at 30°C in BL21 cells the truncated fusion protein from pDJ03 was produced, but no full length fusion protein was visible by SDS-PAGE of *E. coli* extracts carrying pDJ01.

To determine whether small, but usable quantities of the full length fusion protein were being produced, 2 L of pDJ01 (pMALC2:*te1*) and pDJ03 (pMALC2:*te1-5'*) were induced in BL21 cells at 30°C as described for large scale inductions. The cell extracts were affinity purified with amylose columns as described in Section 4.3d. While amylose affinity purified truncated fusion protein (MBP:TE$^{1-286}$) of approximately the expected 79 kD was produced from pDJ03, the full length MBP:TE1 fusion protein obtained from an amylose affinity column was barely detectable in Coomassie stained SDS-PAGE gels (Figure 4.2). Polyclonal antibodies were raised against the MBP:TE1$^{1-286}$ fusion protein (Sections 4.4 and 4.10).

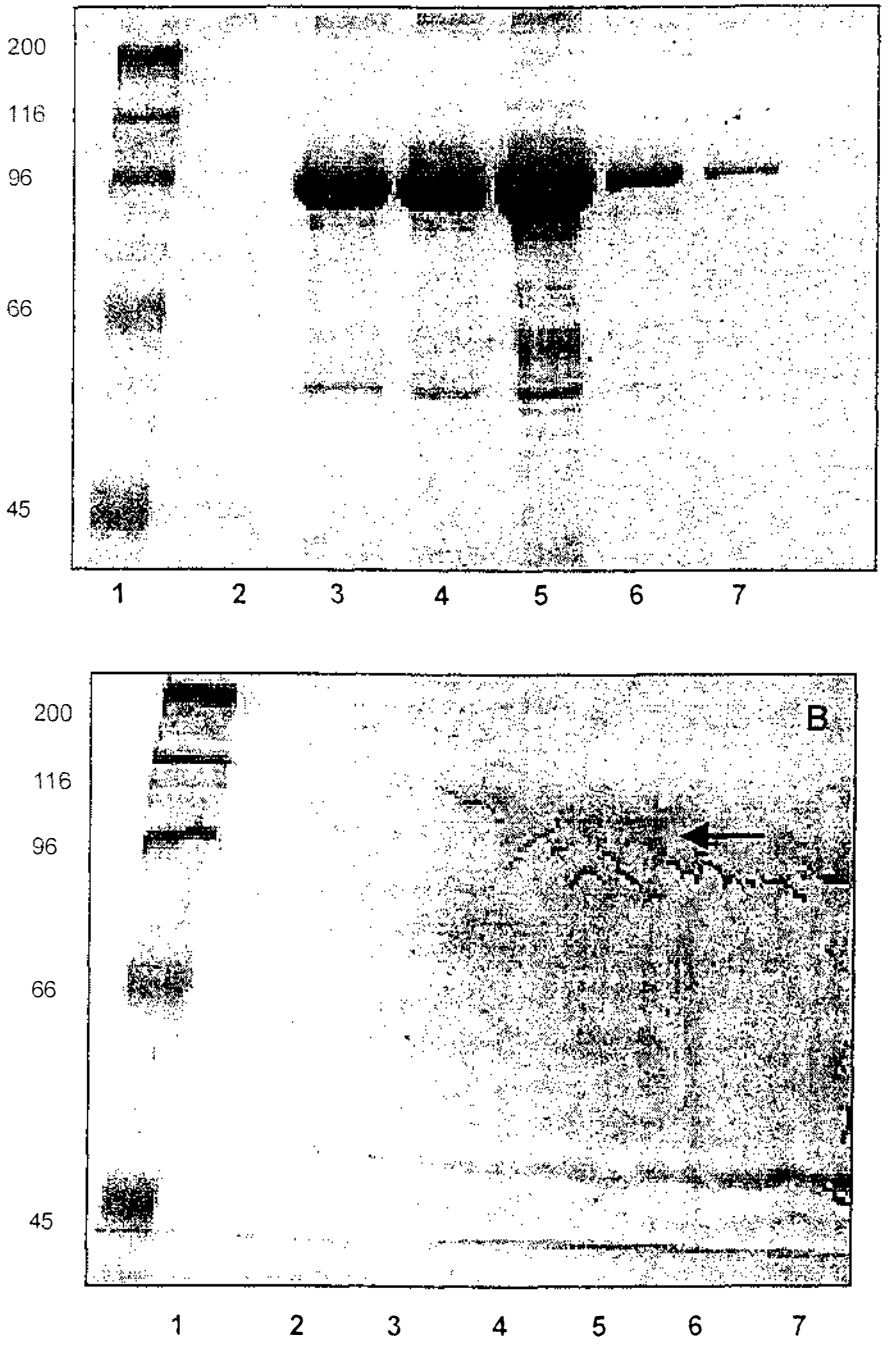


Figure 4.2. SDS-PAGE gel of affinity purified MBP:TE1 proteins. Significant quantities of the 79 kD MBP:TE1$^{1-286}$ protein , could be affinity purified on amylose columns (**A**), but only very small quantities of full length MBP:TE1 fusion protein were purified (arrow, lane 5 in B), expected molecular weight 113 kD. In each case 2 L of *E. coli* BL21 cells were induced at 30°C. The cells extract was affinity purified on an amylose column and eluted in 16 1 ml fractions. The protein concentrations of odd numbered fractions was estimated by Bradford assay, and either 10 µl (MBP:TE1$^{1-286}$) (**A**), or 20 µl (MBP:TE1) (**B**) of the protein-containing fractions was electrophoresed. Lanes are; **1**) molecular weight maker (molecular weights are indicated at left of gel), **2**) fraction 1, **3**) fraction 3, **4**) fraction 5, **5**) fraction 7, **6**) fraction 9.

4.8 Expression of full length TE1 protein from the ArgU tRNA construct and the ArgU tRNA-supplemented pSBET vector

No production of the 6H:TE1 protein was detectable in Coomassie stained SDS PAGE gels from initial small scale inductions of pDJ06 (pPROEX:*te1* containing the *argU* gene) when induced at 37°C in DH5α, BL21 and TB1 cell lines (data not shown). Western blotting using the anti-MBP:$TE1^{1-286}$ revealed that small quantities of the 6H:TE1 protein were being produced (Figure 4.3). Western blots of further inductions indicated that neither SOC media nor Terrific Broth improved the level of 6H:TE1 produced from this vector in these cell lines (data not shown).

The quantity of 6H:TE1 protein produced was further examined by affinity purification of an induced *E. coli* extract. Less than 0.2 μg of 6H:TE1 protein per ml of culture were produced, and the extract contained *E. coli* contaminants that were at least as abundant as the tagged protein (Figure 4.4).

The lack of expression from the pDJ06 construct led us to question whether the construction of pDJ06 was sufficient to overcome translational fidelity problems, specifically whether all the elements of the *ArgU* promoter had been cloned from the pSBET vector by PCR (see discussion). To address this issue, a *Bam*HI fragment of the *te1* cDNA vector pBV407 was cloned[5] into the original pSBET vector. The *argU* construct in this vector had been shown to contain a sufficient proportion of the native *E. coli argU* promoter to overcome translational fidelity problems of other proteins with frequent AGG codons (Schenk 1995). The vector, referred to here as pDJ12 (pSBET:*te1-BamHI*), codes for 577 of the 656 amino acids of the full length cDNA, and includes all 3 RRMs. The predicted molecular weight, including the 13 amino acids derived from the vector is 63 kD. Small scale inductions of pDJ12 in BL21 (DE3) cells at 37°C showed that significant quantities of a protein of approximately the correct molecular weight were produced (Figure 4.5). The cloning stategy used to produce this vector appeared to have some advantages over the pDJ06 strategy where the ArgU tRNA was subcloned into another vector (see Figure 4.1b for vector maps).

---

[5] The pSBET:*te1-BamHI* vector was produced by Carmel Gillman

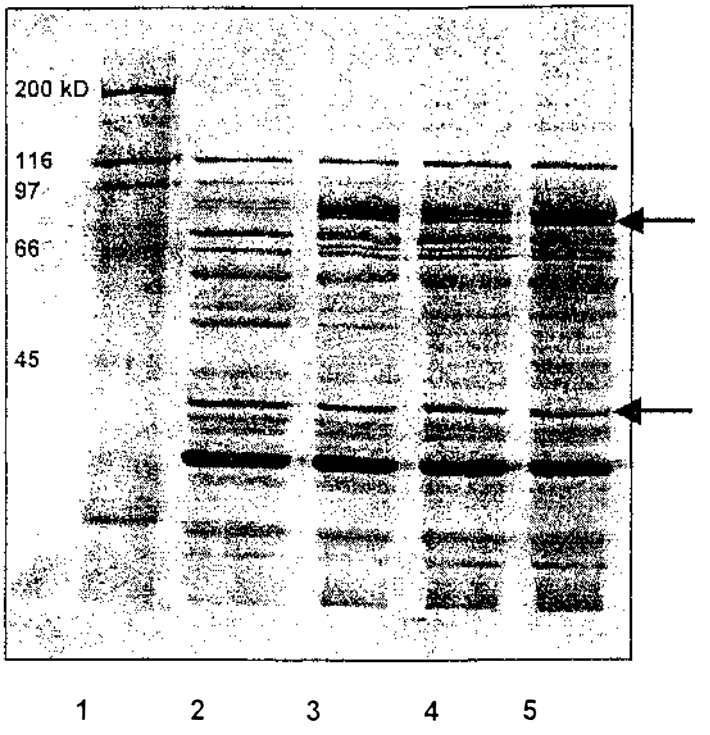


Figure 4.3. Western blot of histidine tagged TE1 protein (6H:TE1) induction from pDJ06 vector. Cell extracts were from inductions of pDJ06 vector in BL21 cell line at 37°C. Lanes are **1**) Biotinylated molecular weight marker (molecular weights are indicated at left), **2**) cell extract prior to induction, **3**) cell extract 1 hour after induction, **4**) cell extract 2 hours after induction, **5**) cell extract 3 hours after induction. The large number of bands that developed were probably due to cross reactivity of the anti-MBP:TE1$^{1\text{-}286}$ primary antibody, which had not yet been affinity purified, with *E. coli* proteins. Despite this 'non-specific' signal, a new band of the expected molecular weight (71 kD) can be seen in induced samples (upper arrow) although samples are equally loaded (lower arrow). Both DH5α and TB1 cell lines showed similar results.

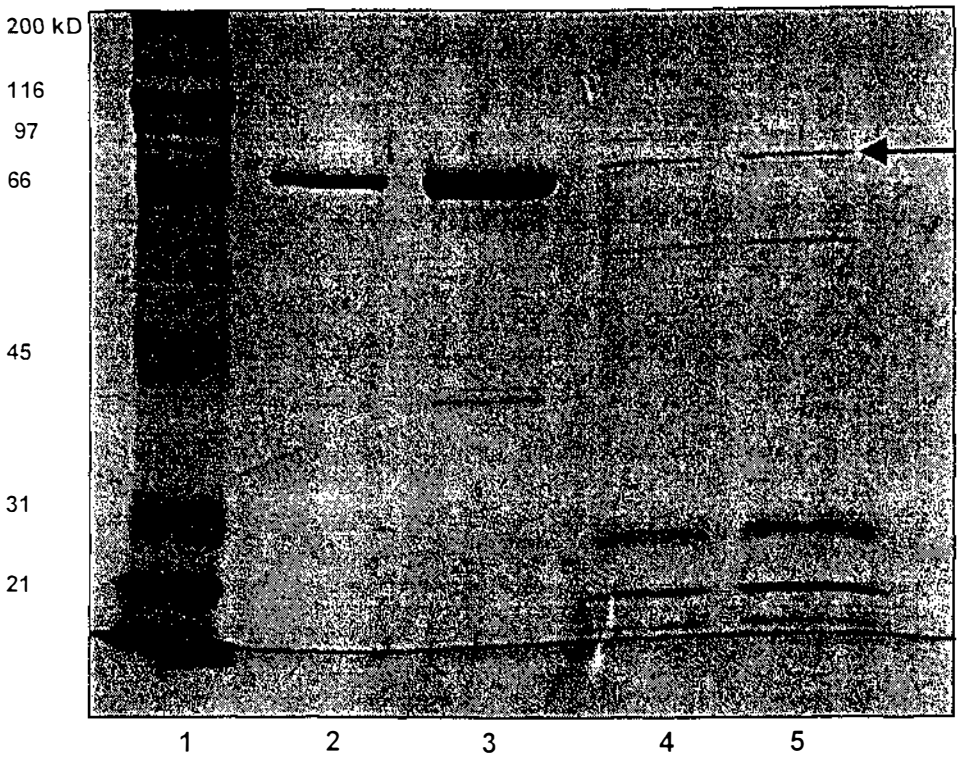


Figure 4.4. Coomassie stained SDS-PAGE gel of affinity purified 6H:TE1 protein. Only very small quantities of the 6H:TE1 protein were obtained from Ni-NTA affinity columns, and significant levels of contaminating proteins were present. Lanes are **1**) molecular weight marker (molecular weights are at left), **2**) 1 μg of affinity purified MBP:TE1$^{1-286}$ fusion protein, **3**) 5 μg of affinity purified MBP:TE1$^{1-286}$ fusion protein, **4**) 10μl of concentrated eluate from Ni-NTA column, **5**) 15μl of concentrated eluate from Ni-NTA column. The column eluates represent the total purified protein from 5 ml of culture

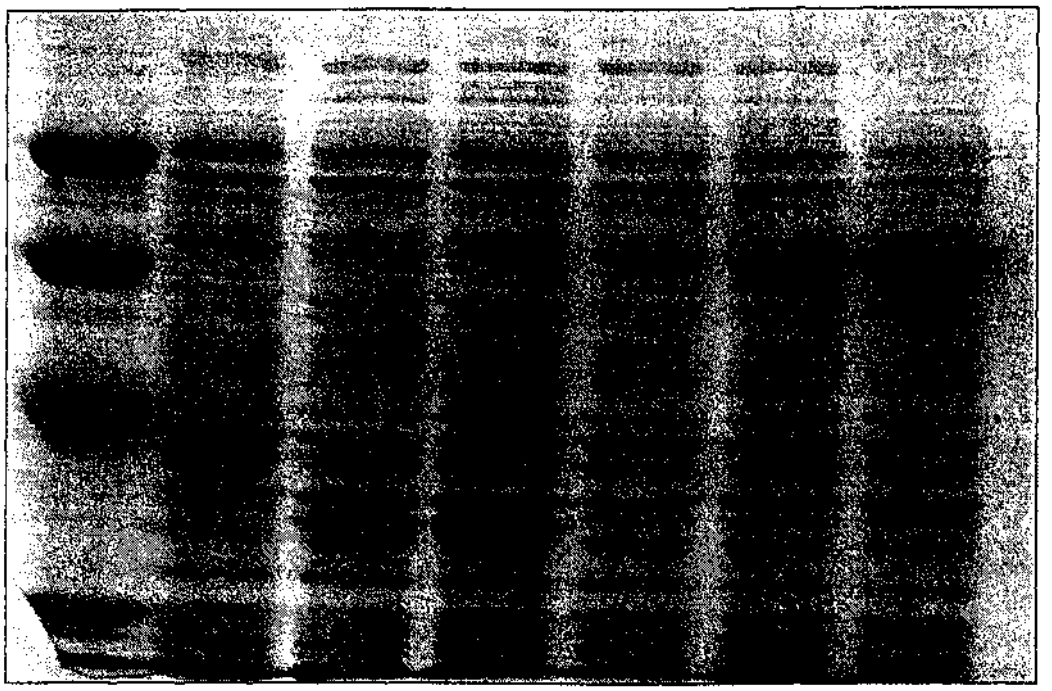


Figure 4.5. The pSBET vector produces abundant TE1$^{1-577}$ fusion protein. A pSBET construct containing a truncated *te1* cDNA coding for the first 577 residues of the TE1 protein was induced as described in Section 4.3b. The 63 kD protein in the extract is clearly visible in lanes 6 and 7. *E. coli* TB1 cells expressing the *Trifolium repens* TRACCO2 protein from the pPROEX vector (NEB) were included as a positive control. Lanes are; **1**) molecular weight marker (Pharmacia), **2**) pPROEX:*TRACCO2* uninduced, **3**) pPROEX:*TRACCO2* 1 hour post-induction, **4**) pPROEX:*TRACCO2* 2 hours post-induction, **5**) pSBET:*te1-BamHI* uninduced, **6**) pSBET:*te1-BamHI* 1 hour post-induction, **7**) pSBET:*te1-BamHI* 1 hours post-induction.

4.9 The *Pichia pastoris* expression system did not produce significant quantities of the TE1 protein

4.9a. Analysis of protein expression by transformed *Pichia pastoris* lines

All *Pichia* expression experiments used derivatives of the KM71 *Pichia pastoris* strain that had been transformed with the integrating vector pPIC9K containing the *te1* cDNA. Clonal lines should contain stable integrations of the vector, and so cultures did not contain antibiotics. Lines were charactersied by the level of resistance to the antibiotic G418 (Table 2.1), and are referred to as DJ41, DJ42 *etc.*

Initially, lines DJ41, DJ47 and DJ48 were chosen for expression trials. Cultures of KM71 were included as a negative control, and the BGL2 line (which secretes the *Candida albicans* BGL2 protein into the media) was included as a positive control. All lines were induced as described in Section 4.6h. Since the α Factor signal sequence encoded by the pPIC9K vector should result in secretion of the fusion protein, the supernatants from cultures induced with methanol were compared with un-induced cultures by SDS-PAGE and western blotting. The control strain BGL2 showed induction of a protein, but no αSF:TE1 protein was detected (Figure 4.6).

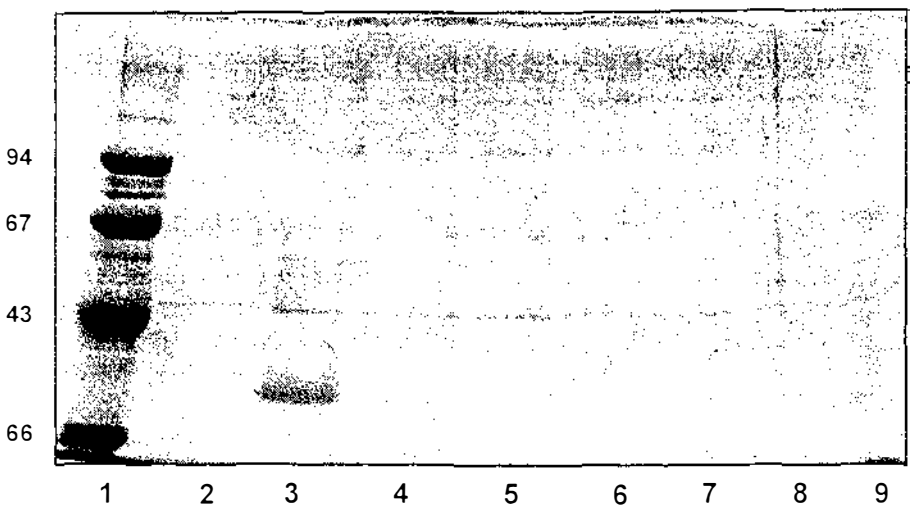


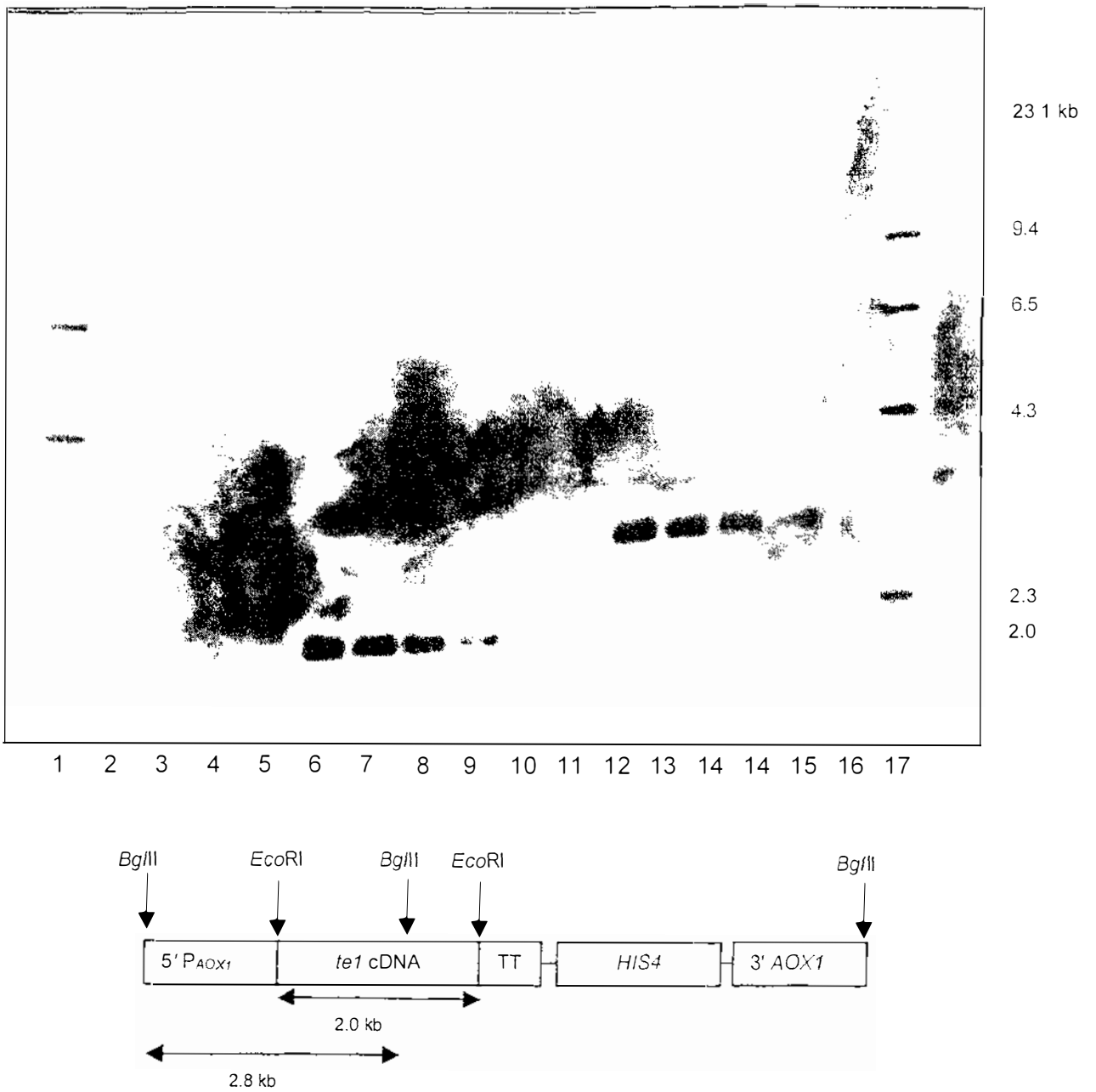
Figure 4.6. SDS-PAGE gel of *Pichia* inductions. The supernatants from three transformed lines of KM71 strain *Pichia* showed no induction of new protein at the expected 73 kD molecular weight. The production of the BGL2 protein, included as a positive control, is clearly visible in lane three. Western blotting of the same extracts also indicated that no significant quantities of the αSF:TE1 protein were produced. Lanes are **1**) molecular weight marker (molecular weights are indicated at left), **2**) BGL2 line, uninduced, **3**) BGL2 induced, **4**) DJ41 line, uninduced, **5**) DJ41 line, induced, **6**) DJ47 line, uninduced, **7**) DJ47 line, induced, **8**) DJ48 line, uninduced, **9**) DJ48 line, induced.

To exclude the possibility that the αSF:TE1 protein was produced, but not excreted, protein extracts of the cell pellets were analsyed by western blotting with the affinity purified anti-MBP:TE1$^{1-286}$ antibody. As with western blots of *E. coli* extracts, multiple bands developed, but no new bands were visible in induced cell pellets that were not present in KM71 and BGL2 lines (data not shown).

Expression trials were conducted with another 16 cell lines that had been transformed with the pPIC9K:*te1* vector (lines DJ1 to DJ16) and had not been selected for high resistance to G418. These 16 lines, the KM71 parent strain, and the BGl2 line were induced as described in Section 4.6h. Samples of the supernatants were pooled into groups of four, and the protein content was concentrated with Nanosep 10K centrifugal concentrators. A sample containing the protein extract from 25 μl of each initial 1 ml of BMMY induction culture was examined by SDS-PAGE and western blotting with the affinity purified anti-MBP:TE1$^{1-286}$ antibody. Although the BGL2 protein was clearly visible by SDS-PAGE, no new bands were visible in either SDS-PAGE gels or western blots (data not shown) from any of the pools. Since the western blotting protocol could detect 10 ng of MBP:TE1$^{1-286}$, the level of αSF:TE1 protein produced was less than 40 μg /L of induction culture.

4.9b. Analysis of the integrations of transformed *Pichia* pastoris lines

The lack of αFS:TE1 protein expressed by the transformed *Pichia* lines led us to question the nature of the integrations. Although the lines that were obtained from the transformation by electroporation could grow on a histidine deficient medium (RDB), and were resistant to the antibiotic G418, it was a formal possibility that the transformation vector had been rearranged or truncated in such a way that precluded expression of the transgene. To address this issue, a Southern blot was produced from transformed lines DJ48, DJ47, DJ44, DJ41 and the KM71 parent strain, loading *Eco*RI and *Bgl* II digests in each case (Figure 4.7). The blot was probed with a DIG-labelled probe produced from the 5' end of the *te1*cDNA with primers SW1 and TE1 as described in Section 2.9e. The Southern blot is shown in Figure 4.7. *Eco*RI digests were expected to produce a 2.0 kb band corresponding to the *te1* cDNA *Eco*RI fragment cloned into the integrating vector pPIC9K. *Bgl* II is present in 5' *AOX1* region (promoter), 1611 bp into the *te1* cDNA, and in the 3' *AOX1* region of the vector, so was expected to produce a 2.8 kb fragment, corresponding to the integrated vector with the *AOX1* promoter, *te1* cDNA and *AOX1* terminator in series. All hybridising fragments were as expected.



igure 4.7. Southern blot of transgenic *Pichia* lines. This Southern blot of four *Pichia* lines transformed ith pPIC9:*te1* (pDJ07), probed with a *te1* cDNA probe (see Section 4.6g), indicated that the vector had ıtegrated without rearrangement. The pPIC9K vector is designed to integrate one or more copies (head-to-ail) at the 3' AOX1 region. The expected integration product of pDJ07 is shown below the blot (TT, anscription terminator). The trend of band intensity in lanes 6-9 and 12-15 of the blot correlates with the ı418 antibiotic resistance of the strains, confirming that they have multiple integrations of the vector. This blot uffered from high background and the scanned image was altered with Photoshop. Lanes are **1**) DIG ıbelled λ *Hind*III ladder (Roche) (molecular weights indicated at right), **2**) 10 pg pPIC9:*te1* vector (not visible ı this rendering), **3**) 1 pg pPIC9:*te1* vector (also not visible), **4**) blank, **5-9) *Eco*RI digests; 5**) KM71 parent ;rain, **6**) strain DJ48, **7**) strain DJ47, **8**) strain DJ44, **9**) strain DJ41, **10**) blank, **11-15) *Bgl*II digests, 11**) M71 parent strain, **12**) strain DJ48, **13**) strain DJ47, **14**) strain DJ44, **15**) strain DJ41, **16**) blank, **17**) DIG ıbelled λ *Hind*III ladder.

Table 4. 2 Resistance of transgenic *Pichia* lines to G418

	G418 (mg/ml)			
Line	0.25	1.0	4.0	Estimated Number of integrations
DJ41	+	-	-	1
DJ42	+	+/-	-	1-2
DJ43	+	+	-/+	2-7
DJ44	+	+	+	>7
DJ45	+	+	+	>7
DJ46	+	+	+	>7
DJ47	+	+	+	>7
DJ48	+	+	+	>7

+	Growth on the stated concentration of G418
+/-	Marginal growth on the stated concentration of G418
2.	No growth on the stated concentration of G418

To confirm that the transforming vector had been constructed as designed so that the vectors α factor secretion signal open reading frame derived form the pPIC9K vector was continuous with the *te1* gene's the insert junction was sequenced using the primer BV36. This sequence confirmed the correct construction of pDJ07 (pPIC9K:*te1*).

4.10. Characterisation of the activity of polyclonal sera

4.10a. Immune sera from both rabbits had activity against the MBP:TE1$^{1-286}$ protein

Enzyme-linked immuno-sorbent assays (ELISAs) indicated that the first collection of immune sera from both rabbits (immune sera A and immune sera B) contained activity against the MBP:TE$^{1-286}$ fusion protein (Figure 4.8). The best ratio of pre-immune sera:immne sera titre was achieved at approximately 1:20 000 dilution for both sera A and B. Further ELISAs indicated that the titre was unchanged after 3 further injections of MBP:TE$^{1-286}$ fusion protein (data not shown).

Figure 4.8 ELISA data from raw sera.
The absorbance at 405 nm of 3 repeats of pre-immune sera, and two of immune sera are shown. For both Rabbit A and Rabbit B, the immune sera can be diluted 1:20 000 before significant reduction in the signal occurs. The difference for each dilution is calculated as [average immune sera signal] - [average pre-immune sera signal].

Rabbit A
Colour development 20 minutes

Sera dilution	Pre-immune sera			average	Immune sera		average	difference
	1	2	3		1	2		
1:100	155	150	205	**170**	1311	1384	**1347**	1177
1:1 000	122	143	125	**130**	1625	1606	**1615**	1485
1:5 000	103	115	527	**248**	1485	1432	**1458**	1210
1:10 000	102	106	169	**125**	1171	1091	**1131**	1006
1:20 000	98	102	104	**101**	1177	1126	**1151**	**1050**
1:50 000	102	97	93	**97**	718	630	**674**	577
1:100 000	95	98	91	**94**	324	327	**325**	231
1:200 000	94	104	86	**94**	353	376	**364**	270

Rabbit B.
Colour development 13 minutes

Sera dilution	Pre-immune sera			average	Immune sera		average	difference
	1	2	3		1	2		
1:100	531	423	448	467	831	700	**765**	298
1:1 000	184	194	301	226	886	1034	**960**	734
1:5 000	116	119	131	122	653	886	**769**	647
1:10 000	121	115	143	126	648	534	**591**	465
1:20 000	104	108	108	106	700	664	**682**	**576**
1:50 000	106	197	88	130	385	354	**369**	239
1:100 000	104	108	90	100	181	189	**185**	85
1:200 000	129	124	107	120	182	175	**178**	58

4.10b. Pre-immune sera B had activity against the MBP protein

To qualitatively characterise the activity of the immune sera, western blots were performed as described in Section 4.2d. Both pre-immune and immune sera A and B

were used as primary antibodies at a dilution of 1:200 (100-fold less than the optimum dilution for ELISA) against membranes that contained the MBP:TE^{1-286} fusion and the maltose binding protein (MBP) that had been produced separately. These western blots indicated that pre-immune sera B had activity against the MBP (Figure 4.9). Because this activity may represent a significant portion of the already low titre of sera B this sera was excluded from further experiments.

4.10c. Antibody activity against MBP and $TE1^{1-186}$ portions of the fusion protein

MBP fusion protein produced from pMAL vectors include a Factor Xa protease signal sequence between the MBP gene and the cloned gene. Pilot experiments indicated that the fusion protein was effectively cleaved by Factor Xa (data not shown), and produced a ~42 kD MBP cleavage product and a ~37 kD $TE1^{1-286}$ cleavage product[6]. To ensure that immune sera A contained activity against the $TE1^{1-286}$ portion of the fusion protein, western blots were performed against separated Factor Xa cleavage products.

Factor Xa-digested MBP:$TE1^{1-286}$ (20μg) was separated into MBP and $TE1^{1-286}$ portions by affinity purification of the MBP portion on an amylose column (NEB). The column wash (flow-through) fractions were expected to contain the $TE1^{1-286}$ cleavage product, and the MBP portion was eluted with maltose. Concentrated column fractions were examined by western blotting (Figure 4.10). No protein was detected in the flow-through fraction by a commercial anti-MBP antibody (NEB), indicating that the MBP fraction was completely retained in the column. The anti-MBP:$TE1^{1-286}$ antibody (sera A) detected a protein from this fraction that was of the expected molecular weight (37 kD) for the $TE1^{1-286}$ peptide. Assuming that the column did not retain any of the $TE1^{1-286}$ peptide, sera A (1:200 dilution) detected 160 ng of this peptide.

[6] The 37 kD $TE1^{1-286}$ cleavage product is predicted to include the following vector-derived 73 amino acids before encountering an in-frame TAG stop codon: GPTSSRVDLQASLALAVVLQRRDWENPGVTQLN RLAAHPPFASWRNSEEARTDRPSQQLRSLNGEWQLGCFGG.

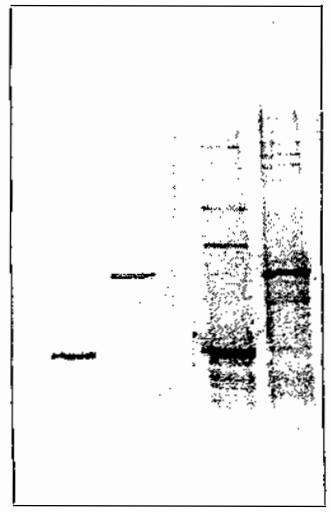


Figure 4.9. Pre-immune sera B showed activity against the MBP. In this western blot pre-immune sera B was the primary antibody for the left panel, and immune sera B for the right panel, both sera were bound at 1:200 dilution. Lanes are **1**) 100 ng of MBP, **2**) 100 ng of MBP:TE1$^{1-286}$.

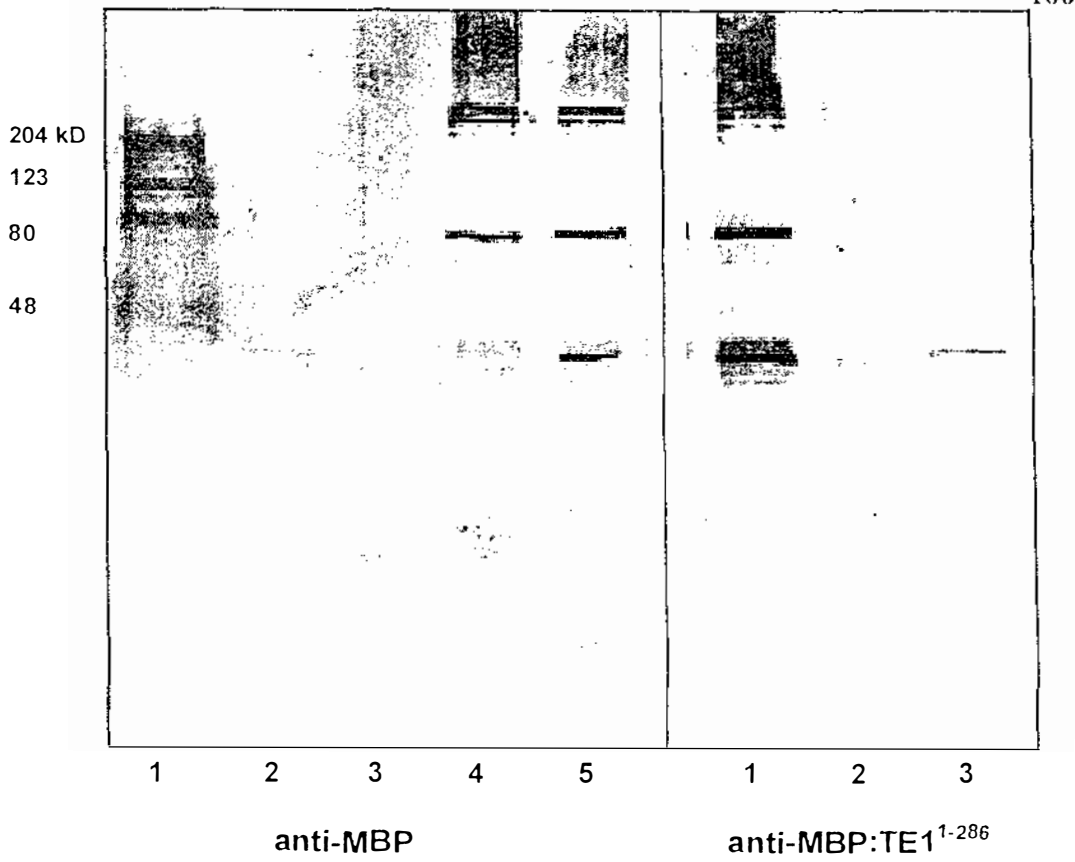


Figure 4.10. Sera A has activity against both MBP and TE1 portions of the fusion protein. The MBP:TE1$^{1-286}$ fusion protein was cleaved into MBP and TE1$^{1-286}$ portions by Factor Xa protease, and the portions separated with an amylose column, which ninds the MBP but allows the TE1$^{1-286}$ to pass through (see Section 4.10c). Western blots were bound with either 1:10,000 anti-MBP antibody (NEB) or 1:200 anti-MBP:TE1$^{1-286}$ immune sera A. **Membrane bound with anti-MBP, 1**) Biorad molecular weight marker (molecular weights at left), **2**) protein eluted from column (MBP), **3**) column flowthrough, **4**) MBP:TE1$^{1-286}$ fusion protein prior to factor Xa digestion, **5**) MBP:TE1$^{1-286}$ fusion protein after partial factor Xa digestion, **Membrane bound with anti-MBP:TE1$^{1-286}$, 1**) MBP:TE1$^{1-286}$ fusion protein after partial factor Xa digestion, **2**) protein eluted from column (MBP), **3**) column flow-through.

4.10d. The affinity purified antibody has activity against all maize tissues

It was known that sera A contained activity against the $TE1^{1-286}$ peptide. To increase the titre, the raw sera was affinity purified against the MBP:$TE1^{1-286}$ fusion protein (Section 4.4b). Western blots indicated that with sera dilutions that developed only a single band (1:6 000 to 1:8 000), the protein detected was present in both normal and *te1-1* mutant SAM, leaf and root tissues (Figure. 4.11). Since *te1-1* transcripts could not be detected by northern blotting (B. Veit, unpublished data), and normal leaf tissues contain at least 100-fold less *te1* transcript than SAM tissues (Section 2.12), the TE1 protein is not expected to be present in these tissues. At lower dilutions of the sera, or with longer exposures of the substrate to film multiple bands were produced (not shown). No band(s) could be identified that were unique to normal maize extracts. The affinity purified antibody therefore contains some activity against a ubiquitous maize protein or proteins, and could not be used for experiments that required specific detection of the TE1 protein.

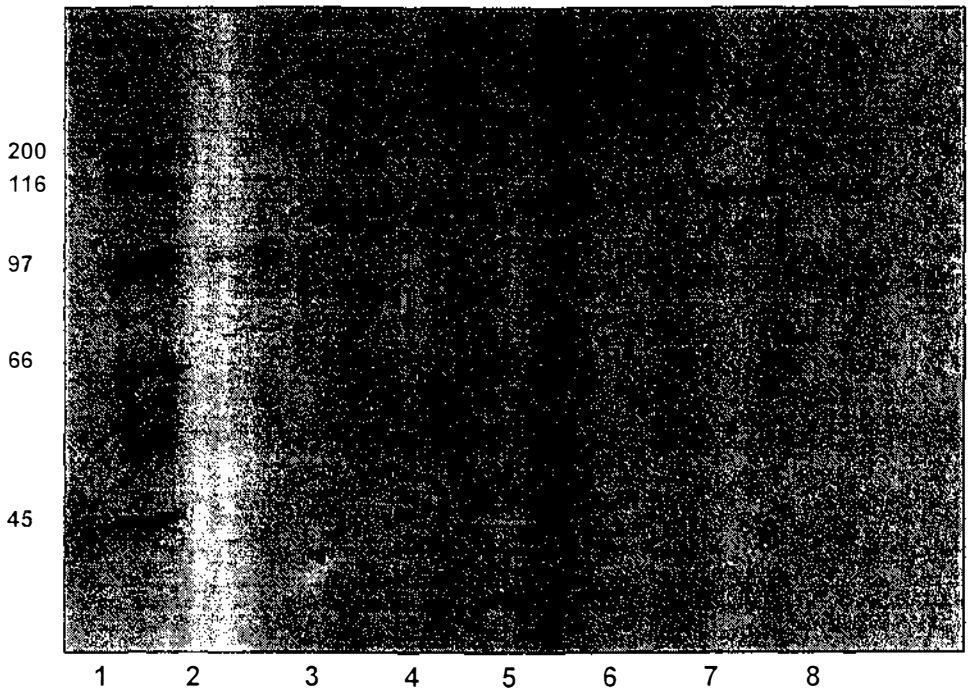


Figure 4.11. The affinity purified antibody had activity against all maize tissues examined. This western blot developed as described in Section 4.2c, the affinity purified anti-MBP:TE1$^{1-286}$ primary antibody was bound at 1:6,000 dilution. Note that the antibody detects at least a protein of the same size in all tissues (lanes 3-8). Lanes are; **1**) Biotinylated molecular weight marker (Biorad) (molecular weights indicated at left), 2) 1 ng MB:TE1$^{1-286}$ protein, **3**) normal (B73) SAM, **4**) *te1-1* mutant SAM, **5**) normal leaf, **6**) *te1-1* leaf, **7**) normal root **8**) *te1-1* root. 10 μg of protein was loaded for each tissue.

DISCUSSION (4)

4.11. Why was the TE1 protein not produced in *E. coli*?

Only the N-terminal portion of the TE1 protein, encoding the first two RRMs of the protein, could be expressed in *E. coli* without the *argU* tRNA gene. Both MBP fusions and histidine-tagged versions of this peptide could be expressed[7]. We first considered two possibilities for the failure of *E. coli* to produce the full length TE1 protein: either the TE1 protein was toxic due to non specific RNA binding activity, or the protein was not efficiently translated, or was rapidly degraded. Since *E. coli* cultures did not appear to grow any more slowly after induction with IPTG in the presence of antibiotic selection for the vector, we focused on codon bias.

The assumption was that a low codon frequency in a particular host indicated that the population of the corresponding tRNA was low. The shortage of $tRNA_{arg4}$ has been shown to cause ribosomal frame shifts at tandem AGA and AGG codons (Spanjaard 1990). The arginine codons AGG and AGA are recognized by the same tRNA (product of the *argU* gene). The *te1* cDNA contains nine AGG codons, including one instance of two consecutive AGG codons (AGG AGG). The argU gene (encoding the tRNA with anticodon AGG) was used to overcome potential problems with the translation of the *te1* gene caused these rare codons.

The first construct (pDJ06), failed to improve the expression levels. This construct included the *E. coli argU* gene that was amplified form the parent vector pSBET by PCR and cloned into another expression vector (pPROEX). There are some issues concerning the construction of pDJ06. The *argU* PCR product included the argU transcript and native *E. coli* promoter from position -44 to (where the transcription start site is +1) to +230bp. While this region includes the minimal promoter and will initiate transcription in *E. coli* (Saxena 1992), it does not include the upstream activation sequence that extends to position -238. Studies of *argU* promoter elements (Saxena 1992) indicated that this region increased transcription four fold. Another issue was the production of the *argU* insert by PCR. It is entirely plausible that a

[7] It was later found that another construct containing the 5' portion of the *te1* cDNA encoding the first 286 amino acids of the TE1 protein also produced a protein of the expected size. The *Eco*RI-*Sma*I fragment of the *te1* cDNA was cloned from the vector pBV246 (a subclone of the *te1* cDNA

single nucleotide mutation during PCR or cloning could cause the tRNA gene to be non-functional. The single cloned *argU* PCR product that was cloned was not sequenced, so we cannot discount the possibility that the pDJ06 vector was used for expression experiments contained a non-functional *argU* gene.

Once the *Bam*HI fragment of the *te1* cDNA was cloned directly into pSBET, which contains the full *argU* promoter, production of the TE1 protein increased dramatically (compare Figure 4.4 to Figure 4.5). The 577 amino acids of the TE1 protein produced from this construct contain eight of the nine AGG codons, and all 3 RRMs (Figure 4.12). The pSBET vector, which contains the ArgU tRNA gene transcribed from its native *E. coli* promoter, appeared to solve the problem. This information has two corollaries. Firstly the protein with all 3 RRMs does not appear to be toxic to *E. coli*, even in relatively large quantities. Secondly, as discussed above, the previous *argU* PCR cloning strategy seems to have been flawed. The full length protein was cloned into this vector due to lack of time.

The commercial development[8] of transgenic *E. coli* containing integrations of both the *argU* gene (AGG, AGA codons), and also isoleucine (AUA), leucine (CUA), and proline (CCC) tRNA genes has made vectors containing such tRNA genes unnecessary. High quantities of the *Arabidopsis* TEL2 protein has been produced in such a cell line (V. Trainor, unpublished).

---

vector pBV432) into pPROEX by Carmel Gillman. No expression was seen from a construct expressining the C-terminal 578 to 655 amino acids.

[8] Stratagene BL21-CodonPlus™ strain.

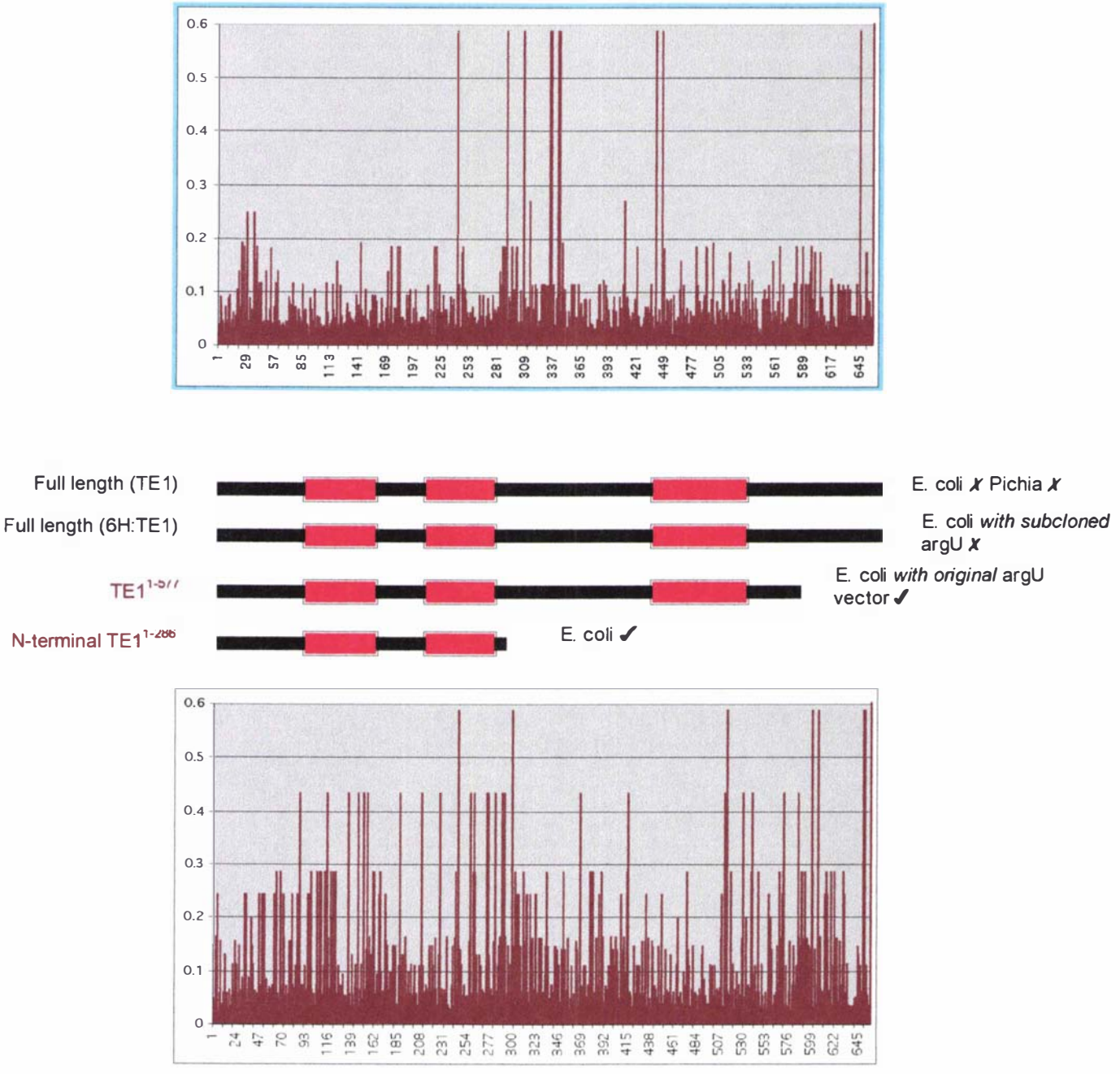


Figure 4.12. The rarity of *te1* codons in heterologous hosts. The two graphs plot the inverse frequency (rarity) of each *te1* codon in *E. coli* and *Pichia pastoris*. Representations of the various portions of the TE1 protein that were expressed (red font), or were not (black font) are shown between the graphs. All the codons above 0.4 rarity are AGG arginine codons, for which ArgU is the tRNA.

4.12. Why was the TE1 protein not produced in *Pichia pastoris*?

It seemed with *E. coli* that successful complementation of the tRNA pool by the *argU* gene improved expression. Given that there are 29 codons in the *te1* gene that are above the 0.4 level of 'codon rarity' in *Pichia pastoris*, including 3 instances of consecutive rare codons (Figure 4.12), it is not surprising that the TE1 protein was not produced. It is clear that codon rarity can preclude heterologous protein expression.

4.13. The 'non specific' activity of the polyclonal antibodies

While the affinity purified sera contained some activity against the $TE1^{1-286}$ protein (Section 4.10c), the development of bands in tissues that contain the significantly lower levels of the *te1* transcript indicated that the antibodies were cross reacting with some other maize protein(s). There are two ways that this problem could be solved. Firstly the remaining sera could be affinity purified against a TE1 fusion protein that does not contain the MBP, so that antibodies with affinity to the MBP will be lost, and conversely anti-TE1 activity may increase.

Secondly, the antibody was raised against the N -terminal region of the TE1 protein, which includes the first two RRMs. BLAST searches indicate that this region is most similar to 'canonical' two-RRM RNA binding proteins as described in Birney (1993), including splicing factors (Figure 3.1), and are present in all eukaryotes. Comparisons of the structures of seven RRM domains whose structure have been solved indicates that this RRM domain is extremely highly conserved, with a number of absolutely conserved residues (Figure 1.2, S Moore pers. comm.). These proteins are likely to be present in much higher abundance than the TE1 protein and so even low affinity of the antibody for these proteins could overwhelm a weak signal for the TE1 protein. The C-terminal RRM domain is less similar to canonical RRM proteins, so may be a better substrate for antibody production.

The structure-based alignments of *Mei2*-like proteins yield good predictions for the locations of the variable, solvent exposed loops of the proteins (Chapter 3). These regions are the most suitable regions to raise or affinity purify antibodies against.

4.14. Concluding remarks and future prospects

Two significant advances have made the study of TEL group protein function much more tractable: the completion of the *Arabidopsis* genome, which allowed *TEL1* and *TEL2* cDNAs to be cloned relatively easily, and the development of codon supplemented *E. coli* strains, which allowed a histidine-tagged TEL2 protein to be expressed routinely (V Trainor, unpublished data). The ease with which *Arabidopsis* can be transformed may also obviate the need to raise antibodies against such proteins; the main applications of such antibodies (immunolocalisation and immunoprecipitation) can be performed by other methods such as GFP fusions and the TAP construct, which allows proteins and bound protein cofactors to be purified from plant extracts (Rigaut 1999).

CHAPTER 5: ECTOPIC EXPRESSION OF *TEL* GENES IN *ARABIDOPSIS*

5.1 ABSTRACT

Two *TEL* group genes have been ectopically expressed in *Arabidopsis*, both with contingent nuclear localisation. The maize *te1* gene was expressed under the control of a copper-inducible promoter, both with and without the SV40 large T antigen nuclear localisation signal (NLS). No abnormal phenotypes that consistently correlated with the expression levels of either the *te1* transgene or the *NLS:te1* transgene were observed. The *Arabidopsis TEL2* gene was expressed from a constitutive promoter, as a glucocorticoid receptor (GR) fusion protein, TEL2:GR. Glucocorticoid fusions are localised to the nucleus in the presence of the synthetic hormone dexamethasone. Preliminary studies have not identified any differences between dexamethasone treated TEL2:GR plants and untreated TEL2:GR plants. The lack of phenotypes from the expression of these genes could be due to technical problems with vector construction or transformation, or to biological factors such as genetic redundancy in the function of TEL group proteins, divergence of structure or function of TEL group proteins over time, or the requirement for some accessory factor(s) for gene function.

METHODS (5)

5.2 Construction of vectors for transformation of *Arabidopsis.*

5.2a. Construction of copper-inducible transformation vectors

The copper inducible promoter system was developed in *Nicotiana* to allow inducible gene expression in plants (Mett 1993). The transformation vector constitutively express the *S. cerevisiae ACE1* gene, that activates transcription from a chimeric promoter only in the presence of μM concentrations of $CuSO_4$. The promoter was contstructed from a fusion of the *S. cerevisiae* metalloregulatory element (MRE) to which the ACE1 protein binds in yeast and the CaMV 25S promoter.

Two vectors in the series were used (Figure 5.1); pPMB7066 is a pUC119 derivative containing a chimeric promoter consisting of one copy of the *Saccharomyces cerevisiae ACE1* transcription factor binding site (MRE) fused to the 90 bp 35S promoter, a short polylinker and a *nos* terminator, and pPMB765 which contains the elements required for *Agrobacterium* mediated plant transformation. The open reading frame of the *te1* cDNA was cloned into pPMB7066 as an *Xba*I *Kpn*I fragment from pBV432 to create pPMB7066/*te1.* The MRE - *te1* - nos 'casette' was cloned into the *Not*I site of pPMB765 by partial digestion of pPMB7066/*te1* with *Not*I, isolation of the 2.8 kb digestion product from a gel, gel extraction and ligation as described in Section 2.3e. This created pDJ09 and pDJ11, which have the MRE - *te1* - nos casette in the same orientation as the *ACE1* gene, and PDJ8 and pDJ10 which have the MRE - *te1* - nos cassette in the opposite orientation. These vectors were screened with informative *Hind*III, *Not*I, and *Xba*I/*Kpn*I digests. The junctions between the MRE/35S promoter, and the *te1* cDNA of pDJ10 and pDJ11 were sequenced as described in Section 2.4c using the BV36 primer.

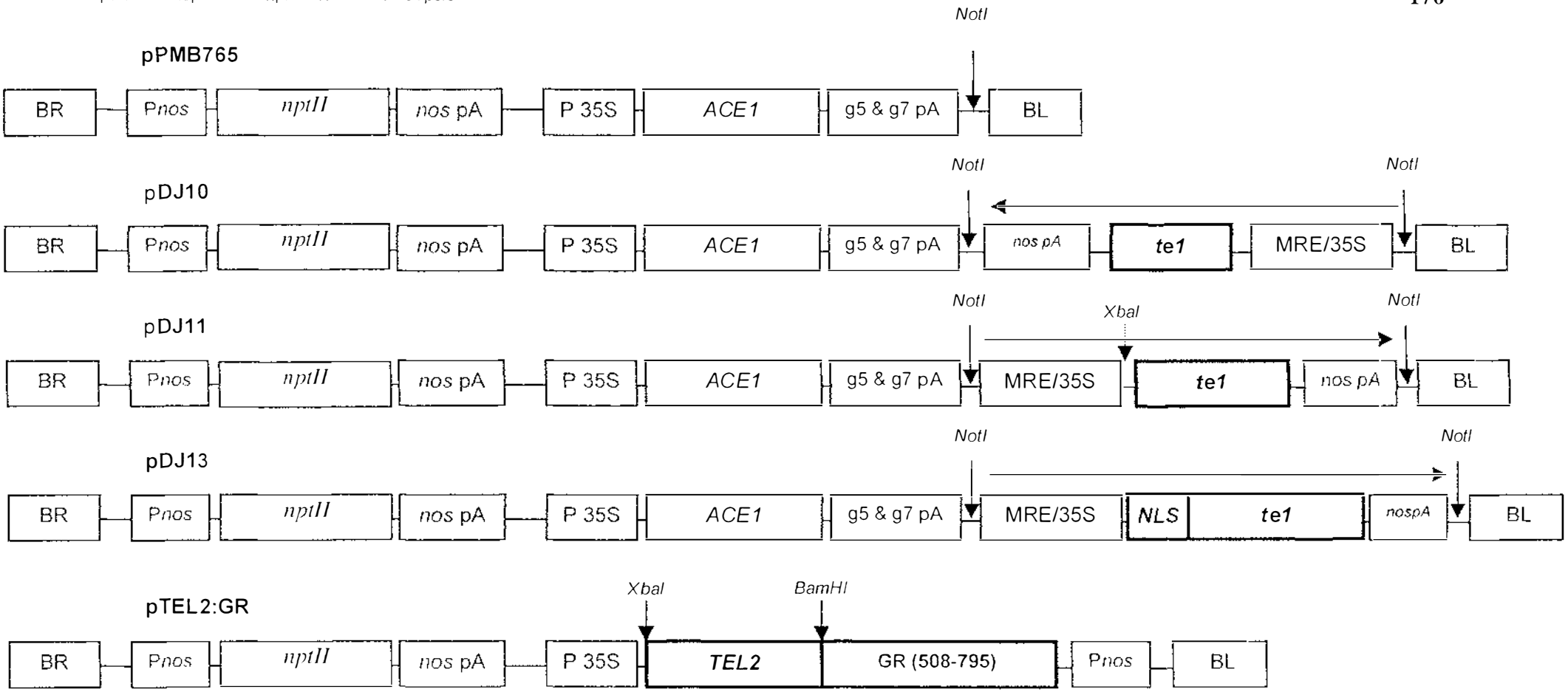


Figure 5.1. Vectors constructed for expression of *te1* and *TEL2* in *Arabidopsis*. Only T-DNA portions of vectors are shown. Other details of pPMB765 and pPMB7066 Mett (1993), and Aoyama (1997) for the glucicorticoid fusion vector. The pPMB765 T-DNA contains only the *NptII* gene (conferring kanamycin resistance in plants), and the *S. cerevisiae ACE1* gene under control of the constitutive 35S promoter. The pDJ10 and pDJ11 T-DNAs consist of the pPMB765 T-DNA with the {MRE/35S promoter - *te1* cDNA - nos terminator} cassette cloned into the NotI site in either the same orientation as the *ACE1* gene (pDJ11) or the in the opposite orientation to the *ACE1* gene (pDJ10). In each case the *te1* gene is correctly positioned to be transcribed form the MRE/35S promoter. The pDJ13 T-DNA is a derivative of pDJ10 with an oligonucleotide encoding the SV40 NLS cloned into the XbaI site (dashed line) in frame with the *te1* reading frame. The TEL2:GR vector T-DNA consists of the *NptII* gene and a 35S-driven fusion of the TEL2 gene and of the a portion of the glucicorticoid receptor gene (GR) encoding the amino acids 508 to 795.

During the course of he research we considered whether the TE1 protein functioned in the nucleus, or the cytoplasm, or both. We chose to examine the effect of obligatory nuclear localisation. To attach an N-terminal nuclear localisation signal (NLS) to the TE1 protein expressed form the copper-inducible construct, a double stranded DNA oligonucleotide encoding the SV40 large T antigen NLS, (amino acids MAPKKKRKV, Kalderon 1984), was inserted into the pPMB765/*te1* constructs as follows. The complementary NLS_SENSE and NLS_ANTI oligonucleotides (Appendix 1) were boiled for 5 minutes at 20 pmol/µl concentration and allowed to cool to room temperature to anneal, then 40 pmol of this annealing reaction was phophatased with 0.8 units of heat labile alkaline phosphatase (TsAP, Gibco BRL). This short DNA fragment was ligated[1] into the *Xba*I site of pDJ10 and pDJ11. Potential colonies were screened from STET plasmid preparations (Section 2.2g) by PCR with the primers NLS_SENSE and a reverse primer close to the 5' end of the *te1* cDNA (BV36) in a standard PCR reaction (Section 2.3c) using the thermocycling program DJ53x40. The MRE/35S promoter - *te1* cDNA junction (including the NLS), of the vectors pDJ14 and pDJ15, were sequenced as above. Sequencing confirmed that the ATG start codon was present in the NLS oligonucleotide insert, and the reading frame was continuous with the *te1* frame.

5.2b. Construction of glucocorticoid transformation vectors

This vector was constructed by Carmel Gillman. The *TEL2* cDNA was cloned into the *Xba*I and *Bam*HI sites of the glucocorticoid fusion vector pBI-ΔGR using a PCR strategy. The TEL2 cDNA was amplified from 500 ng of pVTTEL2 using the primers GLUX-5 and GLUB-3 which span the *TEL2* open reading frame, and contain an *Xba*I and *Bam*HI linkers respectively (see Appendix 1). A PCR program equivalent to DJ55x9 was used in a standard PCR reaction (Section 2.3c) , except that a high fidelity polymerase was used (Roche High Fidelity). The TEL2 PCR product was cloned into pGEM-T Easy (Promega), and the *Xba*I *Bam*HI insert of this vector was then cloned into pBI-ΔGR to create pBIΔGR:TEL2B (see footnote[2]). Initial *E. coli* colonies were confirmed by colony PCR with GLUX-5 and GLUB-3 primers (Section 2.3d), as were *Agrobacterium* transformants. Neither the sequence of the PCR product, or the vector-insert junction were verified by sequencing, but the large

[1] The annealed NLS oligonucleotides were phophatased to prevent multimers forming, ligation reactions contained the NLS oligonucleotide in 1:100 molar excess over the *Xba*I digested vector (not phosphatased).

2 pBI-ΔGR:TEL2A was also constructed from an initial *TEL2* cDNA clone that with a deletion relative to the *TEL2B* clone that was subsequently recovered. The *TEL2B* clone is considered to be full length cDNA (V. Trainor unpublished).

quantity initial cDNA vector template (500 ng), low cycles (9), and high fidelity polymerase were designed to reduce the chances of point mutations.

5.3 Transformation of *Agrobacterium*

The *Agrobacterium tumefaciens* strain GV3101, was maintained on YEB plates (1.5% beef extract, 0.3% yeast extract, 0.3% peptone, 1.5% sucrose, 1.5% agar), with 20 μg/ml gentamycin (YEB Gent^{20}) to maintain selection for the binary vector pMP90. Electrocompetant cells were prepared as for *E. coli* (Section 2.2d), except that YEB Gent^{20} media was used, and cultures were grown at 28°C. Cells were transformed as described for *E. coli* (Section 2.2e) (25 μF, 2.5 kV, 200 Ω), except that cells were 'rescued' in YEB Gent^{20} for 3 hours at 28°C, before being plated onto YEB Gent^{20} plates containing an antibiotic to select for the T-DNA vector (100 μg/ml kanamycin for pPMB765 vectors, 50 μg/ml kanamycin for the pBI-ΔGR vector).

To confirm that pPMB765-based vectors were not rearranged by the *Agrobacterium tumefaciens* strain GV3101, the vector pDJ10 (pPMB765/*te1*) was isolated from this strain by alkaline lysis exactly as for *E. coli* (Section 2.2f). So that a range of individual vector molecules could be examined separately, a sample of this plasmid preparation was transformed into *E. coli* by heat shock. Single colonies were selected, and eight small scale preparations of plasmid were prepared. These samples were compared by informative *Not*I, *Xba*I/*Kpn*I, and *Hind*III digests to a pDJ10 sample that had not been propagated in *Agrobacterium tumefaciens* GV3101. In all cases the digestion products were the same.

5.4 Growth of *Arabidopsis*

Seed was sown onto wet seedling raising mix (Yates Black Magic) covered with a plastic bag, and vernalised for two days at 6°C. Pots were transferred to a growth chamber set to 23°C, 65% relative humidity, and either an eight or a 16 hour photoperiod with 300-400 μmoles m^{-2} s^{-1} PAR light. Plastic covers were removed from pots once cotyledons had emerged.

5.5 Transformation of *Arabidopsis* by vacuum infiltration

5.5a. Overview of the procedure

This protocol involves dipping flowering *Arabidopsis* shoots into a suspension of *Agrobacterium* (Bechtold 1993). The *Agrobacterium* transform approximately 0.2% of ovules (Ye 1999), to produce hemizygous transgenic seed, which are selected by germinating seed on an appropriate antibiotic. Many variations in the protocol are used, the most important parameter appears to be health and developmental stage of the plants.

In this thesis plants that are dipped into *Agrobacterium* cultures are referred to as the T0 generation, primary transformants as the T1 generation, and the progeny as T2 generation. T2 plants were produced by allowing T1 plants to self fertilise. Because primary transformants are usually hemizygous with 1 or 2 inserts of the T-DNA, T2 plants are expected to segregate at approximately 3:1 or 15:1 for the presence of the T-DNA.

5.5b. Vacuum infiltration

Arabidopsis thaliana Columbia ecotype plants were sown into pots covered with a nylon mesh, so that the plants grew up between the mesh. Plants were grown with short days until the inflorescences were 1-10 cm long. Primary inflorescences were then removed to reduce apical dominance, allowing secondary inflorescences to grow. Plants were infiltrated with *Agrobacterium* 4-6 days after clipping.

A 1.5 L YEB Gent[20] culture of *Agrobacterium tumefaciens* strain GV3101 carrying the binary vector was grown at 28°C with antibiotics to select for the T-DNA plasmid until mid log phase (see footnote[3]) (OD_{600} ~ 0.8). Cells were harvested by centrifugation at 4200 xg, and resuspended in 1 L of infiltration media (0.5 x MS salts, 1 x B5 MS vitamins, 5% sucrose, pH 5.8), followed by addition of 50 µl of Silwet L77 (Lehle Seeds). Plants (T0) were upturned in the *Agrobacterium* suspension and a vacuum drawn in a dessicator until the solution boiled for 5 minutes (see footenote[4]). Plants were then stored overnight in a plastic covered tray,

---

[3] Growth of *Agrobacterium* to mid log phase takes approximately 24 hours when started with a 1/1000 dilution of a stationary phase culture.

[4] Later experiments showed that drawing a vacuum is not necessary.

and the cover removed over 3-5 days. Plants were allowed to self fertilise. The vectors that were transformed into *Arabidopsis* are described in Table 5.1.

Table 5. 1 Transgenic *Arabidopsis* constructs

Type	Vector	Construction	Lines generated*	Notes
Copper-inducible	pPMB765	T-DNA contains only *NptII* and *ACE1* genes (Mett 1993). No transgene.	6	Empty vector control
	pDJ10	pPMB765 with [MRE - *te1* cDNA - nos] cassette in 'reverse' orientation. Transgene *te1* cDNA.	23	
	pDJ11	pPMB765 with [MRE - *te1* cDNA - nos] cassette in 'forward' orientation. Transgene *te1* cDNA.	4	
	pDJ13	pPMB765 with [MRE - NLS:*te1* cDNA - nos] cassette in 'reverse' orientation A derivative of pDJ10. Transgene NLS:*te1* cDNA.	9	
GR receptor	pBI-ΔGR:TEL2B	pBI-ΔGR with *TEL2* GLUX-5/GLUB-3 PCR product. Transgene *TEL2:ΔGR*.	3	expression not confirmed

* The number of lines examined for abnormal phenotypes. Less lines have been examined on the molecular level.

5.5c. Selection of transformants

Plants transformed with the copper inducible vector, pPMB765, were selected on MS media lacking $CuSO_4$ containing 100 μg/ml kanamycin and 100μg/ml cefotaxime (MS-Cu kan^{100} $cefo^{100}$). Plants transformed with the glucocorticoid fusion vector pBIΔ-GR were selected on 50 μg/ml kanamycin. T1 seed, were sterilised by rinsing with 80% ethanol for 1 minute, 7.5 % sodium hypochlorite (50% domestic bleach), 0.05% Tween 20 for 5 minutes, and rinsing three times in sterile Milli-Q water. Seed were resuspended in warm 0.5% agar, and poured an MS kan^{100} $cefo^{100}$ plate containing 0.8 % agarose. Up to 2000 seed (40 μg) were germinated per plate. To repress fungus growth 20 μl of fungicides 0.05% Benlate, 0.15% Chlorothalonil were applied to the surface of MS kan^{100} $cefo^{100}$ media before addition of seed. Seeds were germinated at 23°C with either an 8 or 16 hour photoperiod until antibiotic resistant seedlings were clearly distinguished from susceptible seedlings. Resistant seedlings (T1 plants) were transferred to soil and allowed to self fertilise, giving rise to T2 seed.

5.6 Induction conditions for plants with copper-inducible transgenes

5.6a. Induction in synthetic media

It was suspected that expression of the *te1* transgene may terminate development of the SAM, therefore plants were initially sown on MS kan^{100} media containing either 0.1, 5, and 50 µM CuSO4. Plants segregated for kanamycin resistance with the expected ratios. Lines carrying either *te1* transgene or *NLS:te1* transgene were compared to 'empty vector' lines that had been transformed with the parent vector containing no transgene.

5.6b. Induction in soil

Once it was determined that expression of the *te1* transgene was not lethal, plants were grown without selection in soil watered from below with 5 µM CuSO4, and sprayed with 50 µM CuSO4. The numbers and arrangements of vegetative leaves, cauline leaves, and floral parts were examined.

5.7. Dexamethasone application to TEL2:glucocorticoid receptor plants

Three primary TEL2:GR transformants were obtained from initial kanamycin screening (as in Section 4.5c). These primary transformants were allowed to self fertilise, producing the T2 generation. Two trays were sowed with approximately 100 T2 seed of each line, and approximately 100 Columbia seeds as described in Section 5.4. As soon as cotyledons were visible, one tray was sprayed on alternative days with 30 µM dexamethasone (made from a 30 mM dexamthasone (Sigma), stock solution in ethanol, stored at -80˚C), 0.01 % Tween-20, 0.1% ethanol. The control tray was sprayed with 0.01 % Tween-20, 0.1% ethanol. Plants were sprayed until most plants had begun to develop inflorescence shoots.

5.8. Assays of transgene expression

5.8a. Northern blotting

Total RNA was extracted from the aerial portions or leaves of *Arabidopsis* plants as described in Section 2.7c (simple version), electrophoresed and blotted as described in Sections 2.8a and 2.8b. Since the MRE/35S promoter appears to express genes constitutively in *Nicotiana*, root tissue (Vadim Mett, pers. comm.) roots were excluded from expression analysis of Arabidopsis transformants. At least one 'empty vector' line was included on each blot. So that the intensity of bands could be compared between experiments, lanes containing 100 pg, and 10 pg of a synthetic *te1* cDNA transcript were included. Membranes were hybridised in Church/Gilbert hybridisation buffer described in Section 2.8d using radiolabelled probes (see Section 2.8c for a description of probe synthesis, and probe templates) made from the full length *te1* cDNA. Equal loading of samples was shown by either re-probing blots with a probe synthesised from the *ROC1* cyclophilin gene, which is constitutively expressed in *Arabidopsis* (Lippuner 1994), or by visualising rRNA bands in gels by staining with ethidium bromide prior to blotting.

5.8b. Western blotting

Protein extracts were prepared from the aerial portions/leaves of *Arabidopsis* plants using the EZ protein extraction method (Marínez-García 1999). Tissue was ground in E buffer (125 mM Tris-HCl pH 8.8, 1% SDS, 10% glycerol, 50 mM $Na_2S_2O_5$), transferred to an eppendorf tube and stored on ice until all other samples were ground. Extracts were then warmed to room temperature, and centrifuged at 20,000 xg in a benchtop microfuge. The supernatant was removed, and sample (10 μl) was set aside for protein content estimation, and 1/10th volume of Z buffer (125 mM Tris-HCl pH 6.8, 12 % SDS, 10% glycerol, 22% 2-ME, 0.001% bromophenol blue) was added to the remainder.

Protein content was estimated by the Lowry method (Lowry 1951) as follows; solution D was freshly prepared with 48 ml solution A (2% Na_2CO_3 in 0.1 N NaOH), 1 ml solution B (1% NaK tartrate in water), 1 ml solution C (0.5% CuSO4.5H_2O in water). 95 μl of solution D was added to protein samples of 95 μl volume in microtire plate wells, and they were incubated for 10 minutes at room temperature. 10 μl of a 1:3 dilution of Folin reagent (in water) was added and incubated at room temperature for 30 minutes. The absorbance of each sample was read at 600 nm, and compared to a BSA standard curve.

Protein extracts were electrophoresed and blotted as described in Sections 4.2b and 4.2c. Typically 10 µg of protein was loaded, a negative control lane containing a protein extract from an empty vector line was always included. Western blots were performed by the chemiluminescent method (Section 4.2c) using the affinity purified anti-MBP:TE1$^{1-286}$ antibody at 1:600 dilution.

RESULTS (5)

5.10. Initial analysis of copper-inducible transformants

5.10a. Transgenic plants transformed with copper inducible constructs contained one or two integrations of T-DNA

T2 generation plants segregated for kanamycin resistance with ratios of close to either 3:1 or 15:1 (Table 5.2), indicating that the T1 lines were hemizygous with either one or two integrations of the T-DNA. To confirm that the *te1* was integrated as well as the *NptII* gene which confers kanamycin resistance, a Southern blot was prepared including nine of the initial pDJ0 transformants. A *te1* probe indicated the *te1* portion of the T-DNA was present in all lines that were examined (Figure 5.2). Since the pDJ10 T-DNA contains a unique *Xba*I site, and a unique *Xho*I site, tandem head-to-tail insertions result in a hybridising band of the size of the T-DNA (7.4 kb) for each of these digestions. Transformants 30.6 (lanes 7 and 8 in Figure 5.2), and 33.13 (lanes 19 and 20) appeared to contain tandem insertions.

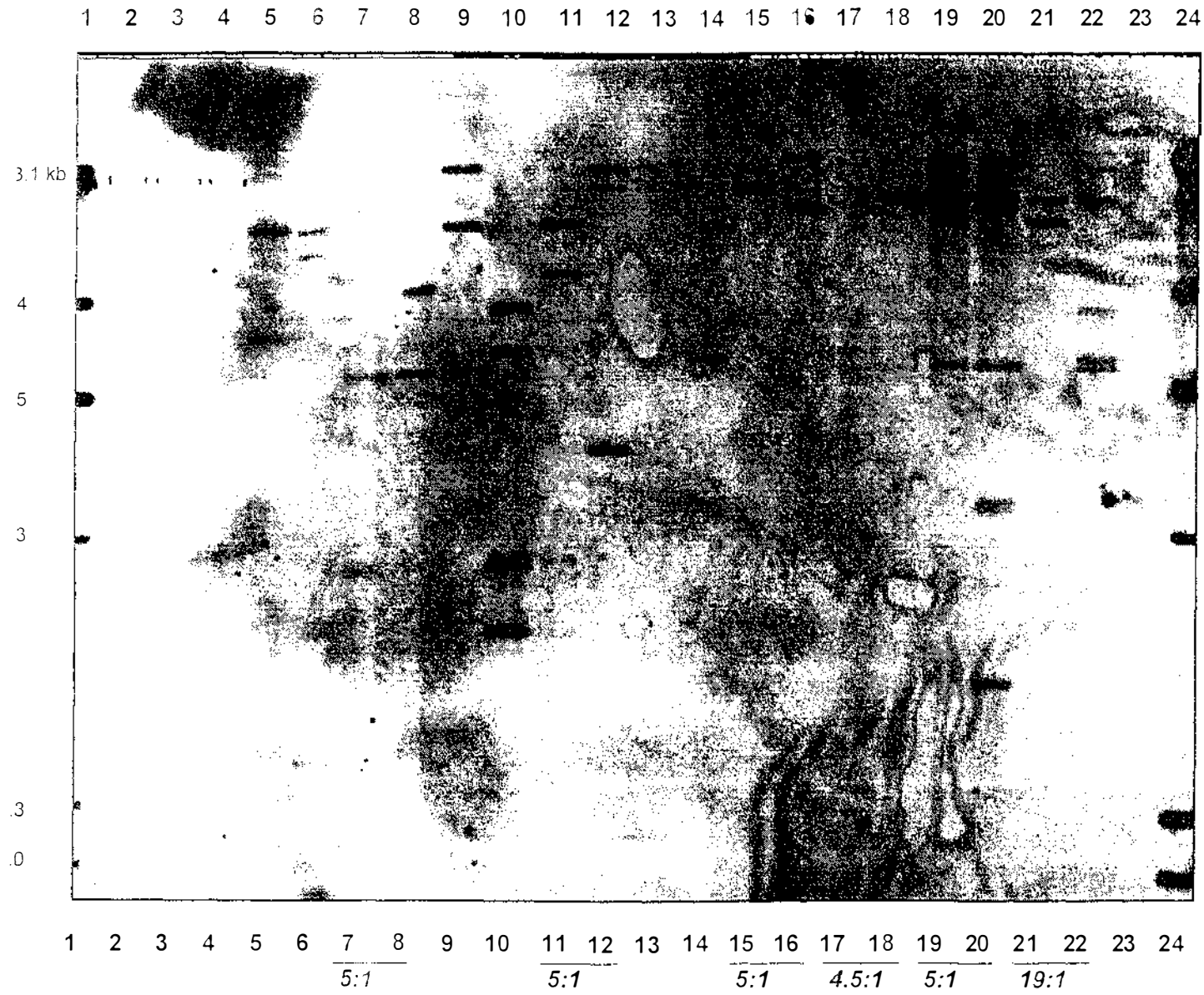


Figure 5.2. Southern blot of pPMB765/*te1* **transgenic lines confirmed independent T-DNA insertion.** DNA extracts (100 ng) from ten kanamycin resistant pDJ10 (pPMB765/*te1*) primary transformants and Columbia are included in this Southern blot. All samples were digested with XbaI and with XhoI, both of which cleave the T-DNA once. The blot was probed with a DIG labelled *te1* probe. Although some samples appear to be partial digests, this blot confirms that kanamycin resistant lines contain independent insertions of the T-DNA, including the *te1* gene. Where the ratio of kanamycin resistant seed in the T2 progeny was established, it is indicated in italics below the lane numbers. Lanes are **1)** DIG labelled λHindIII molecular weight marker (partly obscured) (Roche) (molecular weights are indicated at left), **2)** blank, **3)** Columbia XbaI, **4)** Columbia XhoI, **5)** 30.4 XbaI, **6)** 30.4 XhoI, **7)** 30.6 XbaI 8) 30.6 XhoI, **9)** 30.17 XbaI, **10)** 30.17 XhoI, **11)** 30.20 XbaI 12) 30.20 XhoI, **13)** 31.11 XbaI **14)** 31.11 XhoI, **15)** 31.5 XbaI **16)** 31.5 XhoI, **17)** 33.12 XbaI 18) 33.12 XhoI, **19)** 33.13 XbaI **20)** 33.13 XhoI, **21)** 33.17 XbaI **22)** 33.17 XhoI **23)** empty, **24)** DIG labelled λ HindIII molecular weight marker.

5.10b. The *te1* transgene was expressed in *Arabidopsis* from the copper inducible promoter.

The copper-inducible promoter system was developed and tested in *Nicotiana* (Mett 1993). The authors found that foliar application of 0.5 μM $CuSO_4$ to *Nicotiana* transformants was sufficient to induce a 50-fold increase in GUS reporter gene expression after 5 days. To ensure that high levels of expression were obtained, T2 plants containing both the *NLS:te1* and *te1* transgenes were grown in soil and sprayed daily with 5 μM $CuSO_4$. Since these populations were segregating for the presence of the T-DNA, the aerial parts of at least six T2 plants were pooled for examination by northern blotting. When 5 μM $CuSO_4$ was applied the transgenes were expressed to the order of ~1 pg transcript/μg of total RNA (Figure 5.3), an order of magnitude higher than *te1* expression in maize embryos or SAMs (see Section 2.12).

The relationship between Cu(II) ion concentration and expression levels was not examined. The data produced by Mett *et al.* (1993) indicated that the MRE/35S promoter produced low levels of expression in the absence of applied copper, and this effect has been observed in *Arabidopsis* (H Zhang, pers. comm.). Experiments were therefore designed that did not assume expression was completely repressed in the absence of copper; plants transformed with empty vector (pPMB765), containing no transgenes other than *NptII* and the *ACE1* gene, were always used as controls.

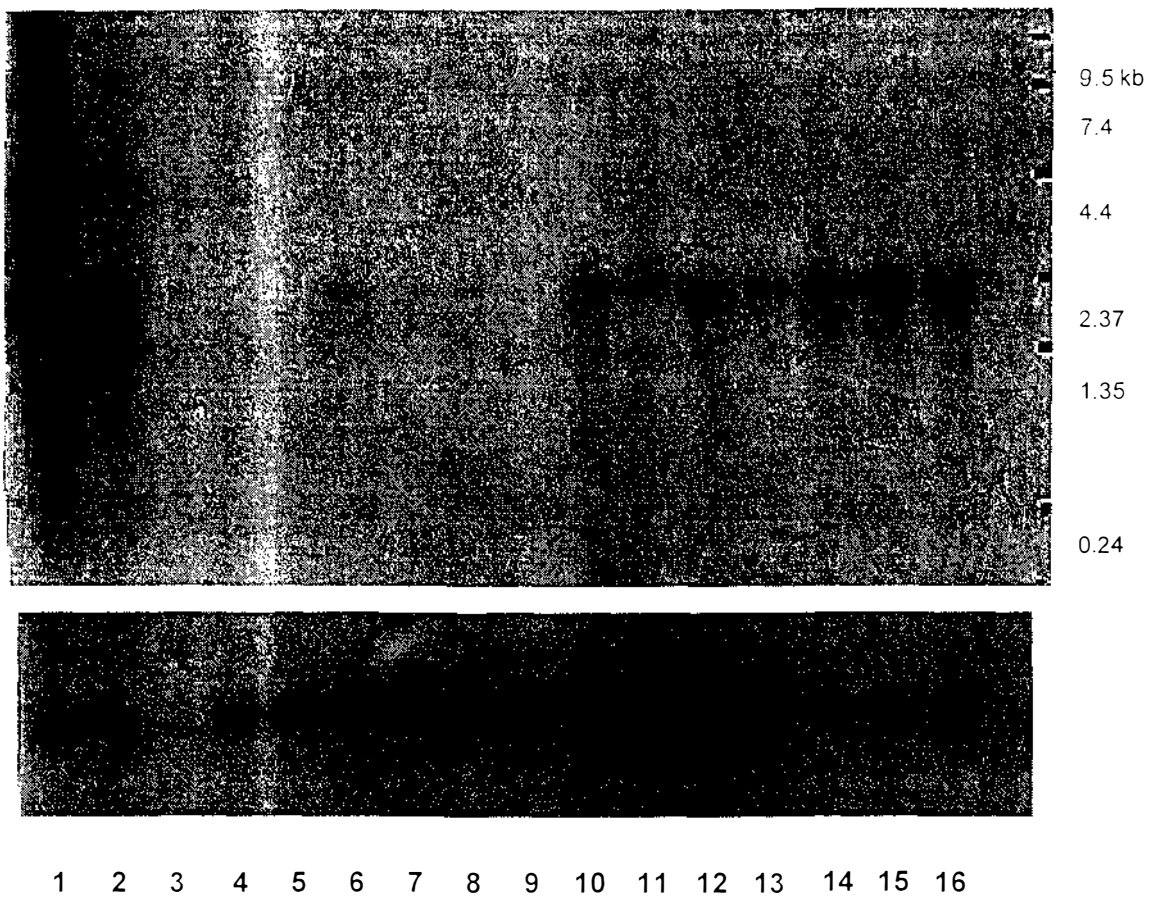


Figure 5.3. Northern blotting indicated that transgenic lines expressed the *te1* transgene at various levels. Transgenic lines expressed the *te1* transgene at different levels. T2 plants arising from the self fertilisation of primary transformants were grown on soil with watered each day with 5 µM $CuSO_4$ for two weeks. Each lane contains 1 µg of total RNA extracted from at least 6 individual T2 plants of a line. The membrane was hybridised with (**A**) a radiolabelled *te1* cDNA probe as described in Section 2.8c, and then (**B**) re-probed with a *ROC1* probe. Note that lane 11 (line 98.2) is overloaded. Expression levels are of the order of 10-100 pg *te1* transcript/µg of total RNA. Lanes are (with the transgene in parenthesis) **1**) 54.2 (empty vector) spiked with 100 pg of synthetic *te1* transcript, **2**) 54.2 spiked with 10 pg of synthetic *te1* transcript, **3**) blank, **4**) 54.2 (empty vector), **5**) 59.2 (empty vector), **6**) 30.7 (*te1*), **7**) 30.9 (*te1*), **8**) 30.20 (*te1*), **9**) 31.11 (*te1*), **10**) 98.1 (*NLS:te1*), **11**) 98.2 (*NLS:te1*), **12**) 98.4 (*NLS:te1*), **13**) 100.1 (*NLS:te1*), **14**) 100.2 (*NLS:te1*), **15**) 100.3 (*NLS:te1*), **16**) 100.4 (*NLS:te1*).

5.11. Expression of the *te1* transgene did not correlate with any observable phenotypes

5.11a. Neither vegetative nor floral morphology was affected by *te1* transgene expression

Once it was known that foliar application of 5 μM $CuSO_4$ was sufficient to induce physiologically high levels of the transgene (Figure 5.3), primary transformants grown in soil watered with 5 μM $CuSO_4$ were examined carefully. No consistent phenotypes were observed from 27 primary pDJ10/pDJ11 transformants (*te1* transgene) and nine pDJ13 (*NLS:te1* transgene) primary transformants. Leaf morphology, inflorescence architecture, and floral structure were indistinguishable from empty vector lines.

Maize plants with *te1* mutant alleles have increased leaf initiation rates, abnormal phyllotaxy, and narrower leaves (Veit 1998, N. Alvarez unpublished data). Overexpression of *te1* in *Arabidopsis* may also perturb this process, so the numbers of vegetative leaves and of inflorescence nodes in these lines were examined. Plants grown in soil watered with 5 μM $CuSO_4$ did not reveal any significant difference in the average numbers of visible vegetative leaves between empty vector lines and lines containing the *te1*transgene, average numbers of leaves for lines transformed with *te1* cDNA-containing copper-inducible vectors fell both above and below empty vector lines (Table 5.3). In all such experiments seeds were germinated and grown concurrently.

T2 generation plants were also grown in MS media containing kanamycin to select for those plants containing the T-DNA. Lines carrying both the *te1* transgene and the *NLS:te1* transgene were grown on MS media containing 0.1, 0.5, 5 and 50 μM $CuSO_4$ and compared to empty vector (pPMB765) lines. Live seedlings were periodically examined under a dissecting microscope from germination until inflorescence shoots were visible in the apex. This type of experiment allowed closer examination of young seedling phenotypes from precise copper concentrations, and ensured that all individuals contained the T-DNA. No consistent phenotype was observed, although 50 μM $CuSO_4$ was toxic to plants.

Table 5. 2 Leaf number frequencies in *te1* expressing *Arabidopsis* plants

Sorted in ascending order by the average number of leaves.

vector	line	plants with indicated numbers of leaves						*n*	average number of leaves
		4	5	6	7	8	9		
pDJ11	61.1	2	2	3	1	0	0	8	5.3
pPMB765	54.3	2	3	4	2	0	0	11	5.5
pDJ10	30.10	4	1	5	6	0	0	16	5.8
pDJ10	31.6	3	1	9	5	0	0	18	5.8
Columbia	-	2	2	8	10	0	0	22	6.1
pDJ11	61.3	1	0	4	4	1	0	10	6.4
pDJ10	3.06	1	0	8	8	1	0	17	6.4
pPMB765	54.1	0	3	3	3	2	1	13	6.5
pDJ10	31.5	0	0	6	9	2	0	17	6.7
pDJ10	31.12	0	0	2	5	10	0	17	7.4
average									***6.2***

5.11b. The wrinkled leaf phenotype did not correlate with transgene expression

Because addition of EDTA may help to reduce copper toxicity (H. Zhang, pers. comm.) later induction experiments of soil-grown plants applied a foliar application of 50 μM $CuSO_4$:EDTA, and sub-irrigated with 5 μM $CuSO_4$. In T-DNA segregating populations of T2 plants some individuals developed an abnormal 'wrinkled leaf' morphology (Figures 5.4a-c). Wrinkled leaf plants grew smaller malformed leaves, that were buckled in the dorsoventral axis and developed fewer trichomes. Both *NLS:te1* lines, and *te1* lines without the NLS developed wrinkled leaves. Because this phenotype could easily be the result of copper toxicity northern blotting was used to examine whether this phenotype correlated with transgene expression.

In the first experiment RNA was extracted from individual wrinkled or normal plants. Northern blotting indicated that in two of the three lines the wrinkled plant was expressing significantly more of the transgene (Figure 5.5). However, when transgene expression from small pools of wrinkled/normal plants was compared, this correlation was not consistent[5] (Figure 5.6). This phenotype was sporadic, it was not seen in every planting of lines that had been shown to express the transgene (see Table 5.2). The segregation of wrinkled and normal plants did not correspond to the kanamycin segregation (data not shown).

[5] In this analysis *NLS:te1* expression in both lines examined (98.1, 100.2), correlated with the phenotype, expression of the unaltered *te1* cDNA did not. Including both northern blots this included nine *NLS:te1* expressing plants.

Table 5. 3 Copper-inducible Arabidopsis transformant lines

plant line	kan-R	kan-S	ratio:1	vector	Expression observed by Northern?*	Wrinkled leaf phenotype?** [$CuSO_4$]	Notes
te1 lines							
30.5	70	12	5.8	pDJ10		✗ [5 μM]	
30.7	64	4	16	pDJ10	✓	✗ [5 μM], ✓ [50 μM]	
30.9	78	18	4.3	pDJ10	✓	✓ [5 μM]	
30.11	30	12	2.5	pDJ10	✓	✗ [5 μM]	
30.15	101	24	4.2	pDJ10		✗ [5 μM]	Aerial rosette in T1
30.20	113	21	5.4	pDJ10	✓	✓ [5 μM], ✓ [50 μM]	Aerial rosette in T1
31.2	62	13	4.8	pDJ10		✗ [5 μM]	Aerial rosette in T1
31.3	86	3	28.7	pDJ10		✗ [5 μM]	Aerial rosette in T1
31.11	9	11.4	103	pDJ10	✓	✓ [5 μM], ✓ [50 μM]	
31.13	67	13	5.2	pDJ10		✗ [5 μM]	Tandem insertion of T-DNA
33.15	72	16	4.5	pDJ10		✗ [5 μM]	
33.16	73	9	8.1	pDJ10		✗ [5 μM]	
60.1	107	15	7.1	pDJ11		✗ [5 μM]	
60.2	111	12	9.3	pDJ11		✗ [5 μM]	
NLS:te1 lines							
98.1	71	27	2.6	pDJ13	✓	✗ [5 μM], ✓ [5 μM], ✓ [50 μM]	
98.2	81	28	2.9	pDJ13	✓	✗ [5 μM]	
98.3	82	33	2.5	pDJ13		✗ [5 μM]	
98.4	89	32	2.8	pDJ13	✓	✗ [5 μM]	
100.1	4	28	112	pDJ13	✓	✓ [5 μM], ✗ [5 μM]	
100.2	nd			pDJ13	✓	✗ [50 μM]	
100.3	nd				✓		
100.4	nd			pDJ13	✓		
empty vector lines							
54.2	nd			pPMB765	✗	✗ [5 μM]	
59.2	nd			pPMB765	✗	✗ [5 μM]	

* A tick indicates that expression was observed, a cross that expression was not observed. All other lines were not examined. ** Only those lines whose T2 pants were examined carefully are given a ✓ or ✗ designation, numbers [5 μM]/ [50 μM] refer to the $CuSO_4$ watering regime in which the phenotype was observed. Each tick or cross refers to a specific experiment.

Figure 5.4a. Wrinkled leaf individuals were smaller than normal siblings. The one normal plant in this pot is at the top. All other plants had wrinkled leaves. Graduations on the ruler marked "30.20D" are 1 mm. This image shows T2 siblings of line 30.20 that were watered with 5 μM $CuSO_4$, and sprayed with 50 μM $CuSO_4$

Figure 5.4b. The wrinkled leaf phenotype. The leaves of normal plants (above) were larger, had a more ordered phyllotaxy, and developed more trichomes than wrinkled siblings (below). The leaves of wrinkled individuals appear to be folded in the dorsoventral axis. These images show T2 siblings of line 30.20 that were watered with 5 μM $CuSO_4$, and sprayed with 50 μM $CuSO_4$. This line was transformed with the *te1* transgene without the SV40 NLS.

Figure 5.4c. The wrinkled leaf phenotype from a line containing the *NLS:te1* transgene. Several transgenic lines showed the wrinkled leaf phenotype, including lines with and without an NLS. These images show T2 siblings of line 98.1 that were watered with 5 μM $CuSO_4$, and sprayed with 50 μM $CuSO_4$. The images are at the same scale.

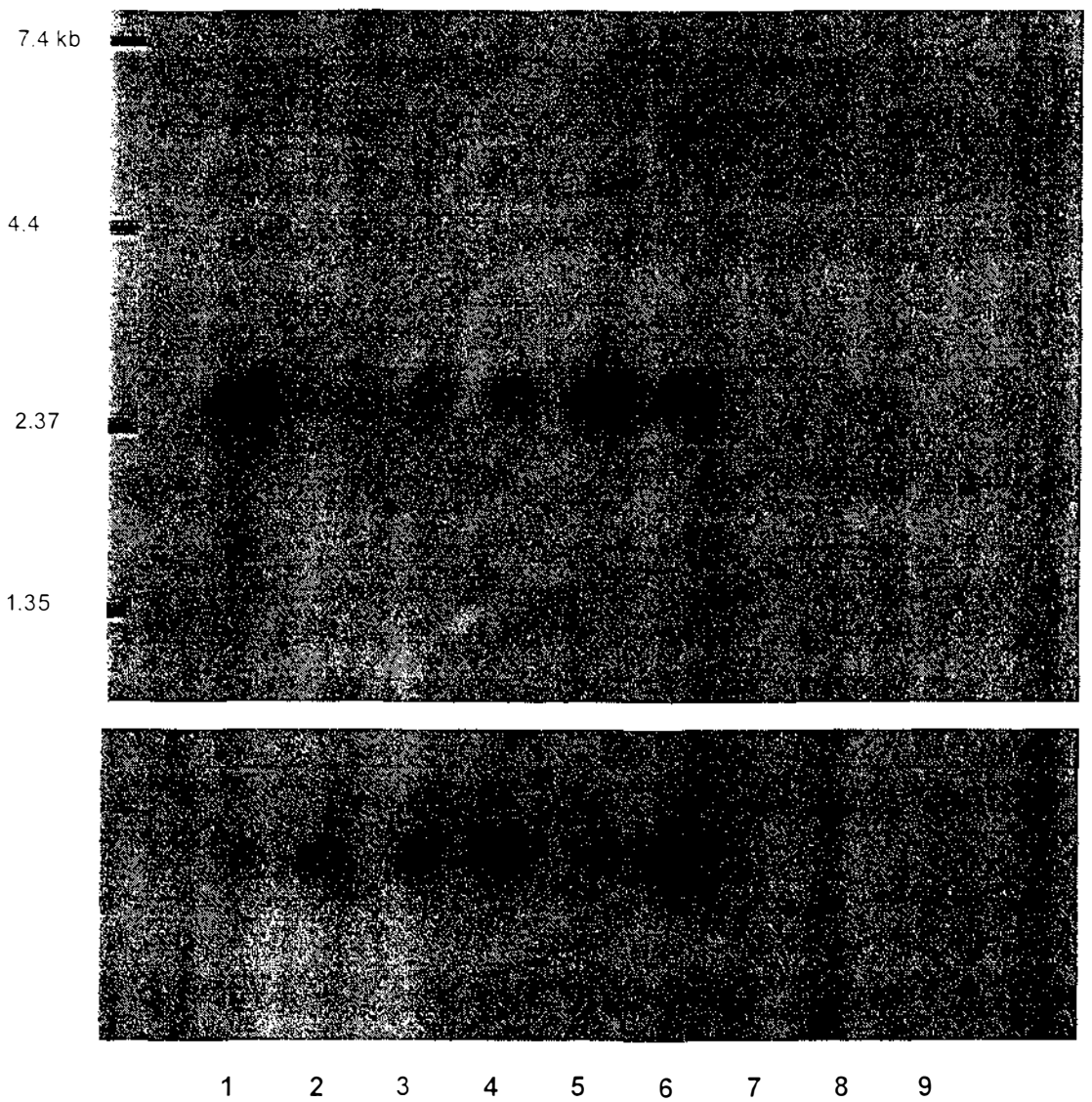


Figure 5.5. Northern blot showing *te1* transcript expression in individual wrinkled and normal sibling pairs. In this experiment, the *te1* transcript appears to be more highly expressed in wrinkled plants than in normal siblings. RNA was extracted from one wrinkled plant and one normal plant from three transgenic lines that had been grown in soil with 5 μM/50μM CuSO4, subjected to northern blotting, and probed with (a) radiolabelled *te1* probe, and then (b) re-probed with a *ROC1* probe. Molecular weight markers are shown at left, lanes are **1**) line 30.7 (pPMB765/*te1*) wrinkled, **2**) 30.7 normal, 3) 98.1 (pPMB765/NLS:*te1*) wrinkled, **4**) 98.1 normal, **5**) 100.2 (pPMB765/NLS:*te1*) wrinkled **6**) 100.2 normal, **7**) blank, **8**) 100 pg synthetic *te1* transcript, **9**) 10 pg synthetic *te1* transcript.

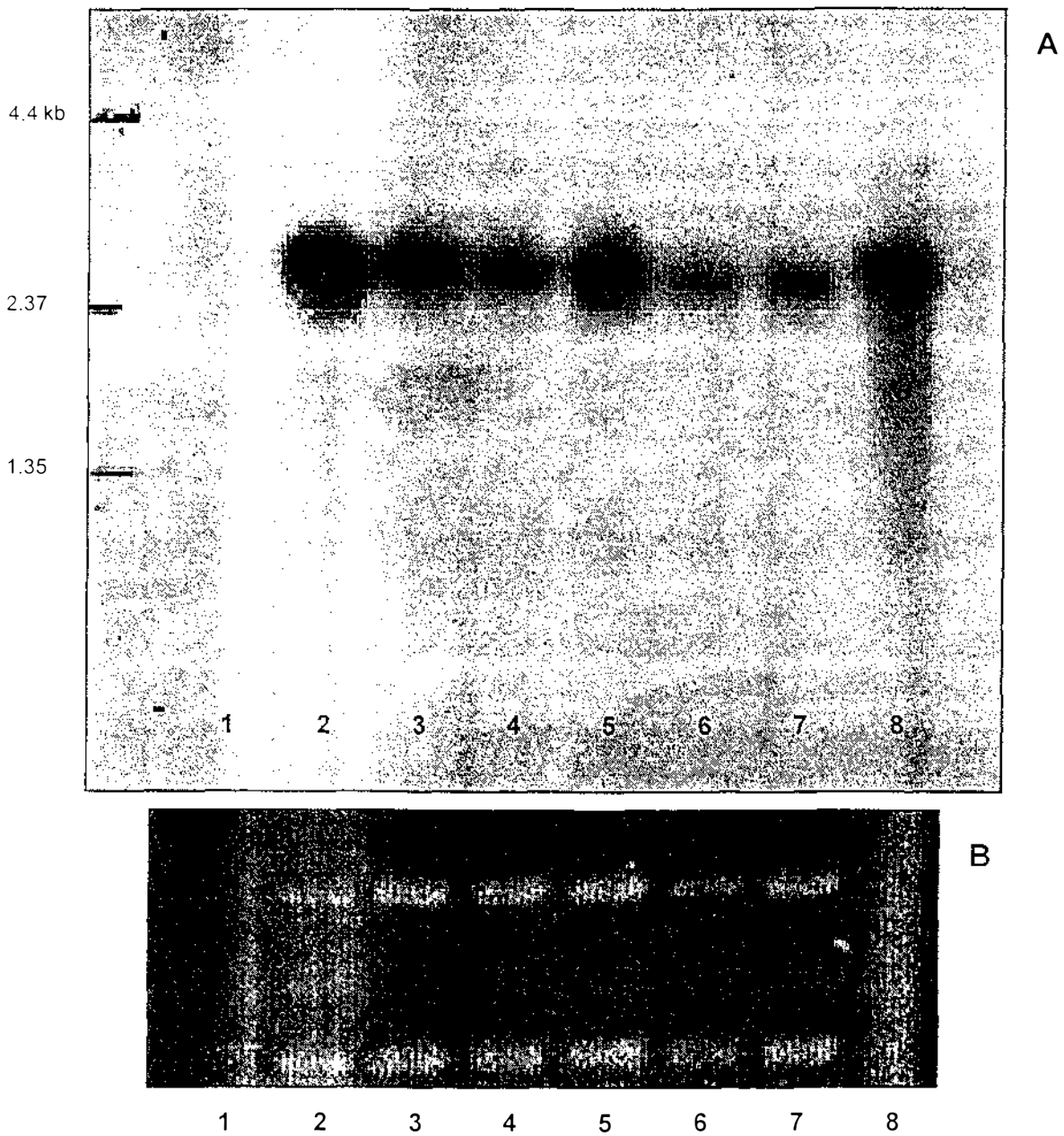


Figure 5.6. The wrinkled phenotype did not consistently correlate with expression of the *te1* transgene. Total RNA was extracted from small pools of 2 to 3 wrinkled plants and 2 to 3 normal plants of three genotypes. RNA (10 µg) was subjected to northern blotting, and probed with a radiolabelled *te1* cDNA probe as described in Section 2.8c. in only one line (98.1) of the three lines examined here did expression levels correlate with the wrinkled phenotype. Lanes are (with the transgene in parenthesis) **1**) one 54.2 plant (empty vector), **2**) two wrinkled 98.1 plants (*NLS:te1*), **3**) three normal 98.1 plants (*NLS:te1*), **4**) two wrinkled 31.11 plants (*te1*), **5**) three normal 31.11 plants (*te1*), **6**) three wrinkled 30.20 plants (*te1*), **7**) two normal 30.20 plants (*te1*), **8**) 100 pg synthetic *te1* transcript. **A**) shows northern blot, **B**) rRNA bands visualised by ethidium staining.

5.11c. The anti-MBP:TE1$^{1-286}$ antibody did not detect the TE1 protein in transgenic plants

Western blotting was used in an attempt to compare the levels of the TE1 protein in wrinkled and normal plants. Protein was extracted from the same tissue samples of wrinkled/normal plants that had been shown to express the *te1* transcript (Figure 5.5). Protein (10 μg) was subjected to SDS-PAGE, and Western blotting using the affinity purified anti-MBP:TE1$^{1-286}$ antibody (characterised in Section 4.10) at 1:6000 dilution as described in Section 4.7b. Even though high levels of *te1* transcript were detected from these tissues, no protein of the expected 71 kD was unique to lines expressing the *te1* transgene (data not shown). A simplistic calculation (below) indicates that even the highest expressing transgenic plants probably contain of the order of 1 ng of TE1 or NLS:TE1 protein per 10μg, which may be below the level of detection of this antibody.

Levels of TE1 protein expected in transgenic plants:

- The *te1* cDNA was present in 1:100,000 cDNA clones of a maize SAM library (B. Veit, unpublished data)
- This is equivalent to ~100 fg/μg of total RNA (Section 2.12)
- Highest expressing transgenic lines (Figure 5.5) express at ~10 pg/μg (1:1,000 cDNAs)
- Assume transcript levels ~= protein levels,
- Then transgene is 1:1,000 proteins, = 1 ng/10 μg
- Titration of MBP:TE1$^{1-286}$ fusion protein positive controls for western blots has shown that the affinity purified anti- MBP:TE1$^{1-286}$ antibody detects 10 ng of MBP:TE1$^{1-286}$ fusion protein easily, and barely detects 1 ng (data not shown).

5.11d. Aerial rosettes formed, but they were not due to expression of the *te1* transgene

A number of the first primary transformants that were recovered[1] developed extra leaves in the axils of inflorescence shoots (Figure 5.7). These leaves had a morphology and phyllotaxy that was more similar to vegetative leaves that to cauline leaves. Such 'aerial rosettes' were only seen in primary transformants, and were never observed in the T2 progeny. The phenotype was also observed in an empty vector line (line 59.1) (data not shown).The phenotype may have been induced by

[1] Lines that developed aerial rosettes were: 30.6, 30.10, 30.15, 30.20, 31.2, 31.3, 31.4, 31.5, 31.6, 30.10, 31.12 (all pDJ10 transformants), and the empty pPMB765 transformant line 59.1.

the growth conditions of these plants but was clearly not the due to ectopic *te1* expression.

Figure 5.7. The aerial rosette phenotype. Many primary transformants, including some transformed with 'empty' pPMB765 vector, developed rosette-like structures from the axils of cauline leaves. This phenotype was not seen in the progeny of these primary transformants, and is tentatively attributed to conditions of growth or transplantation treatments. **Above**) Normal inflorescence nodes develop up to three leaves with a characteristic morphology and phyllotaxy, **Below**) 'aerial' or axillary rosettes developed more leaves that were more similar to vegetative morphology and phyllotaxy, the rosette arises on an 'inflorescence shoot' from the axil of a cauline leaf cauline leaf.

5.12. Initial observations of TEL2 glucocorticoid fusion plants showed no clear phenotype[2]

Three kanamycin resistant lines were recovered from pBI-ΔGR:TEL2 transformations. The pBI-ΔGR:TEL2 construct produces a fusion of the TEL2 gene and a derivative of the glucocorticoid receptor protein (*TEL2:GR* transgene). Glucocorticoid fusions are able activate transcription factors, when localised to the nucleus in the presence of the inducer dexamethasone (Aoyama 1997, Wagner 1999).

In a pilot experiment approximately 100 Columbia ecotype plants, and 100 T2 plants from each line that had been sprayed with dexamethasone were compared to populations that had been mock sprayed (described in Section 5.6). These plants were sprayed until inflorescences had developed. No clear differences were seen between sprayed plants and mock sprayed plants.

These lines have not been analysed at the molecular level.

[2] This work was conducted in collaboration with Vernon Trainor and Carmel Gillman.

DISCUSSION (5)

5.12 The aerial rosette phenotype is the result of growth conditions

Some of the primary transformants that were initially recovered developed extra leaves in the axils of cauline leaves (Figure 5.7). These leaves were arranged with the spiral phyllotaxy of a vegetative rosette, and their morphology was more similar to vegetative leaves than to cauline leaves. These 'aerial rosettes' were formed from an axillary meristem that grew with a vegetative program rather than a inflorescence program. The phenotype was not due to expression of the *te1* transgene, since this phenotype only developed in T1 plants that were not watered with $CuSO_4$, was not seen in any T2 plants that were shown to be expressing the *te1* transgene, and was present in an empty vector line.

The development of aerial rosettes may have been the result of a long period of growth in short days. The first primary transformants that were recovered (and subsequently developed aerial rosettes) were grown for 38 to 40 days with an eight hour photoperiod. During this time light intensity was relatively low, 30-50 $\mu mol/m^2/$ s, produce from Sylvania Cool White fluorescent bulbs. These plants were then transferred to soil and grown with a 16 hour photoperiod under natural lighting supplemented with sodium lights. The light intensity in these conditions was 300-500 µM/m2/s (depending on the intensity of sunlight).

The primary transformants recovered in this way had begun to develop an inflorescence shoot by the time they were transferred to soil, they had developed many vegetative leaves, and probably initiated axillary meristems. It is likely that the axillary meristems initiated in these conditions were not competent to be induced into a floral developmental program.

Very similar aerial rosettes have been described in a late flowering ecotype of *Arabidopsis*, Sy-0 (Grbic 1996). The authors propose that the axillary meristems of Sy-0 maintain a prolonged vegetative phase, even though the primary shoot apical meristem has already converted to reproductive development. In the case of Sy-0 genetic evidence indicates that aerial rosette phenotype arises due to the interaction

between dominant alleles of two genes: *ART*, aerial rosette gene and *EAR*, enhancer of aerial rosette. *EAR* has been tentatively identified as a new allele of the *FRI* locus.

This study unwittingly showed that aerial rosettes can form in a Columbia background. An explanation for the effect is that once axillary meristems form, they are usually induced into an inflorescence program. In the conditions that were used in these experiments, it seems that a prolonged period of growth in low light intensity with a short photoperiod precluded floral evocation, and the axillary meristems developed with a vegetative program. Reversion from floral back to vegetative development has been observed in other species, and can even result floral structures with leaf-like appendages (Battey 1990).

5.13. Expression of the maize *te1* gene or the *Arabidopsis TEL2* gene in *Arabidopsis* did not produce any consistent phenotype

5.13a. Technical issues with *te1* transgene expression

Technical issues concerning transgene expression at the RNA level have been thoroughly investigated; both the *NLS:te1* and *te1* transcripts of the expected sizes were detected, at a variety of levels in different lines. The promoter-transgene junctions were confirmed by sequencing the transforming vector, and it was shown that the vector was not rearranged in the *Agrobacterium tumefaciens* GV3101 (Section 5.3). While T-DNA insertions are known to be truncated and rearranged, the *te1* cDNA portion of the T-DNA was confirmed to be present in ten of the initial transformants. Two of these lines appeared to contain head-to-tail multimers of the T-DNA (Section 5.8, Figure 5.2). The promoter-transgene junction in these lines does not appear to have been deleted, since the unique *Xba*I site in the T-DNA is between the *te1* gene and the MRE/35S promoter. Overall 27 lines with the *te1* transgene, and 8 *NLS:te1* lines were recovered, so it is unlikely that all of these lines expressed aberrant transcripts.

There is a possibility that expression from the MRE/35S promoter *Arabidopsis* does not include the shoot apical meristem (SAM). All *TEL* group genes that have been examined by *in situ* hybridisation to date are expressed in the SAM (Veit 1998 (maize *te1*), N. Alvarez (*Arabidopsis TEL1* and *TEL2*)). The SAM may be the only cellular context for expression that perturbs development. It has been shown that β-glucuronidase expression in *Nicotiana* from this promoter includes the shoot apical

meristem (H Zhang, pers. comm.), but similar experiments have not been conducted in *Arabidopsis*.

Although northern blots indicated that high levels of the *te1* transcript were expressed in the aerial parts of transgenic plants, it is unknown whether the TE1 protein was produced. The affinity purified anti-MBP:TE1$^{1-286}$ antibody did not detect the protein from plants that were known to express either *NLS:te1* or *te1* transcripts, but this result was limited by the sensitivity of the antibody (Section 5.9c). The *te1* transgene gene is expressed at 10 to 100-fold higher levels in *Arabidopsis* than in maize vegetative SAMs (see Section 2.12), so even low levels of the translation should be sufficient for function.

5.13b. Technical issues with *TEL2:GR* transgene expression

The initial TEL2:GR transformants that were recovered have not been analysed at the molecular level. Because the pBI-ΔGR vector was present in low copy numbers in *E. coli*, the construction of the pBI-ΔGR:TEL2B vector was not confirmed by restriction digestion or sequencing. The T-DNA integrations of the three transgenic lines have not been confirmed by Southern blotting. Another consideration is the possibility of co-suppression of the *TEL2:GR* transgene mediated by endogenous *TEL2* transcripts. Given current models of co-suppression (Vaucheret 1998, Meins 2000) it is possible that the chimeric transgene is not accumulating.

5.13c. Biological aspects of *TEL* genes

Assuming that the TE1 protein was produced in shoot meristems, there are at least four aspects of *TEL* gene biology that could account for the lack of a clear phenotype: protein nuclear localisation, differences between *Arabidopsis* and maize TEL protein structure, redundancy in *Arabidopsis* TEL proteins, and cofactor (protein/RNA) requirements.

The *S. pombe* MEI2 protein requires nuclear localisation for function, localisation is dependant on its association with the cognate RNA, meiRNA (Watanabe 1994). The protein retains full function in the absence of the meiRNA when the SV40 large T antigen NLS is attached at the C-terminus of the protein. GFP fusions have was shown that the SV40 NLS caused RNA-independent nuclear localisation of the MEI2 protein in both *S. pombe* and mammalian cells (Watanabe 1994, Yamashita 1998). Given the functioning and localisation of MEI2, the TE1 protein,and the related *Arabidopsis* TEL proteins may also function when localised to the nucleus. Because

monocots and dicots diverged ~ 200 Mya (Wolfe 1989), it is possible that the maize TE1 protein was unable to interact with the *Arabidopsis* RNA(s) that were required for nuclear localisation. The SV40 large T antigen NLS was used to address this concern because of the precedent for its use with MEI2 (Watanabe 1994, Yamashita 1998). This NLS also functions in plants: N- and C-terminal SV40 NLS fusions cause nuclear localisation of GUS in tobacco cells (Chua 1991).

Another possibility is that aspects of TE1 function other than nuclear localisation may be impaired due to divergence of maize and *Arabidopsis* protein/RNA structure. The MEI2 protein appears to require its third RRM irrespective of nuclear localisation (Watanabe 1994), suggesting that it interacts with other RNA(s). The TE1 protein may not interact with other *Arabidopsis* factors that are required for function, RNAs or other proteins.

A third (related) possibility is that the molecular mechanisms of the two TE1-like proteins (TEL1 and TEL2) in *Arabidopsis* are redundant to such an extent that overexpression of one type does not significantly perturb the process. There are several lines of evidence that are consistent with functional divergence. *TEL1 and TEL2* have different, but overlapping domains of mRNA expression in *Arabidopsis* SAMs (N Alvarez, unpublished). A T-DNA insertion in the first exon of *TEL1* does not show any obvious phenotype (Carmel Gillman, unpublished data), unlike the nine *te1* alleles that have been recovered.

Fourthly, it is possible that TEL proteins require some specific cofactors for function. Even if the TE1 protein is positioned within the correct tissues, and correctly localised to the nucleus, these cofactors that are required could restrain function to particular developmental conditions. Function of TEL proteins probably require a specific RNA (since the third RRM is so highly conserved, see Section 3.5) and may require phosphorylation like MEI2 (Watanabe 1997), or specific protein cofactors. Any of these factors could be in limiting quantities or expressed in a restricted domain, thereby limiting the activity of the protein. The lack of cofactors as an explanation for the lack of overexpression phenotype has been proposed for *PHABULOSA* (J Bowman, pers. comm.), and TEL proteins may require a specific RNA to be expressed before a constitutive phenotype is observed.

5.14d. A reinterpretation of overexpression phenotypes based on recent data

While this thesis was it its final stages some abnormal phenotypes were observed in *Arabidopsis* plants that transformed with a green fluorescent protein-tagged TEL2 construct under the control of the 35S promoter (V Trainor, unpublished data). Plants were small, produced abnormal 'tendril shaped' leaves that lacked green colour. Small growths of tissue grew from beneath trichomes that developed more than normal numbers branches. Since *TEL2* is expressed in the central zone of the *Arabidopsis* vegetative SAM, it may confer SAM central zone identity, by inducing *STM* expression, for example. One aspect of central zone identity is slow cell division (hence small leaves), and presumably lack of plastid development.

Because the GFP:TEL2 plants are so small, there is a strong possibility that the GR:TEL2 plants produced abnormal phenotypes that were not seen. The T2 populations planted were self fertilised T1 plants, and expected to segregate for the transgene. Small, malformed plants may have been overgrown by healthy plants not containing the T-DNA.

The 'wrinkled leaf' phenotype produced in the *te1* expressing are similar to the malformed leaves of GFP:TEL2 plants. The *te1* overexpression phenotype could be envisaged as a *partial* transformation of leaf cell identity into SAM central zone identity. The reasons for the inconsistency between *te1* expression and phenotype are not understood.

5.15. Concluding remarks

The expression of maize *te1* and *Arabidopsis TEL2* in *Arabidopsis* was designed to examine two questions:

1. If the lack of *te1* expression in maize allows de-repression of leaf primordia initiation sites, does the over expression of *TEL* genes cause repression leaf initiation sites?
2. What effect does nuclear localisation of the protein have on the function of TEL proteins?

The first question cannot be answered unequivocally, because the *te1* gene was expressed in a heterologous system, and the TEL2:GR plants have not yet been

carefully examined on a molecular level. However, we can be fairly sure that expression of the TE1 protein, either with or without a strong NLS, does not repress leaf initiation sites in *Arabidopsis*. It is not clear whether this is due to the protein being divergent from the *Arabidopsis* orthologs, or to other factors.

Without a *te1* overexpression-phenotype it is difficult to draw any conclusions about nuclear localisation. The TEL2:GR fusion system is much better suited to an investigation of this issue for the following reasons.

1. *TEL2* is a native *Arabidopsis* protein, thus functional divergence is not an issue
2. The fusion can be expressed in a TEL2 mutant line that has recently been recovered (C Gillman, unpublished data), which will ensure that the fusion is functioning in a normal manner
3. The glucocorticoid dexamethasone induction system allows direct comparison of the phenotypes produced from cytoplasmic or nuclear-localised TEL2 protein from the same transgenic line (or the same plant)

There are potential problems with dexamethasone treatment, however. Kang *et al.* (1999) noted severe growth defects in dexamethasone-treated *Arabidopsis* plants that had been transformed with a GR fusion. Growth defects included yellow cotyledons, failure to produce true leaves, and a severe retardation of overall growth. The failure to develop true leaves in particular would confuse issues of TEL gene function These growth defects were observed in transgenic empty vector lines (containing the GR gene) when cultured on media containing little as 0.1 μM dexamethasone, but were not seen on un-transformed plants that were treated with as much as 30 μM dexamethasone.

Despite this issue it may be possible to analyse the effect of transitory TEL protein nuclear localisation with the GR system, and a variety of other approaches are being investigated by other members of the Veit laboratory (discussed in Chapter 6)

CHAPTER 6. SYNOPSIS

6.1. Recapitulation of the aims of the project

The main aim of this project was to gain further understanding of the function of the *te1* gene. While sequencing the genomic clones did not produce much more information about the genetics or biology of the gene than the intron structure, it did facilitate the further molecular analysis of the gene by 5' RACE, RT-PCR, and the study of the mutant alleles (Chapter 2). The 5' RACE results are equivocal; the predicted start site was not confirmed by RT-PCR, and the predicted transcript appears to contain an in-frame stop codon (Figure 2.3). Initial attempts during this project to analyse the expression of the *te1* gene used northern blotting, but they were unsuccessful (data not shown). RT-PCR proved to be a more tractable method, and because each primer was designed to be complimetry to an exon boundry (ie, one end of the primer on one exon, and extending across another into another) the results are unlikely to be artefactual. Since and the PCR product extends across the region encoding the C-terminal RRM, which is the most highly conserved in Mei2-like proteins (Chapter 3) the transcripts are probably full length and produce functional proteins.

One of the initial aims of the project was to study the function of the TE1 protein *in vitro* (Chapter 4), and at the cellular level in maize. Two main aspects of the TE1 proteins function were considered: the cellular location of the protein would provide some clues as to function, and knowledge of the RNA target would allow us to follow the next step in the biochemical process. The problems that were encountered with expressing the protein, and the non-specific activity of the polyclonal sera meant that these aspects of the protein could not be studied. The work did provide an understanding of the problems with heterologous TE1 expression, and this allowed further work with the *Arabidopsis* TEL2 protein proceed routinely (V Trainor, unpublished data). Raising of polyclonal sera to other plant MEI2-like proteins could also draw from this work, and the analysis of protein structure presented in Chapter 3. Variable solvent exposed loops may be the best regions of proteins to affinity purify antibodies against.

The lack of obvious phenotype associated with *te1, NLS:te1,* or *GR:TEL2* expression in *Arabidopsis* is difficult to interpret. This author favours three explanations. Firstly, the TE1 protein is simply too structurally different from the *Arabidopsis* orthologs to function. Secondly there may well be technical problems with the GR:TEL2 construct, expression or induction. A third possibility is that the GR:TEL2 was not able to function because it could not be transported intracellularly. The GR fusion is always limited to a particular location, either in the cytoplasm in the absence of dexamethasone, or in the nucleus with dexamethasone. If translocation was essential for function as it is in MEI2 (Yamashita 1998), the method used would inhibit function.

The ability to predict and analyse the *Arabidopsis* Mei2-like genes *in silico* was an unexpected area of research (Chapter 3). While *in silico* biology does not provide proof of function it did enable experiments to be designed more quickly, more precisely, and with more confidence. Once the *Arabidopsis* orthologs of the maize *te1* gene were predicted from the genomic sequence, the cDNAs for *TEL1* and *TEL2* were subsequently cloned routinely by RT-PCR (V Trainor, unpublished data). Just as importantly, the genome project allowed the complete family of *Arabidopsis* Mei2-like genes to be identified. We can be certain that *TEL1* and/or *TEL2* are the orthologs of the maize *te1* gene, and direct experiments accordingly. The analysis of the protein homology and structure will provide a useful guide to future RNA binding and functional studies. Residues can be selected to modify that may generate partial loss-of-function alleles, for example. The analysis also showed that Mei2-like proteins contain novel structures that differentiate them from canonical RRM proteins. These regions probably have some important functional and structural properties, since they are conserved in all MEI2-like proteins.

6.2. Homology, orthologs and genome duplications

An important finding of this project was that there are two *Arabidopsis* genes (*TEL1* and *TEL2*) that are equally good candidates for the ortholog (similar gene that has the same function) of the maize *te1* gene. It is not known if maize contains another *te1*-like gene, both *Zea mays* genes that were identified from the TIGR gene index (EST assemblies) were of the AML group (Figure 3.9). However the maize *brachytic*

mutants look very like *terminal ear1* mutants, and each of the 3 *brachytic* genes are located on different chromosomes to *te1* (see footnote[1]) (Neuffer 1997).

The monocot-dicot divergence is estimated to be have occurred before 170 million years ago (Mya) (Yang 1999), and early monocot lineages were certainly present 100 Mya (Bremer 2000). Since this divergence, separate duplications of large portions of the *Arabidopsis* genome (112 Mya; Ku 2000), and the maize genome (20 Mya and 11 Mya; Gaut 1997) have occurred. Functional divergence probably followed this duplication.

6.3. Models of *TEL* gene function

Our model for *TEL* gene function is being expanded as we discover more about the *Arabidopsis* orthologs. *TEL1* is expressed throughout vegetative, inflorescence and floral meristems, while *TEL2* expression is restricted to the central zone of vegetative apical meristems and floral meristems, and shows a distinctive spiral pattern in inflorescence meristems (Nena Alvarez, unpublished data). This data is consistent with an organ repression function for *TEL2*, but another role for *TEL1*. The differences between expression patterns of *TEL2* and *te1* can be understood in terms of the differences in SAM morphology and the degree to which internodes are extended in *Arabidopsis* and maize. *Arabidopsis* vegetative leaves have a spiral phyllotaxy and compressed internodes, and we do not observe any phyllotaxy-related pattern to the *TEL2* expression pattern (N Alvarez, unpublished data). It is only in the inflorescence that we observe a spiral pattern of *TEL2* expression, and the internodes become elongated at this stage in *Arabidopsis* development. Combined with the analysis of *te1* expression levels in maize tissues (Figure 2.15), these results are consistent with the timing and extent of *TEL* gene expression being a major factor in the evolution of plant architecture. A similar proposal has been made for *teosinte branched1* expression levels in the alteration of maize and teosinte morphology, although the changes are less extensive (Doebly 1997). A genomic region of the *te1* gene was compared with *Zea* genus to test the hypothesis that *te1* had been a major determinant of the morphological evolution of the genus (White 1999). The authors found no evidence of past selection, however, this analysis did not analyse expression levels or the promoter regions of the gene.

[1] *brachytic1* (*br1*) is on 1L-81, *br2* is on 1L near *hm1* and *br3* is on chromosome 5. The *te1* gene is on 3L (Neuffer 1997)

We know very little about the cellular activity of plant Mei2-like genes. The *te1* expression pattern in maize SAMs, and the *te1* mutant phenotype have been explained by a model wherein the TE1 protein limits the positions of leaf initiation (Veit 1998). Given that RNas are trafficked through plasmodesmata (lucas 1995, Rinne 1998) one possibility is that some leaf initiation promoting transcript acts to initiate leaves at a specific longitudinal distance from the apex except in positions where the TE1 protein is present. In these cells the TE1 protein binds to this transcript and negatively regulates its activity. It would certainly be a significant breakthrough to identify the RNA target(s) of a plant Mei2-like protein. RRM proteins are involved in many and diverse cellular activities (Section 1.7b), so the cellular mechanism of their activity is impossible to predict. While the MEI2 protein is understood well at the cellular and molecular level (Reviewed in Ohno 1999), and plant Mei2-like gene probably share some biochemical (RNA binding) activity, the cellular and biological contexts may be very different.

6.4. A final summary and some new questions

The ultimate aim of a developmental study is the understanding of the mechanisms that give rise to form. To what extent Do we understand the role of *te1* and *TEL* genes? We have seen that the expression of the *te1* gene in maize relates to the spacing of leaf primordia in the SAM, and that the mutant phenotype appears to have longitudinally misregulated primordia initiation (Veit 1998). Expression levels in maize inflorescence meristems generally correlate to internode spacing, so the function appears to regulate the positioning of shoot lateral organs (this study). Very little is known about the biochemistry of TEL genes except that the proteins probably regulate some aspect of RNA metabolism or localisation. Since the structures are non-canonical, the biochemical interactions of the proteins may include a few surprises, there is certainly no closely related example to compare to. The *Mei2*-like gene family contains seven members in *Arabidopsis* that have different domains of expression (N Alvarez, unpublished data). These expression patterns include expression in the embryo and root quiescent centre implying a broader role than expected. Many questions remain unanswered, or have been brought to light by these studies:

- What is the biochemical activity of TEL proteins? Knowledge of the cellular localisation of TEL proteins would provide help to answer this question. The identity of the target RNA(s) would be another important clue.
- Do Mei2-like proteins have a structure not seen before? What does it do?
- How is the expression of TEL genes regulated? An interesting facet of *TEL* gene expression in both *Arabidopsis* and maize is that it is precisely patterned. What determines the pattern?
- What is the role of *TEL* genes in roots? Both *te1* and *TEL1* are expressed in roots at various stages of development, and it is unclear how the spatial positioning model applies to roots

REFERENCES

Aida M, Ishida T, Tasaka M (1999). Shoot apical meristem and cotyledon formation during *Arabidopsis* embryogenesis: interaction among the *CUP-SHAPED COTYLEDON* and *SHOOT MERISTEMLESS* genes. Development 126:1563-1570.

Alba MM, Pages M (1998). Plant proteins containing the RNA recognition motif. Trends in Plant Science 3:15-21.

Allain FHT, Gilbert DE, Bouvet P, Feigon J (2000). Solution structure of the two N-terminal RNA-binding domains of nucleolin and NMR study of the interaction with its RNA target. J. of Molecular Biology 303:227-241.

Aoyama T, Chua N (1997). A glucicorticoid-mediated transcriptional induction system in transgenic plants. The Plant Journal 11:605-612.

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smiith JA, Struhl K (1998). Current Protocols in Molecular Biology. John Wiley & Sons, New York.

Bandziulis RJ, Swanson MS, Dreyfuss G (1989). RNA-binding proteins as developmental regulators. Genes and Development 3:431-437.

Battey NH, Lyndon RF (1990). Reversion of flowering. Botanical Reviews 56:162-189.

Bauer P, Rolfs A, Regitz-Zagrosek V, Hildebrandt A, Fleck E (1997). Use of Manganese in RT-PCR eliminates PCR artefacts resulting from DNase I digestion. Biotechniques 22:1128-1132.

Birney E, Kumar S, Krainer A (1993). Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. Nucleic Acids Research 25:5803-5816.

Blom N, Gammeltoft S, Brunak S (1999). Sequence- and structure-based prediction of eukaryotic protein phosphorylation Sites. Journal of Molecular Biology 294: 1351-1362.

Bode PM, Bode HR (1984). Patterning in hydra. In: Malacinski GM, Bryant SV. Pattern Formation. Macmillan Publishing Company, New York.

Bourdon, V. and Tinland, B. Unpublished, Direct Submission to Genbank. Isolation and characterization of *MEI2*, the *Arabidopsis* homolog of the Schizosaccharomyces pombe mei2 gene. Submitted (22-OCT-1998). Department of Developmental Biology, Institute of Plant Sciences, ETH Zurich, Universitatstrasse, 2, Zurich 8092, Switzerland

Bowman JL (2000). Axial patterning in leaves and other lateral organs. Current Opinion in Genetics and Development 10:399-404.

Bowman JL, Smyth DR, Meyerowitz EM (1991). Genetic interactions among floral homeotic genes of *Arabidopsis*. Development 114:599-615.

Bremer K (2000). Early cretaceous lineages of monocot and dicot flowering plants. Proceedings of the National Academy of Sciences USA 97:4707-4711.

Brownlee C, Berger F, Bouget F-Y (1997). Signals involved in control of polarity, cell fate and developmental pattern in plants. In: Greenland AJ, Meyerowitz EM, Steer M (Eds). Control of Plant Development: Genes and Signals. Symposia for the Society for Experimental Botany 51, The Company of Biologists Ltd., Cambridge UK.

Brunak S, Engelbrecht J, Knudsen S (1991). Prediction of Human mRNA Donor and Acceptor Sites from the DNA Sequence, Journal of Molecular Biology 220:49-65.

Burd CG, Dreyfuss G (1994). Conserved structures and diversity of functions of RNA-binding proteins. Science 265:615-621.

Burge C, Karlin S. (1997). Prediction of complete gene structures in human genomic DNA. Journal of Molecular Biology 268: 78-94.

Burge CB (1998). Modeling dependencies in pre-mRNA splicing signals. In Salzberg S, Searls D, and Kasif S, (eds). Computational Methods in Molecular Biology, Elsevier Science, Amsterdam, p127-163.

Buvat R (1952). Structure, evolution et functionnement du meristem apical de quelques dicotyledones. Ann. Sci. Nat Bot. Ser. 11. 13:199-300.

Carpenter R, Cohen ES (1995). Transposon induced chomeras show that *floricaula*, a meristem identity gene, acts non-autonomously between cell layers. Development 121:19-26.

Chasan R (1992). Developing somatic embryos. The Plant Cell 4:367-368.

Chomczynski P, Sacchi N (1987). Single -step method of RNA isolation by acid guadinium thiocyanate-phenol-chloroform extraction. Annals of Biochemistry. 162:156-159.

Christensen AR, Sharrock R, Quail P (1992). Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. Plant Molecular Biology 18:675-689.

Church GM, Gilbert W (1984). Genomic sequencing. Proceedings of the National Academy of Sciences USA 81:1991-1995.

Clark SE, Rinning MP, Meyerowitz EM (1995). *CLAVATA3* is a specific regulator of shoot and floral meristem development affecting the asme processes as *CLAVATA1.* Development 121: 2057-2067.

Conte MR, Grune T, Ghuman J, Kelly G, Ladas A, Matthews S, Curry S (2000). Structure of tandem RNA recognition motifs from polypyrimidine tract binding protein reveals novel features of the RRM fold. The European Molecular Biology Organisation Journal 19:3132-3141.

Cosgrove DJ (2000). Loosening of plant cell walls by expansins. Nature 407:321-326.

Crowder SM, Kanaar R, Rio DC, Albert T (1999). Absence of interdomain contacts in the crystal structure of the RNA recognition motifs of Sex-lethal. Proceedings of the National Academy of Sciences USA 96:4892-4897.

Deo RC, Bonanno JB, Sonenberg N, Burley SK (1999). Recognition of polyadenylate RNA by the poly(A)-binding protein. Cell 98:835-845.

Dermen H (1953). Periclinal cytochimeras and origin of tissues in stem and leaf of peach. American Journal of Botany 40:154-168.

Ding JZ, Hayashi MK, Zhang Y, Manche L, Krainer AR, Xu RM (1999). Crystal structure of the two-RRM domain of hnRNP A1 (UP1) complexed with single-stranded telomeric DNA. Genes and Development 13:1102-1115.

Doebley J, Stec A (1991). Genetic analysis of the morphological differences between maize and teosinte. Genetics 129:285-295.

Doebley J, Stec A, Hubbard L (1997). The evolution of apical dominace in maize. Nature 386:485-488.

Dudley M, Poethig RS (1993). The heterochronic *teopod1* and *teopod2* mutations of miaze are expressed non-cell-autonomously. Genetics 133:389-399.

Ehlers K, Binding H, Kollmann R (1999). The formation of symplasmic domains by plugging of plasmodesmata: a general event in plant morphogenesis? Protoplasma 209:181-192.

Elliot RC, Betzner AS, Huttner E, Oakes MP, Tucker WQJ, Gerentes D, Perez P, Smyth DR (1996). *AINTEGUMENTA*, an *APETALA2*-like gene of *Arabidopsis* with pleiotropic roles in ovule development and floral organ growth. The Plant Cell 8:155-168.

Engval E (1980). Enzyme Immunoasay ELISA and EMIT. Methods in Enzymology 70:419-439.

Fleming AJ, McQueen-Mason S, Mandel T, Kulemeier C (1997). Induction of leaf primordia by the cell wall protein expansin. Science 276:1415-18

Fletcher JC, Meyerowitz EM (2000). Cell signaling in the shoot meristem. Current Opinion in Plant Biology 3:23-30

Foard DE (1971). The initial protusion of a leaf primordium can form without concurrent cell divisions. Canadian Journal of Botany 49:694-702.

Foster T, Veit B (2000). Genetic anlysis of leaf development and differentiation. In: Marshall B, Roberts JA (Eds.). Leaf Development and Canopy Growth. Sheffield Academic Press, UK.

Foster T, Veit B, Hake S (1999). Mosaic analysis of the dominant mutant, *Gnarley1-R*, reveals distinct lateral and transverse signalling pathways during maize leaf development. Development 126: 305-313.

Freeling M (1992). A conceptual framework for maize leaf development. Dev. Biol. 153:44-58

Furner IJ, Pumfrey JE (1992). Cell fate in the shoot apical meristem of *Arabidopsis thaliana*. Development 115:755-764.

Galinat (1959). The phytomer in relation to floral homologies in the American *Maydeae.* Bot. Mus. Leafl., Harvard University 19:1-32.

Gaut BS, Doebley JF (1997). DNA sequence evidence for the segmental allotetraploid origin of maize. Proceedings of the National Academy of Sciences USA 94: 6809-6814.

Gebauer F, Merendino L, Hentze MW, Valcarcel J (1997). Novel functins for nuclear factos in the cytoplasm - the sex-lethal paradigm. Seminars in Cell and Developmental Biology 8:561-566.

Genefinder http://dot.imgen.bcm.tmc.edu:9331/gene-finder/gf.html

Genscan http://genes.mit.edu/GENSCAN.html

Gisel A, Barella S, Hempel FD, Zambryski PC (1999). Temporal and spatial regulation of synplastic trafficking during development in *Arabidopsis thaliana* apices. Development 126:1879-1889.

Goethe JW (1790). Versuch die Metamorphose der Pflanzen zu erklären. CW Ettinger. Gotha, Germany.

Grbic V, Bleecker AB (1996). Altered body plan is conferred on *Arabidopsis* plants carrying dominant alleles of two genes. Development 122:2395-2403.

Green B, Walko R, Hake S (1994). Mutator insertions in an intron of the maize knotted1 gene result in dominant suppressible mutations. Genetics 138:1275-1285.

Green PB, Steele CS, Rennich SC (1996). Phyllotactic patterns: a biophysical mechanism for their origin. Annals of Botany. 77:515-527.

Hake S, Char BR (1997). Cell-cell interactions during plant development. Genes and Development 11:1087-1097.

Handa N, Nureki O, Kurimoto K, Kim I, Sakamoto H, Shimura Y, Muto Y, Yokoyama S (1999). Structural basis for recognition of the *tra* mRNA precursor by the Sex-lethal protein. Nature 398:579-585.

Hebsgaard SM, Korning PG, Tolstrup N, Engelbrecht J, Rouze P, Brunak S (1996). Splice site prediction in *Arabidopsis thaliana* DNA by combining local and global sequence information. Nucleic Acids Research 24:3439-3452.

Hemerly A, De Almeida Engler J, Bergounieux C, Van Montagu M, Engler G, Inze D, Ferreira P (1995). Dominant negative mutants of Cdc2 kinase uncouple cell division from iterative plant development. The European Molecular Biology Organisation Journal 14:3925-3936.

Inoue M. Muto Y. Sakamoto H. Kigawa T. Takio K. Shimura Y. Yokoyama S (1997). A characteristic arrangement of aromatic amino acid residues in the solution structure of the amino-terminal RNA-binding domain of *Drosophila* sex-lethal. Journal of Molecular Biology 272:82-94.

Itoh JI, Kitano H, Matsuoka M, Nagato Y (2000). *SHOOT ORGANISATION* genes regulate shoot apical meristem organisation and the pattern of leaf primordium initiation in rice. The Plant Cell 12:2161-2174.

Jackson D, Veit B, Hake S (1994). Expression of maize *KNOTTED1* related homeobox genes predicts patterns of morphogenesis in the vegetative shoot. Development 120:405-413.

Joshi CP, Zhou H, Huang X, Chiang VL (1997). Context sequences of translation initiation codon in plants. Plant Molecular Biology 35(6):993-1001

Kalderon D, Richardson WD, Markham AT, Smith AE (1984). Sequence requirements for nuclear localisation of simian virus 40 large T antigen. Nature 311:33-38.

Kishino, H, Hasegawa M (1989). Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. Journal of Molecular Evolution 29: 170-179.

Koes R, Souer E, van Houwelingen A, Mur L, Spelt C, Quattrocchio F, Wing J, Oppedijk B, Ahmed S, Maes T (1995). Targeted gene inactivation in petunia by PCR-based selection of transposon insertion mutants. Proceedings of the National Academy of Sciences USA 92:8149-53.

Koutz PJ, Davis GR, Stillman G, Barringer K, Cregg JM, Thill G (1989). Structural comparison of the *Pichia pastoris* alcohol oxidase genes. Yeast 5: 167-177

Kranz JK. Hall KB (1999). RNA recognition by the human U1A protein is mediated by a network of local cooperative interactions that create the optimal binding surface. Journal of Molecular Biology 285:215-231.

Ku HM, Vision T, Liu J, Tanksley SD (2000). Comparing sequenced segments of the tomato and *Arabidopsis* genomes: Large scale duplication followed by selective gene loss creates a network of synteny. Proceedings of the National Academy of Sciences USA 97:9121-9126.

Kusaba S, Kano-Murakami Y, Matsuoka M, Tamaoki M, Sakamoto T (1998). Alteration of hormone levels in transgenic tobacco plants overexpressing a rice homeobox gene *OSH1*. Plant Physiology116:471-76.

Laufs, Dockx J, Kronenberger J, Traas J (1998). *MGOUN1* and *MGOUN2*: two genes required for primordium initiation at the shoot apical and floral meristems in *Arabidopsis thaliana.* Development 125: 1253-1260.

Lenhard M, Laux T (1999). Shoot meristem formation and maintenance. Current Opinion in Plant Biology 2:44-50

Lippuner V, Chou IT, Scott SV, Ettinger WF, Theg SM, and Gasser CS (1994). Cytosolic and chloroplast forms of cyclophilin from *Arabidopsis thaliana*. Journal of Biological Chemistry 269:7863-7868

Long JA, Moan EI, Medford JI, Barton MK (1996). A member of the KNOTTED class of homeodomain proteins is encoded by the STM gene of *Arabidopsis*. Nature 379: 66-69.

Lorkovic ZL, Wieczorek Kirk DA, Lambermon MHL, Filipowicz W (2000). Pre-mRNA splicing in higher plants. Trends in Plant Science 5: 160-167.

Lowry OH, Rosebrough NJ, Farr AL, Randall RJ (1951). Protein measurement with the Folin-Phenol reagents. Journal of Biological Chemistry 193:265-275.

Lucas WJ, Bouchepillon S, Jackson DP, Nguyen L, Baker L, Ding B, Hake S (1995). Selective trafficking of KNOTTED1 homeodomain protein and its mRNA through plasmodesmata. Science 270:1980-1983.

Lucas WJ, Bouche-Pillon S, Jackson DP, Nguyen L, Baker L, Ding B, Hake S (1995). Selective trafficking of KNOTTED1 and its mRNA through plant plasmodesmata. Science 270:1980-1983.

Lucas WJ, Ding B, van der Schoot C (1993). Plasmodesmata and the supracellular nature of plants. New Phytologist 125:435-476.

Lyndon RF (1983). The mechanism of leaf initiation. In: Dale JE, Milthorpe FL. The Growth and Functioning of Leaves. Cambridge University Press, Cambridge UK.

Macknight R, Bancroft I, Page T, Lister C, Schmidt R, Love K, Westphal L, Murphy G, Sherson S, Cobbett C, Dean C (1997). *FCA,* a gene controlling flowering time in *Arabidopsis*, encodes protein containing RNA-binding domains. Cell 89:737-745.

Marc J, Hackett WP (1991). Gibberelin-induced reorganization of spatial relationships of emerging leaf primordia at the shoot apical meristem in *Hedera helix* L. Planta 185:171-178.

Marínez-García JF, Monte E, Quail PH (1999). A simple, rapid and quantitative method for preparing *Arabidopsis* protein extracts for immunoblot analysis. The Plant Journal 20:251-257.

Marx GA (1987). A suite of mutants that modify pattern formation in pea leaves. Plant Molecular Biology Reporter 5:311-335.

Mathews DL, Grogan CO, Manchester CE (1974). Terminal ear mutant of maize. J. agric. Sci. Camb. 82:433-435.

McCarty DR, Chory J (2000). Conservation and Innovation in Plant Signaling Pathways. Cell 103:210-209.

McDaniel CN (1984). Competence, determination and induction in plant development. In: Malacinski GM, Bryant SV. Pattern Formation. Macmillan Publishing Company, New York.

McDaniel CN Poethig RS (1988). Cell-lineage patterns in the shoot apical meristem of the germinating maize embryo. Planta 175:13-22.

McKeown M (1992). Sex differentiation: the role of alternative splicing. Current Opinion in Genetics and Development 2:229-303.

Meicemheimer RD (1981). Changes in *Epilobium* phyllotaxy induced by *N*-1-napthylpthamic acid and α-4-chlorophenoxy-isobutyric acid. American Journal of Botany 68:1139-1154.

Mett VL, Lochhead LP, Reynolds PHS (1993). Copper-controllable gene expression system for whole plants. Proceedings of the National Academy of Sciences USA 90:4567-4571

Meyerowitz EM (1996). Plant development: local control, global patterning. Current Opinion in Genetics and Development 5:475-479

Milan M, Campuzano S, Garcia-Bellido (1996). Cell cycling and patterned cell proliferation in the wing primordium of Drosophila. Proceedings of the National Academy of Sciences USA 93:640-645.

Mittwoch U (2000). Genetics of sex determination. Exceptions that prove the rule. Molecular Genetics & Metabolism 71:405-410.

Mizukami Y, Fischer RL (2000). Plant organ size control: *AINTEGUMENTA* regulates growth and cell numbers during development. Proceedings of the National Academy of Sciences USA 97:942-947.

Münster T, Pahnke J, Di Rosa A, Kin JT, Martin W, Saedler H (1997). Floral organ homeotic genes were recruited from homologous MADS-box genes pre-existing in the common ancestor of ferns and seed plants. Proceedings of the National Academy of Sciences USA 94:2415-2420.

Nagata T. Kanno R. Kurihara Y. Uesugi S. Imai T. Sakakibara S. Okano H. Katahira M (1999). Structure, backbone dynamics and interactions with RNA of the C-terminal RNA-binding domain of a mouse neural RNA-binding protein, Musashi1. Journal of Molecular Biology 287:315-330.

Netgene2 http://www.cbs.dtu.dk/services/NetGene2/

NetPhos http://www.cbs.dtu.dk/services/NetPhos

Neuffer MG, Coe EH, Wessler SR (1997). Mutants of maize. Cold Spring Harbour Laboratory Press, New York.

Neurospora database. http://www-genome.wi.mit.edu/annotation/fungi/neurospora/

Ohno M, Mattaj IW (1999). Meiosis: MeiRNA hits the spot. Current Biology 9:R66-R69.

Oubridge C, Ito N, Evans P, Hiang Teo C, Nagal K (1994). Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. Nature 372:432-438.

Pfam database RRM entry http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF00076

PLACE database Version 6.01 http://www.dna.affrc.go.jp/htdocs/PLACE/

Predict Protein Server http://www.embl-heidelberg.de/predictprotein/predictprotein.html

Reinhardt D, Wittwer F, Mandel T, Kuhlemeier C (1998). Localised upregulation of a new expansin gene predicts the site of leaf formation in the tomato meristem. The Plant Cell 10:1427-1437.

Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann, Seraphin B (1999). A generic protein purification method for protein complex characterization and proteome exploration. Nature Biotechnology 17:1030-1032.

Rost B (1996). PHd: predicting one-dimensional protein structure by profile based neural networks. Methods in Enzymology. 266: 525-539

Ruiz-Medrano R, Xoconostle-Cázares B, Lucas WJ (1999). Phloem long-distance transport of *CmNACP* mRNA: implications for supracellular regulation in plants. Development 126: 4405-4419.

Sachs T (1984). Controls of cell patterns in plants. In: Malacinski GM, Bryant SV. Pattern Formation. Macmillan Publishing Company, New York.

Sambrook J, Fritsch EF, Maniatis T (1989). Molecular Cloning: a laboratory manual. Cold Spring Harbour Laboratory Press, New York.

Saxena P, Walker JR (1992). Expression of argU, the E. coli gene coding for a rare Arginine tRNA. Journal of Bacteriology 174:1956-1964

Schaefer BC (1995). Revolutions in Rapid Amplification of cDNA ends: new strategies for polymerase chain reaction cloning of full length cDNAs. Analytical Biochemistry 227:255-273.

Schwabe WW (1971). Chemical modification of phyllotaxis and its implications. Symposia of the Society for Experimental Biology 25:301-322.

Seigfreid KR, Eshed Y, Baum SF, Otsuga D, Drews GN, Bowman JL (1999). Members of the YABBY gene family specify abaxial cell fate in *Arabidopsis*. Development 126:4117-4128.

Shamoo Y, Abdul-Manan N, Williams K (1995). Multiple RNA binding domains (RBDs) just don't add up. Nucleic Acids Research 23:725-728.

Shinozaki-Yabana S, Watanabe Y, Yamamoto M (2000). Novel WD-repeat protein Mip1 facilitates function of the meiotic regulator *Mei2*p in fission yeast. Molecular Cell Biology 20:1234-1242.

Simpson GG, Filipowicz W (1996). Splicing of precursors to mRNA in higher plants: mechanism, regulation and subnuclear organisation of the spliceosomal machinery. Plant Molecular Biology 32:1-41.

Sinha N (1999). Leaf development in angiosperms. Annual Review of Plant Physiology and Plant Molecular Biology 50:419-46.

Sinha N (1999). Leaf development in Angiosperms. Annual Review of Plant Physiology and Plant Molecular Biology 50:419-446

Smith LG, Hake S, Sylvester AW (1996). The *tangled-1* mutantations alters cell division orietations throughout maize leaf development without altering leaf shape. Development 122:481-489.

Snow M, Snow R (1931). Experiments on phyllotaxis. I. The effect of isolating a primordium. Philosophical Transactions of the Royal Society of London, Series B 221:1-43.

Snow M, Snow R (1959). The dorsoventrality of leaf primordia. New Phytologist 58:188-207.

Souer E, Houwelingen AV, Kloos D, Mol J, Koes R (1996). The *No Apical Meristem* gene of Petunia is required for pattern formation in embryos and flowers and is expressed at meristem and primordia boundaries. Cell 85:159-70.

Spanjaard RA, Chen K, Walter JR, van Duin J (1990). Frameshift supression at tandem AGA and AGG codons by cloned tRNA genes: assigning a codon to argU tRNA and T4 $tRNA^{Arg}$. Nucleic Acids Research 18:5031-5036

Steeves TA, Briggs WR (1958) Morphogenetic studies on *osmundia cinnamomea* L. - The origin and early development of vegetative fronds. Phytomorphology 8:60-72.

Steves TA, Sussex IM (1989). Patterns in Plant Development. Cambridge University Press, New York.

Sussex IM (1955). Morphogenesis in *Solanum tuberosum* L.: Experimental investigations of leaf dorsoventrality and orientation in the juvenile shoot. Phytomorphology 5:286-300.

Templeton A (1983). Phylogenetic inference from restriction endonuclease site maps with particular reference to the evolution of humans and the apes. Evolution 37:221-244.

The *Arabidopsis* Information Resource (TAIR). http://www.Arabidopsis.org/home.html

The *Arabidopsis* Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana.* Nature 408: 796 – 815.

Traas J, Bellini C, Nacry P, Kronenberger J, Bouchez D, Caboche M (1995). Normal differentiation in cells lacking microtubular preprophase bands. Nature 375:676-677.

Trotochaud AE, Hao T, Wu Guang, Yang Z, Clark SE (1999). The CLAVATA1 receptor-like kinase requires CLAVATA3 for its assembly into a signalling complex that includes KAPP and a Rho-related protein. The Plant Cell 11:393-405.

Tschopp JF, Brust PF, Cregg JM, Stillman C, Gingeras TR (1987). Expression of the LacZ gene from two methanol regulated promoters in *Pichia pastoris.* Nucleic Acids Research 15:3859-3876

Tsuge T, Tsukaya H, Uchimiya H (1996). Two independent polarised processes of cell elengation regulate leaf blade expansion in *Arabidopsis thaliana* (L) Heyh. Development 122:1589-1600.

van den Berg C, Willemsen V, Hendricks G, Weisbeek P, Scheres B (1997). Short-range control of cell differentiation in the *Arabidopsis* root meristem. Nature 390:287-289.

van der Schoot C, Rinne P (1999). Networks for shoot design. Trends in Plant Science 4:31-37.

Veit B, Briggs SP, RJ Schmidt, Yanofsky MF, Hake S (1998). Regulation of leaf initiation by the *terminal ear1* gene of maize. Nature 393:166-168.

Wagner D. Sablowski RWM. Meyerowitz EM (1999). Transcriptional activation of *APETALA1* by *LEAFY*. Science 285(5427):582-584.

Waigman 1998 E, Cohen Y, McLean G, Zambryski P (1997). Plasmodesmata: gateways for information transfer. In: Greenland AJ, Meyerowitz EM, Steer M (Eds). Control of Plant Development: Genes and Signals. Symposia for the Society for Experimental Botany 51, The Company of Biologists Ltd., Cambridge UK.

Walbot V, Holder N (1987). Developmental Biology. Random House, New York.

Wang XQ, Hall TMT, Structural basis for recognition of AU-rich element RNA by the HuD protein (2001). Nature Structural Biology. 8:141-145.

Wardlaw CW (1949). Further experimental observations of the shoot apex of *Dryopteris aristata*. Philosophical Transactions of the Royal Society of London, Series B. 233:415-51.

Watanabe Y, Iino Y, Keito F, Shimoda C, Yamamoto M (1988). The *S. pombe* gene encoding a crucial molecule for commitment to meiosis is under the regulation of cAMP. The European Molecular Biology Organisation Journal 7: 761-767.

Watanabe Y, Shinozaki-Yabana, S, Chikashige Y, Hiraoka Y, Yamamoto M (1997). Phosphorylation of RNA-binding protein controls cell cycle switch from mitotic to meiotic in fission yeast. Nature 386: 187-190.

Watanabe Y, Yamamoto M (1994). S. pombe mei2 encodes an RNA-binding protein essential for premeiotic DNA synthesis and meiosis I, which cooperates with a novel RNA species meiRNA. Cell 78: 487-98.

White SE, Doebley JF (1999). Molecular evolution of *terminal ear1*, a regulatory gene in the genus *Zea*. Genetics 153:1455-1462.

Wolf S, Deom RN, Beachy RN, Lucas WJ (1989). Movement protein of tobacco mosaic virus modifies plasmodesmata size exclusion limit. Science 246:377-379.

Wolfe KH, Gouy M, Yand YW, Sharp PM, and Li WH (1989). Date of monocot-dicot divergence estimated from the chloroplast DNA sequence data. Proceedings of the National Academy of Sciences USA 86:6201-6205.

Wolpert L (1969). Positional information and the spatial pattern of cellular differentiation. Journal of Theoretical Biology 25:1-47.

Wolpert L (1998). In: Principles of development. Wolpert L, Beddington R, Brocks J, Jessel T, Lawrence P, Meyerowitz E. Current Biology Ltd (London). Oxford University Press (Oxford).

Xu RM. Jokhan L. Cheng XD. Mayeda A. Krainer AR (1997). Crystal structure of human up1, the domain of hnRNP A1 that contains two RNA-recognition motifs. Structure 5:559-570.

Yamashita A, Watanabe Y, Nukina N, Yamamoto (1998). RNA-assisted nuclear transport of the meiotic regulator *Mei2*p in fission yeast. Cell 95: 115-123.

Yang YW, Lai KN, Tai PY, Li WH. Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between *Brassica* and other angiosperm lineages. Journal of Molecular Evolution 48:597-604.

Ye GN, Stone D, Pang SZ, Creely W, Gonzalez K, Hinchee M (1999). *Arabidopsis* ovule is the target for Agrobacterium in planta vacuum infiltration transformation. The Plant Journal 19:249-57.

APPENDICES

Appendix 1. Oligonucleotide Primers

Primer	Sequence (5' TO 3')	Notes/position*
ANCHOR	GTAACTGCAG TCGACAAGCG AG	5' RACE
ANCHORT16V	GTAACTGCAG TCGACAAGCG AG$(T)_{15}V$	5' RACE
BV28	CTTCCTGTTC AAGGAGCC	*te1* genomic pos. 4674F
BV30	CCGGTGGCAG GTAAAGAAGC TGTA	*te1* genomic pos. 3619R
BV34	GGGAAGCAAA CCCAGAAAGG CGTC	*te1* genomic pos. 4597F
BV36	AAGGTGTACC CCGGCTGCGG CGCCATG	*te1* genomic pos. 3435R
BV41	ACTTACAGGA GCTCCTCCTC AGAAG	*te1* genomic pos. 4447R
DJ1	GCGCATATGC ATAATGGTTC GTAGGTA TGCTC	5' primer to amplify *ArgU* tRNA gene from pSBET vector.
DJ101	GGTGGGGGAG GGAGTGGAAT	Extreme 5' end of cloned *te1* cDNA (pBV432). *te1* genomic pos. 3223F
DJ108	CCGAACAAGT ACAGCCAGAA GC	Exon 3/exon4 border in *te1* cDNA.
DJ109	GCGCTTCCAG GCCTTGTA	Exon5/exon6 border in *te1* cDNA.
DJ2	GCGCATATGC ATTCAAAAGC CATTGAC TCAGC	3' primer to amplify *ArgU* tRNA gene from pSBET vector.
DJ201	CAGTATAGCT GTTGCGGAAG C	3' primer for primary PCR of 5' RACE. *te1* genomic pos. 3333R
DJ203	CACAGACCGT AGGGTGGAAC	3' primer for secondary PCR of 5' RACE. *te1* genomic pos. 3294R
DJ204	GTTTCTGGCA GCAGTGGTG	
GLUB-3	TAGGATCCTA ACGCCGCCGT TTTCGAT	3' TEL2 primer, BamHI linker
GLUX-5	ATTCTAGAGA TGTCTGTCAC CGGACCA	5' TEL2 primer, Xbal linker
M13 Forward	GTAAAACGAC GGCCAGT	M13 phage primer
M13 Reverse	GGAAACAGCT ATGACCATG	M13 phage primer
MUEND	GCGAATTCTT CGTCYATAAT GGCAATTATC	Degenerate consensus for *Mutator* element end
MUSEL	TCTATAATGG CAATTATCTC	*Mutator* element end primer
NLS_ANTI	CTAGAATGGC TCCTAAGAAG AAGAGAAAGG TTT	Encodes SV40 NLS anti-sense strand, with Xbal overhang
NLS_SENSE	CTAGAAACCT TTCTCTTCTT CTTAGGAGCC ATT	Encodes SV40 NLS sense strand, with Xbal overhang
OLIGOdT-20	TTTTTTTTTT TTTTTTTTTT	Reverse transcription primer for competitive RT-PCR

table continues on next page

Primer	Sequence (5' TO 3')	Notes/position*
RACE-RT	CTGTAACATG GCTATTTG	Reverse transcription primer for 5'RACE
SW1	ACAGCCGCTT CCGCAACAG	*te1* genomic pos. 3327F
SW4	TCCAGCATGT TGAGCAGCAG	*te1* genomic pos. 4899R
SW5	AGCTCGCCTT GACCTGTGTC	*te1* genomic pos. 3653F
TE1	AGCCATGAGCGCAAGACCA	*te1* genomic pos. 4537R
TE15	TCGGGCTCCT TGAACAGG	*te1* genomic pos. 4676R
TE19	GTCGCACGGG AACTTGGA	*te1* genomic pos. 6345R
TE22	GTGAACTCGA CGACGAG	*te1* genomic pos. 4169R
TEC10	GCCTAAGCTA GCTATAGC	*te1* genomic pos. 6963R
TG1	TAGCTATGGA GATGGAGCTG	*te1* genomic pos. 5603F
TG10	GGAGGCTAGT ACGTGCGTT	*te1* genomic pos. 1436F
TG11	GCCTGTGCAG TGACCA	*te1* genomic pos. 1889R
TG12	GACTTGGGAG AGGATTTG	*te1* genomic pos. 396F
TG2	TCATTGGGCA TTGATAGGTC	*te1* genomic pos. 5822R
TG3	CAATCCAGCA CACGGCCA	*te1* genomic pos. 2760F
TG4	GAGAACGGGA GAAGTGAGAC	*te1* genomic pos. 3063F
TG5	CACCATGATT AAGGGTAAGA	*te1* genomic pos. 7654R
TG7	GATTGTGGAC TTGGGAGAG	*te1* genomic pos. 389F
TG8	GACCGTCAAG CAGGAAAC	*te1* genomic pos. 1960R
TG9	GCTAGTACGT GCGTTTGATA	*te1* genomic pos. 1431R

* The 5' position in either the *te1* genomic sequence, or the pBV432 *te1* cDNA sequence, with F/R suffix to indicate orientation of primer.

Appendix 2. PCR Programs

Standard PCR program templates and non-standard programs.
Most PCR programs described in this thesis are vary only in annealing temperature, number of cycles, and the length of the extension time. I found it convenient to use a common system to name PCR programs during my work, this system is used in this thesis. Programs are named with the format DJT_axN_C(L), where T_a indicates the annealing temperature in degrees centigrade, N_C the number of cycles, and L is an optional parameter that indicates a longer than usual extension time. The length of the extension time is always specified. For example, the program DJ60x45L has an annealing temperature of 60∞C, 45 cycles, and a long extension time. The general parameters are described in the DJT_axN_C(L) template below.

Programs that differ significantly from the standard template are named and described individually.

DJTaxNc(L) template

step	Condition	note
1	98°C for 3:00 min.	initial denaturation
2	95°C for 0:30 min.	denaturation, addition of enzyme during first cycle at 95∞C for hot starts
3	Ta °C for 1:00 min.	annealing, temperature varies with program
4	70°C for 1:00 min.	extension, programs with the suffix L will have a longer extension time
5	GOTO step 2, Nc times	number of cycles varies
6	95°C for 0:30 min.	
7	Ta °C for 1:00 min.	
8	70°C for 10:00 min.	final extension, time invariant
9	hold at 4°C	

RACE 6

step	Condition	note
stage 1		PCR reactions contain only ANCHORT16V primer
1	98°C for 3:00 min.	Stage 1: initial denaturation
2	95°C for 0:30 min.	denaturation
3	33°C for 1:00 min.	annealing
4	68°C for 0:45 min.	Second strand cDNA synthesis, note multiple cycles
6	GOTO step 2, 10 times	
5	hold at 4°C	
stage 2		ANCHOR and DJ201 primers added at 95∞C
6	95°C for 0:30 min.	
7	66°C 1:00 min.	annealing
8	70°C for 1:00 min.	extension
9	GOTO step 6, 30 times	
10	95°C for 0:30 min.	
11	66°C 1:00 min.	
12	70°C for 10:00 min.	final extension
13	hold at 4°C	

RACE 7

step	Condition	note
stage 1		PCR reactions contain only ANCHORT16V primer
1	98°C for 3:00 min.	initial denaturation
2	95°C for 0:30 min.	denaturation
3	36°C for 10:00 min.	annealing
4	72°C for 2:00 min.	second strand cDNA synthesis
5	95°C for 0:30 min.	
6	72°C for 1:00 min. -1°C per cycle	'touchdown' annealing of ANCHORT16V primer
7	70°C for 1:00 min	extension
8	GOTO step 5, 9 times	
9	hold at 4°C	
stage 2		ANCHOR and DJ201 primers added at 95∞C
10	95°C for 0:30 min.	denaturation
11	66°C 1:00 min.	annealing
12	70°C for 1:00 min.	extension
13	GOTO step 10, 34 times	
14	95°C for 0:30 min.	denaturation
15	66°C 1:00 min.	annealing
16	70°C for 10:00 min.	extension

17	hold at 4°C	final extension

AUTOSEQ

step	condition	note
1	96°C for 10 seconds	denaturation
2	50°C for 5 seconds	annealing
3	60°C for 4 minutes	extension
4	GOTO step 1, 25 times	
5	hold at 4∞C	

DJ1

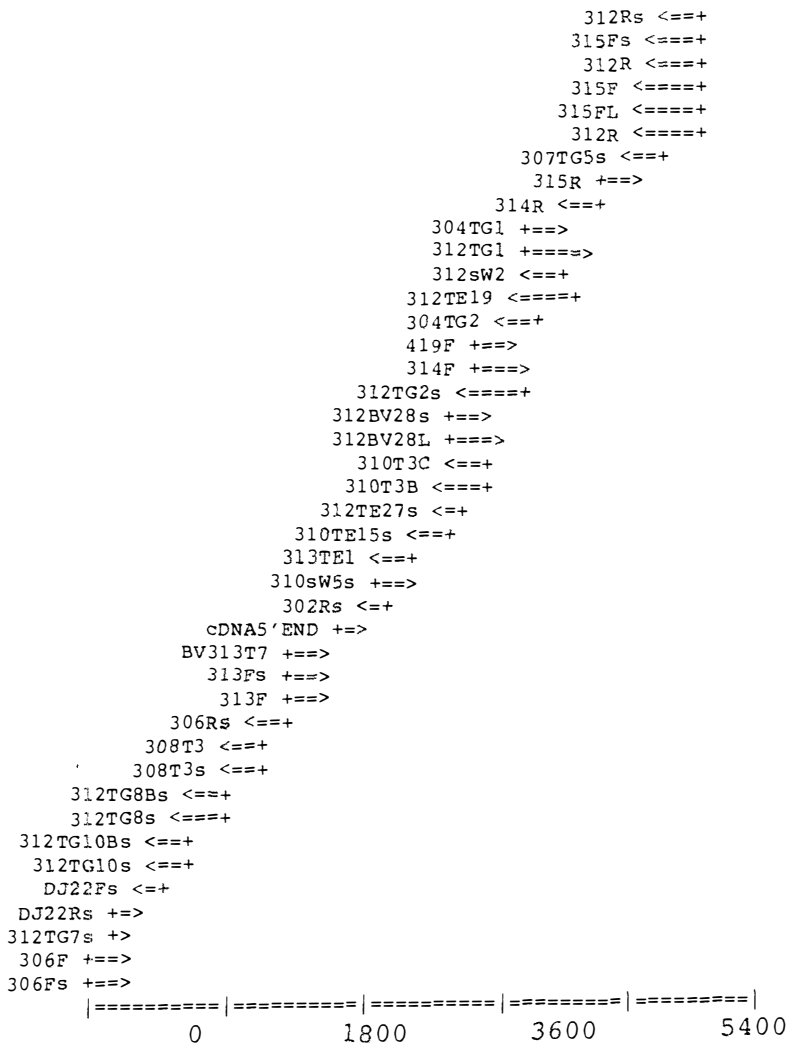
step	condition	note
1	96°C for 2:00	
2	96°C for 0:30	
3	50°C for 0:30	
4	68°C for 0:30	
5	GOTO 2, 25X	

Appendix 3. List of Suppliers

Supplier	Contact Details
EquiBio Ltd.	The Wheelwrights, Boughton Monchelsea Kent, MEA 4LT, UK
Qiagen	P.O. Box 25 Clifton Hill Victoria 3068 Austrailia www.qiagen.com
The Wisconsin Package Version 9.1.	Genetics Computer Group (GCG), Madison, Wisconsin USA
Ultra lum Inc.	Paramount, CA 90723 USA
Cole Parmer Instrument Co.	Chicago Ill USA
Anthos labtec instruments	Jakob Haringerstr, 8 A-5022 Salzburg.
Andrew Industrial Ltd.	3 Porana Rd. North Shore City, Auckland New Zealand.
Dynal, A.S.	P.O. Box 158 Sklyen,. N0212 Oslo, Norway www.dynal.no
Medicell International Ltd.	239 Liverpool Rd London N1 1LX England
Merk-Schuchardt.	Schuardt 85662 Hohenbrunn, Germany
Gibco BRL Anthos labtec instruments.	Jakob Haringerstr, 8 A-5022 Salzburg, Germany
Stratagene Cloning Systems.	11011 North Torrey Pines Rd La Jolla, CA, USA. www.stratagene.com

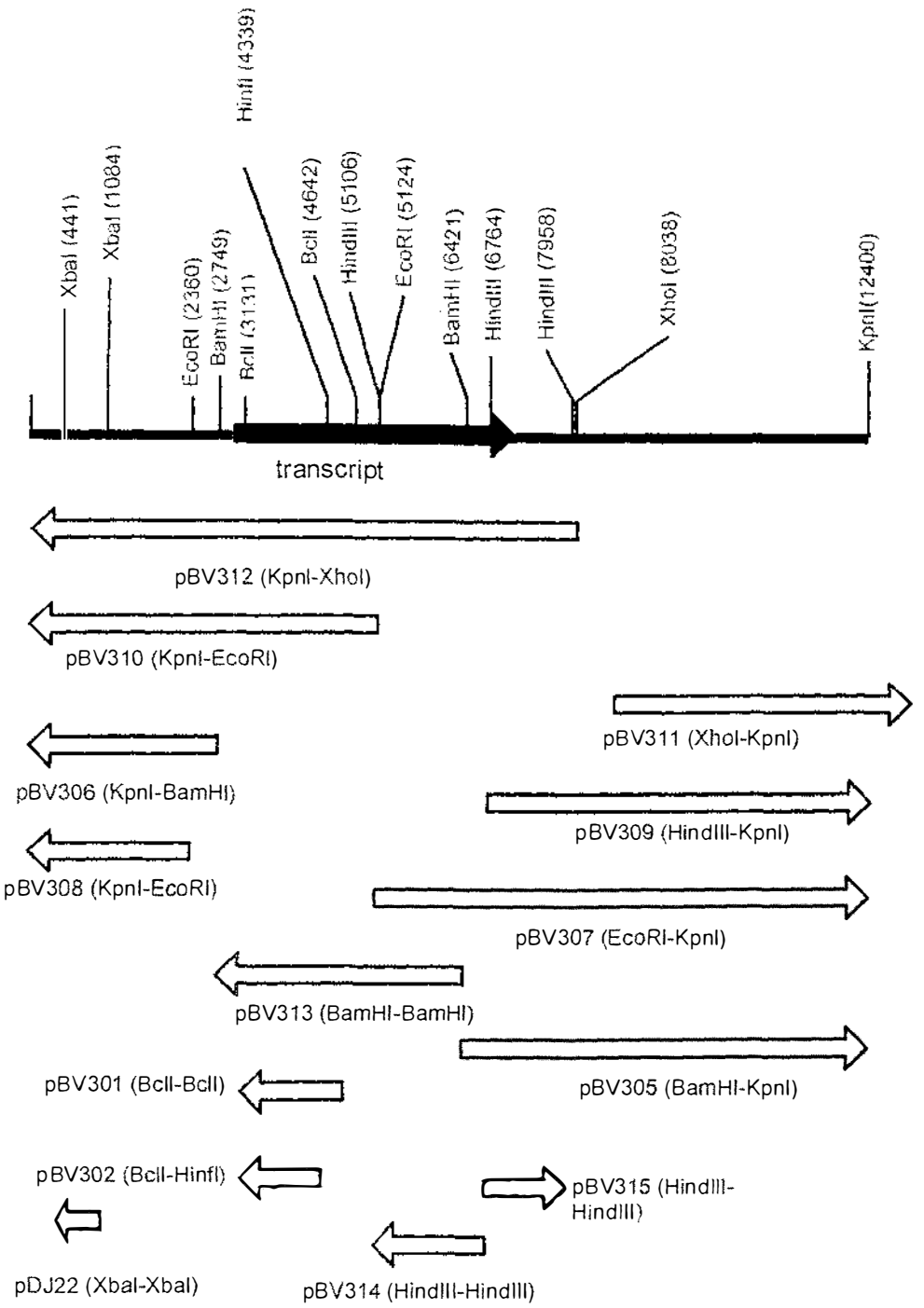
Appendix 4. Assembly of *te1* gene sequencing reactions.

Each sequencing reaction is denoted by an arrow representing the direction of the run, and named according to the *te1* genomic subclone used as a template, followed by the primer sued for sequencing (R, M13 reverse primer, F M13 forward primer, T7/T3 primers are present in the pBluescriptSK^{-} (Stratagene) vector. Other primers were designed to existing genomic sequence. The suffix s indicates that the sequence was trimmed to exclude ambiguous sequence. For example 312Rs used the pBV312 subclone as a template, the M13 reverse primer, and was trimmed.



720

Appendix 5. Subclones of the 12.4 kb KpnI *te1* genomic clone.



Arrows delineate the position of the M13 forward primer in the vector.

Appendix 7. Pedigrees of maize stocks.

The Microsoft Excel file (supplementary file: *DJMaizeStocks.xl*) describes the pedigrees, alleles, and sources of all maize stocks used n this study. Further details can be obtained from Bruce Veit.

Appendix 8. PCR products from *te1* mutant alleles

Allele	Primer set	Range	Result	Notes
te1-1	TG7/SW4	389 - 4899	no product	
	TG3/TE15	2760 - 4676	no product	
	DJ101/TE15	3323 - 4676	Faint 0.4 kb band	expect 1.6 kb band
	SW1/TE15	3327 - 4676	no product	
	SW1/SW4	3327 - 4899	no product	
	SW5/SW4	3653 - 4899	as B73	
	TG1/TE19	5603 - 6345	as B73	
	TG1/TG5	5603 - 7654	as B73	
te1-ems	TG3/TE15	2760 - 4676	as B73	
	SW1/SW4	3327 - 4899	as B73	
	BV28/TEC10	4674 - 6963	no product	inconsistent results
	BV34/TEC10	4597 - 6963	no product	inconsistent results
	TG1/TG5	5603 - 7654	as B73	
*te1-mum3**	MUEND/BV41	Mutator insertion - 4447	~500 bp product	BV family 2154-5 (hetrozygote). PCR product cloned and sequenced: *te1* genomic position 4017.
*te1-mum4**	TG7/SW4	389-4899	no product	
	TG3/TE15	2760-4676	no product	
	SW1/TE15	3327-4676	no product	
	SW1/SW4	3327-4899	no product	
	SW5/SW4	3653-4899	as B73	
te1-mum5	SW1/SW4	3327-4899	~ 2.4 kb product	Expected size, produced by B73, is 1.5 kb. Product analysed by restriction digestion (Section 2.13c).
te1-mum6	TG7/SW4	389-4899	no product	
	TG3/TE15	2760-4676	no product	
	SW1/SW4	3327-4899	no product	
	SW5/SW4	3653-4899	as B73	
	TG1/TE19	5603-6345	as B73	
te1-mum7	TG7/SW4	389-4899	no product	
	TG3/TE15	2760-4676	as B73	
	SW1/SW4	3327-4899	as B73	
	BV34/TEC10	4597-6963	as B73	
	TG1/TG5	5603-7654	as B73	

* The numbering used by the Veit group for *te1-mum3* and *te1-mum4* alleles was interchanged. The new numbering is used in this thesis.

Appendix 9. Southern Blot Data

Appendix 9.1. Southern blot one.

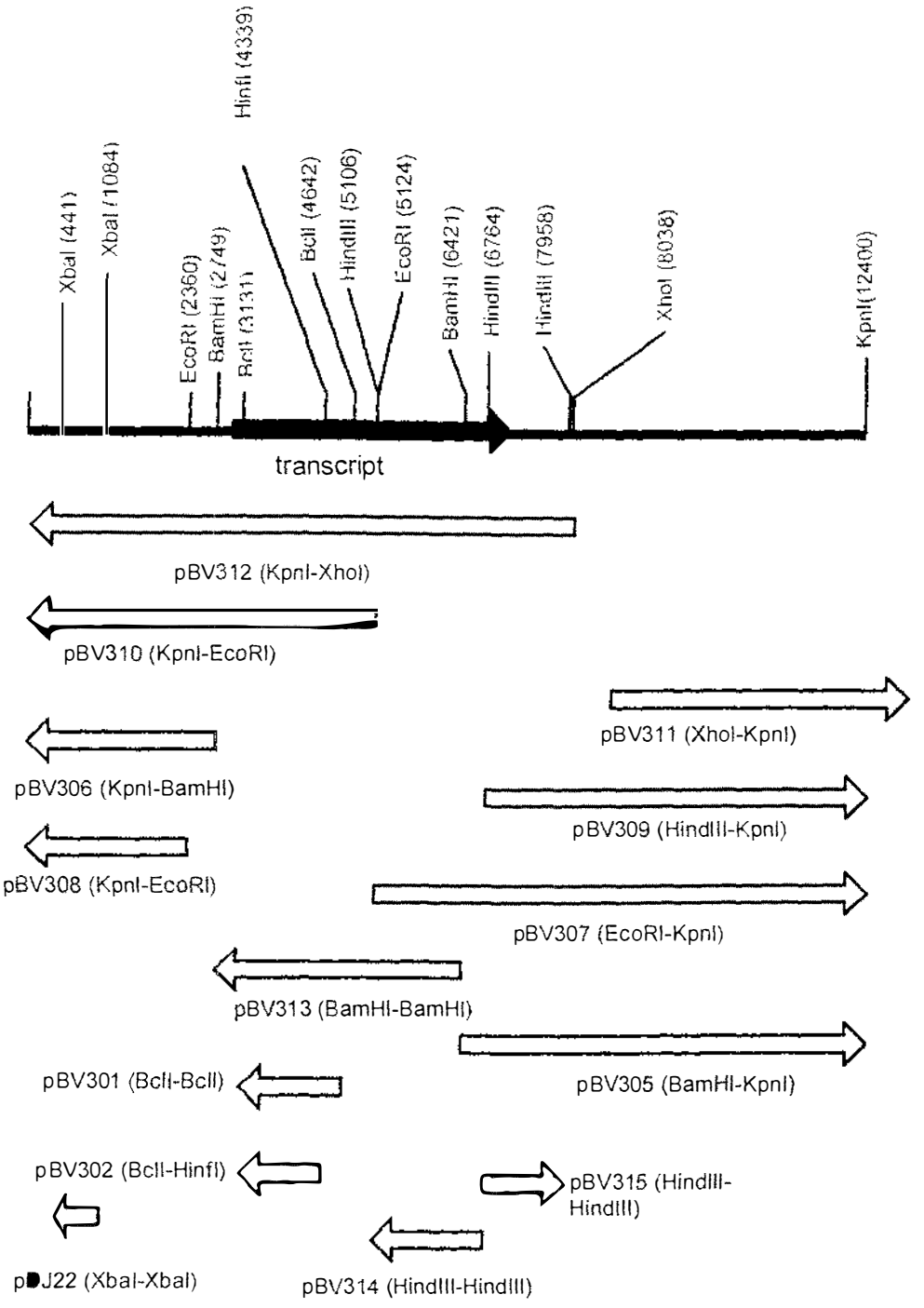
Enzymes: HindIII.
Probe middle *te1* cDNA (4597-6345)

Lane	Digest	Allele	Family	Molecular weights of fragments		Notes
	HindIII	*B73*	*Expected*	*5106+*	*1658*	
3		B73	-	7700	1600	as expected
4		*te1*-1 (A188)	DJ96:17	2200		
5		*te1*-1	DJ96:22	7800	2200	
6		*te1*-1 or *te1*-mum7	DJ96:16	5400	1600	
7		*te1*-ems (A188)	DJ96:15	7800	1600	
8		*te1*-ems	DJ96:26	7800	1600	
9		*te1*-mum1	DJ96:23	9200	1600	
10		*te1*-mum1	DJ97:18	9200	1600	
11		*te1*-mum1 or 2	DJ97:51	2800		not clear, does not seem to be either –mum1 or mum2
12		*te1*-mum2	DJ96:25	9200	1600	
13		*te1*-mum2	DJ97:57	9200	1600	
14		*te1*-mum3	DJ97:43	-		not clear
15		*te1*-mum3 or 6	DJ96:4	2200	1600	is *te1-mum3*
16		*te1-mum3* or 6	DJ96:3	2200	1600	is *te1-mum3*
17		*te1*-mum3 or 6	DJ96:2	2200	1600	is *te1-mum3*
18		*te1*-mum3 or 4	DJ97:22	2200	1600	is *te1-mum3*
19		*te1*-?(W22)	DJ96:40	2700	1600	

Appendix 9.2. Southern blot two.
Enzymes: KpnI, EcoRI, HindIII.
Probe: the entire *te1* cDNA (corresponding to *te1* genomic region 3327 to 6980)

Lane	Digest	Allele	Family	Molecular weights of fragments				Notes
	KpnI	*B73*	*Expected*	*12400*				
3		B73	-	12900				as expected
4		*te1*-1	BV97:32	15000				
5		*te1*-1 (A188)	DJ96:22	16100				
6		*te1*-1	NA97:32	12900				
7		*te1*-mum1	DJ96:23	14400				
8		*te1*-mum2	DJ96:25	14700				
9		*te1*-mum3	DJ96:2	14700				
10		*te1*-mum3 or 5	DJ98:17	13500	12400			is *te1-mum5*
11		*te1*-1 or mum7	DJ96:16	13800	12400			is *te1-mum7*
12		*te1*-mum7	NA2331	13200	12100			
	EcoRI	*B73*	*Expected*	*7280+*	*2760*			
13		B73	-	10800	10400	9500	3200	extra bands not as expected
14		*te1*-1	BV97:32	10600	4600			
15		*te1*-1 (A188)	DJ96:22	10600	4600			
16		*te1*-1	NA97:32	12600	8200			
17		*te1*-mum1	DJ96:23	12100	10200	9500	5400	
18		*te1*-mum2	DJ96:25	12100	10200	9500	5400	
19		*te1*-mum3	DJ96:2	12600	10200	4500		
20		*te1*-mum3 or 5	DJ98:17	13800				
21		*te1*-1 or mum7	DJ96:16	13800				
22		*te1*-mum7	NA2331	13200	9900	9300		
	HindIII	B73	*Expected*	*5106+*	*1658*			
23		B73	-	9300	1600			as expected
24		*te1*-1	BV97:32	-				no bands visible
25		*te1*-1 (A188)	DJ96:22	9500	8700			
26		*te1*-1	NA97:32	13200				
27		*te1*-mum1	DJ96:23	10800	1600			
28		*te1*-mum2	DJ96:25	10800	1600			
29		*te1*-mum3	DJ96:2	11300	2300			
30		*te1*-mum3 or 5	DJ98:17	6400				
31		*te1*-1 or mum7	DJ96:16	6400				
32		*te1*-mum7	NA2331	6300				

Appendix 5. Subclones of the 12.4 kb KpnI *te1* genomic clone.



Arrows delineate the position of the M13 forward primer in the vector.

Appendix 7. Pedigrees of maize stocks.

The Microsoft Excel file (supplementary file: *DJMaizeStocks.xl*) describes the pedigrees, alleles, and sources of all maize stocks used n this study. Further details can be obtained from Bruce Veit.

Appendix 8. PCR products from *te1* mutant alleles

Allele	Primer set	Range	Result	Notes
te1-1	TG7/SW4	389 - 4899	no product	
	TG3/TE15	2760 - 4676	no product	
	DJ101/TE15	3323 - 4676	Faint 0.4 kb band	expect 1.6 kb band
	SW1/TE15	3327 - 4676	no product	
	SW1/SW4	3327 - 4899	no product	
	SW5/SW4	3653 - 4899	as B73	
	TG1/TE19	5603 - 6345	as B73	
	TG1/TG5	5603 - 7654	as B73	
te1-ems	TG3/TE15	2760 - 4676	as B73	
	SW1/SW4	3327 - 4899	as B73	
	BV28/TEC10	4674 - 6963	no product	inconsistent results
	BV34/TEC10	4597 - 6963	no product	inconsistent results
	TG1/TG5	5603 - 7654	as B73	
*te1-mum3**	MUEND/BV41	Mutator insertion - 4447	~500 bp product	BV family 2154-5 (hetrozygote). PCR product cloned and sequenced: *te1* genomic position 4017.
*te1-mum4**	TG7/SW4	389-4899	no product	
	TG3/TE15	2760-4676	no product	
	SW1/TE15	3327-4676	no product	
	SW1/SW4	3327-4899	no product	
	SW5/SW4	3653-4899	as B73	
te1-mum5	SW1/SW4	3327-4899	~ 2.4 kb product	Expected size, produced by B73, is 1.5 kb. Product analysed by restriction digestion (Section 2.13c).
te1-mum6	TG7/SW4	389-4899	no product	
	TG3/TE15	2760-4676	no product	
	SW1/SW4	3327-4899	no product	
	SW5/SW4	3653-4899	as B73	
	TG1/TE19	5603-6345	as B73	
te1-mum7	TG7/SW4	389-4899	no product	
	TG3/TE15	2760-4676	as B73	
	SW1/SW4	3327-4899	as B73	
	BV34/TEC10	4597-6963	as B73	
	TG1/TG5	5603-7654	as B73	

* The numbering used by the Veit group for *te1-mum3* and *te1-mum4* alleles was interchanged. The new numbering is used in this thesis.

Appendix 9. Southern Blot Data

Appendix 9.1. Southern blot one.

Enzymes: HindIII.
Probe middle *te1* cDNA (4597-6345)

Lane	Digest	Allele	Family	Molecular weights of fragments		Notes
	HindIII	*B73*	*Expected*	*5106+*	*1658*	
3		B73	-	7700	1600	as expected
4		*te1*-1 (A188)	DJ96:17	2200		
5		*te1*-1	DJ96:22	7800	2200	
6		*te1*-1 or *te1*-mum7	DJ96:16	5400	1600	
7		*te1*-ems (A188)	DJ96:15	7800	1600	
8		*te1*-ems	DJ96:26	7800	1600	
9		*te1*-mum1	DJ96:23	9200	1600	
10		*te1*-mum1	DJ97:18	9200	1600	
11		*te1*-mum1 or 2	DJ97:51	2800		not clear, does not seem to be *either* –mum1 *or* mum2
12		*te1*-mum2	DJ96:25	9200	1600	
13		*te1*-mum2	DJ97:57	9200	1600	
14		*te1*-mum3	DJ97:43	-		not clear
15		*te1*-mum3 or 6	DJ96:4	2200	1600	is *te1-mum3*
16		*te1*-mum3 or 6	DJ96:3	2200	1600	is *te1-mum3*
17		*te1*-mum3 or 6	DJ96:2	2200	1600	is *te1-mum3*
18		*te1*-mum3 or 4	DJ97:22	2200	1600	is *te1-mum3*
19		*te1*-?(W22)	DJ96:40	2700	1600	

Appendix 9.2. Southern blot two.
Enzymes: KpnI, EcoRI, HindIII.
Probe: the entire *te1* cDNA (corresponding to *te1* genomic region 3327 to 6980)

Lane	Digest	Allele	Family	Molecular weights of fragments				Notes
	KpnI	*B73*	*Expected*	*12400*				
3		B73	-	12900				as expected
4		*te1*-1	BV97:32	15000				
5		*te1*-1 (A188)	DJ96:22	16100				
6		*te1*-1	NA97:32	12900				
7		*te1*-mum1	DJ96:23	14400				
8		*te1*-mum2	DJ96:25	14700				
9		*te1*-mum3	DJ96:2	14700				
10		*te1*-mum3 or 5	DJ98:17	13500	12400			is *te1-mum5*
11		*te1*-1 or mum7	DJ96:16	13800	12400			is *te1-mum7*
12		*te1*-mum7	NA2331	13200	12100			
	EcoRI	*B73*	*Expected*	*7280+*	*2760*			
13		B73	-	10800	10400	9500	3200	extra bands not as expected
14		*te1*-1	BV97:32	10600	4600			
15		*te1*-1 (A188)	DJ96:22	10600	4600			
16		*te1*-1	NA97:32	12600	8200			
17		*te1*-mum1	DJ96:23	12100	10200	9500	5400	
18		*te1*-mum2	DJ96:25	12100	10200	9500	5400	
19		*te1*-mum3	DJ96:2	12600	10200	4500		
20		*te1*-mum3 or 5	DJ98:17	13800				
21		*te1*-1 or mum7	DJ96:16	13800				
22		*te1*-mum7	NA2331	13200	9900	9300		
	HindIII	B73	*Expected*	*5106+*	*1658*			
23		B73	-	9300	1600			as expected
24		*te1*-1	BV97:32	-				no bands visible
25		*te1*-1 (A188)	DJ96:22	9500	8700			
26		*te1*-1	NA97:32	13200				
27		*te1*-mum1	DJ96:23	10800	1600			
28		*te1*-mum2	DJ96:25	10800	1600			
29		*te1*-mum3	DJ96:2	11300	2300			
30		*te1*-mum3 or 5	DJ98:17	6400				
31		*te1*-1 or mum7	DJ96:16	6400				
32		*te1*-mum7	NA2331	6300				

Appendix 9.3. Southern blot three[1]
Enzymes: HindIII, EcoRI,
Probe: 5' end *te1* cDNA (corresponding to *te1* genomic region 3327 to 4537)

Lane	Digest	Allele	Family				
-	HindIII	*B73*	*Expected*	*5106+*			
2		*te1*-mum3	BV99:5	-			bands not visible
3		*te1*-mum6 (A188)	BV99:9	10600	1900		
4		*te1*-mum7	BV99:13	5500			
5		A188	BV99:16	11200			
6		*te1*-? (mum6?)	BV99:20	11200			
7		*te1*-mum1	BV99:21	10200			
8		*te1*-mum2	BV99:23	10400			
9		*te1*-mum3	BV99:25	12900			
10		*te1*-mum6	BV99:26	11400	2100		
11		*te1*-mum7	BV99:29	6000			
12		*te1*-mum5	BV99:32	12600	6300		
13		B73	-	8600			as expected
	EcoRI	*B73*	*Expected*	*2760*			
21		*te1*-mum3	BV99:5	6600	5100		
22		*te1*-mum6 (A188)	BV99:9	9500	2500		
23		*te1*-mum7	BV99:13	11200			
24		A188	BV99:16	11200	4500		
25		*te1*-? (mum6?)	BV99:20	7100	5600		
26		*te1*-mum1	BV99:21	12600	6900	5500	
27		*te1*-mum2	BV99:23	9900	4000	2400	
28		*te1*-mum3	BV99:25	9500	3700	2300	
29		*te1*-mum6	BV99:26	9700	3900	2400	
30		*te1*-mum7	BV99:29	12900			
31		*te1*-mum5	BV99:32	12900			
32		B73	-	9500 (faint)	2500		extra (faint) band not as expected

[1] This Southern blot was produced by Carmel Gillman

Appendix 10. Supplementary file disk

A CD is included in the back cover of this thesis. It contains the following Macintosh format files (all are Microsoft Word Files, except DJ maize lines):

Thesis chapters 1 to 6

Alignments

Gene index alignment
Pneumocystis alignment

Gene predictions
AML3-prediction
AML4-prediction
AML5-prediction

Maize stocks
DJ maize lines (Microsoft Excel Document)

Sequences (all in FASTA format)
AML1-cDNA
AML2-cDNA
te1-cDNA
TEL1-cDNA
TEL2b-cDNA
AML1-genomic
AML2-genomic
AML3-genomic
AML4-genomic
AML5-genomic
AML6-genomic
AML7-genomic
TEL1-genomic
TEL2-genomic
te1-genomic
MEI2-like proteins - FASTA

Appendix 11. The concurrent RNA world project

A number of papers were published as part of an unrelated project that investigated the evolution of early life from RNA world ancestors. This work produced three papers, and a peer-reviewed internet page. These papers are **not** submitted as a requirement for the for the degree, and should not be examined as such. They are included to indicate that the author followed an interest in other aspects of biology.

Poole AM, Jeffares DC, Penny D (1998). The Path from the RNA world. Journal of Molecular Evolution 46:1-17.

Jeffares DC, Poole AM, Penny D (1998) Relics from the RNA world. Journal of Molecular Evolution 46:18-36

Poole AM, Jeffares DC, Penny D (1999) Early Evolution: Prokaryotes, the New Kids on the Block. BioEssays 21(10):880-889.

Jeffares DC, Poole AM (2001). Were bacteria the first forms of life on earth? http://www.actionbioscience.org/ageofbio/jeffares_poole.html