The Identification and Characterisation of Key Genes Involved in Somatic Embryogenesis in Wheat (*Triticum, aestivum L.*).

Ben Michael Minogue

Submitted in accordance with the requirements for the degree of PhD

January 2007

The University of Leeds

School of Biological Sciences

THESIS CONTAINS

CD

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others

'We shall not cease from exploration

And the end of all our exploring

Will be to arrive where we started

And know the place for the first time.'

T.S. Eliot, 'Little Gidding'

Acknowledgements

I would like to thank my supervisors Dr Hanma Zhang at The University of Leeds and Dr Dong-Fang Chen at Syngenta for providing me with the opportunity to undertake such an interesting and challenging project and for all their support and advice during the research and writing processes. I would also like to thank the wheat transformation team at Syngenta, Jealotts Hill, Berkshire for their advice and help in undertaking some of the practical aspects of the project, in particular Dr Paul Drayton and Duncan Oliver. I would also like to give special thanks Dr Sinead Dray and Dr John Doonan at the John Innes Centre, Norwich for facilitating the high-throughput in situ hybridisation experiments. My thanks again go to Dr John Doonan and also to Dr Andrew Cuming for agreeing to examine this work and for providing their comments and suggestions.

Finally, I would like to thank my friends and family for their friendship and patience during the research and writing up process. In particular I would like to give special thanks to my wife Sally for her constant support and encouragement.

This research would not have been possible without funding from the BBSRC and Syngenta.

Ben Minogue.

Abstract

One of the most remarkable aspects of plants cells is that on application of the plant hormone auxin, they possess the unique ability to de-differentiate and return to embryonic development, via a process known as 'somatic embryogenesis' (SE). On removal of auxin these totipotent cells are then capable of re-differentiating and regenerating into full adult plants. Although well characterised physiologically, the molecular mechanisms that regulate this important biological process are still very poorly understood. The research aim of this project was to improve our understanding of SE, by designing a series of cDNA microarray experiments to identify the genetic components responsible for the de-differentiation of somatic cells to an embryonic state and the initiation, development and differentiation of somatic embryos towards adult structures. To achieve this aim, suitable embryogenic and non-embryogenic systems were established and used for subtractive microarray analysis that has identified 701 differentially expressed genes that are specifically related to SE in wheat. Following bioinformatic and expression analysis, 57 of these genes were short listed and further characterised by *in situ* hybridisation (ISH) to provide further information of their temporal and spatial patterns of expression. A small subset of genes that demonstrated interesting patterns of expression, were further characterised and functionally analysed via RNAi silencing studies. Here it is demonstrated that all of the genes tested for RNAi silencing results in a reduction of the transformation and regeneration efficiencies. Furthermore the targeted silencing of several of these novel genes was shown to have an observable phenotypic effect on the development of somatic embryos and regeneration in wheat. Of particular interest is a wheat gene related to the rice embryo specific protein OSE 731, which appears to play an essential role in the transition of somatic embryos into differentiating adult structures. The silencing of this gene resulted in the continuous proliferation of somatic embryos (SEs) and arrested the development of these SEs into differentiated adult structures.

Abbreviations

A-B	Apical-Basal
ABA	Abscisic Acid
ABC	ATP Binding Cassette
ABI	Abscisic Acid Insensitive
ABREBF	Auxin Binding Responsive Element Binding Factor
AGL	Agamous-Like
AGO	Argonaut
AGP	Arabinogalactan Protein
AP	Alkaline Phosphatase
AP2	Apetala 2
ARF	Auxin Response Factors
ASF	Anti-Silencing Function
ATP	Adenosine Triphosphate
At	*Arabidopsis thaliana*
AUX/IAA	Auxin/Indole-3-Acetic Acid
BBM	Baby Boom
BCIP	5-Bromo-4-Chloro-3'-Indolyphosphate p-Toluidine Salt
BDL	Bodenlos
BW	Bobwhite
BSA	Bovine Serum Albumin
BZip	Basic Leucine Zipper
Ca	Calcium
Ca_2Cl	Calcium Chloride
CDC2	Cell Division Cycle 2
CDPK	Calcium Dependent Protein Kinase
cDNA	Complementary Deoxy-ribonucleic Acid
CHD3	Chromodomain Helicase DNA-binding protein
CMV	Cauliflower Mosaic Virus
CRK	Cyclin Responsive Kinase
CTAB	Hexadecyltrimethylammonium Bromide
Cy3	Cyan3
Cy5	Cyan5
dATP	2'-deoxyadenosine 5'-triphosphate
dCTP	2'-deoxycytidine 5'-triphosphate
dGTP	2'-deoxyguanosine 5'-triphosphate
dTTP	2'-deoxythymidine 5'-triphosphate
dNTPs	Deoxynucleotide triphosphates
DAP	Days After Pollination
Dc	*Daucus Carota*
DEV	Differential Expression Value
DEPC	Diethyl Pyrocarbonate
DGE	*Dactilys Glomerata* Embryogenesis
DIG	Digoxigenin
DMSO	Dimethyl Sulfoxide
DNA	Deoxy-ribonucleic Acid
DSPTP1	Dual Specific Protein Phosphatase
DTT	Dithiothreitol
ECFP	Enhanced Cyan Fluorescent Protein

E.coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic Acid
EIF	Eukaryotic Initiation Factor
ESC	Embryonic Stem Cell
EST	Expressed Sequence Tag
EtOH	Ethanol
FK	Fackel
Fe	Iron
GA/GA3	Gibberellic Acid
GBPLN	Genbank Plant Database
GEM	Gene Expression Microarray
GFP	Green Fluorescent Protein
GK	Gurke
GlcNAc	N-acetyl-D-glucosamine
GM	Genetically Modified
GN	Gnom
GSP	Gene Specific Primer
H_20	Water
H_2SO_4	Sulfuric Acid
HCl	Hydrochloric Acid
HD	Homeodomain
HMG	High Mobility Group
hpRNA	Hair Pin RNA
HSP	Heat Shock Protein
HTH3	*Hordeum vulgare* gamma-thionin
HYL1	Hyponastic Leaves
IA	Isoamyl Alcohol
IAA	Indole-3-Acetic Acid
IEs	Immature Embryos
ISH	*In Situ* Hybridisation
K	Potassium
KAPP	Kinase Associated Protein Phosphatase
KOH	Potassium Hydroxide
LEA	Late Embryogenesis Abundant
LEC	Leafy Cotyledon
LiCl	Lithium Chloride
LRRK	Leucine Rich Receptor Kinase
LTP	Lipid Transfer Protein
MAPK	Mitogen Activated Protein Kinases
MCM	Minichromosome Maintenance
MP	Monopterous
MPF	Maturation-Promoting Factor
miRNA	Micro RNA
miRNP	miRNA Containing Effector Complex
mRNA	Messenger RNA
MS	Murashige and Skoog
N	Nitrogen
NAA	Naptheleneacetic acid
NADPH	Nicotinamide Adenine Dinucleotide Phosphate
NaOAc	Sodium Acetate

NaOH	Sodium Hydroxide
NaCl	Sodium Chloride
NaOCl	Sodium Hypochlorite
$NaPO_4$	Sodium phosphate
NBT	Nitro-Blue Tetrazolium Chloride
NTP	Nucleotide Triphosphates
NOS	Nopaline Synthase
OC	Organising Centre
OD	Optical Density
OPT	Oligopeptide Transporter
OSE	*Oryza Sativa* Embryogenesis
P	Phosphate
PBS	Phosphate Buffered Saline
PCNA	Proliferating Cell Nuclear Antigen
PCR	Polymerase Chain Reaction
PEM	Pro-Embryogenic Mass
pH	Power of Hydrogen
PIN1	Pin Formed 1
Pi-PLC	Phosphoinositide-specific phospholipase C
PKL	Pickle
PMI	Phosphomannose Isomerase
PTGS	Post Transcriptional Gene Silencing
Poly A RNA	Poly Adenylated Ribonucleic Acid
PPI	Polyphosphoinositides
PP2C	Protein Phosphatase 2C
PVP	Polyvinyl Pyrrolidone
QC	Quiescent Centre
RAM	Root Apical Meristem
RDR	RNA Dependent RNA Polymerase
RISC	RNA-Induced Silencing Complex
RITS	RNA-Induced Transcriptional Silencing Complex
RNA	Ribonucleic Acid
RNAi	RNA Interference
RNAse	Ribonuclease
rRNA	Ribosomal Ribonucleic Acid
RPM	Revolutions Per Minute
RT	Room Temperature
RT-PCR	Reverse Transcription-Polymerase Chain Reaction
SAM	Shoot Apical Meristem
SAR	Systemic Acquired Resistance
SAV	Savannah
SCAMP	Secretory Carrier Membrane Protein
SCR	Scarecrow
SDS	Sodium Dodecyl Sulphate
SE	Somatic Embryogenesis
SEs	Somatic Embryos
Secs	Seconds
SERK	Somatic Embryogenesis Receptor Kinase
SET	Su(var)3-9, Enhancer-of-zeste, Trithorax
SIP	Seed Imbibition Protein

siRNA	Short Interfering RNA
SSC	Saline Sodium Citrate
STM	Shoot Meristemless
STPK	Serine/Threonine Protein Kinase
Taq	*Thermus Aquaticus*
TE	TRIS-EDTA
TBS	Tris-buffered Saline
TLK	Tousled Like Kinases
TSL	Tousled
TP	Timepoint
TRIS	Hydroxymethylaminoethane
UDP	Uridine-diphosphate
UK	United Kingdom
UV	Ultra Violet
VIP	Vacuum Infiltration Processor
VP1	Viviparous 1
Zn	Zinc
^{0}C	Degrees Celsius
%	Per Cent
2,4-D	2,4-Dichlorophenoxyacetic Acid
2-D PAGE	Two-Dimensional Polyacrylamide Gel Electrophoresis
ß-ME	Beta-Mercaptoethanol
ml	Millilitre
µl	microlitre
nl	nanolitre
mM	millimole
µM	micromole
nM	nanomole
mg	milligram
µg	microgram
ng	nanogram
mm	millimetre
µm	micrometre
nm	nanometre

Table of Contents

List of Figures

List of Tables

Chapter One

Introduction

1 Chapter One - Introduction

1.1. Introduction

One of the most interesting aspects of plants is that their highly differentiated somatic cells possess the unique ability to de-differentiate and return to embryonic development, via a process known as 'somatic embryogenesis' (SE). Theoretically, under the right experimental conditions, all living somatic plant cells have the potential to undergo SE and to form embryos which, like their zygotic counterpart, can germinate and develop into fully-grown adult plants. In some plant species, such as *Kalanchoë,* SE occurs naturally and somatic embryos (SEs) can be observed to form spontaneously on the edge of leaves (Dodeman *et al.,* 1997). For the majority of plant species, however, SEs are only formed under specific culture conditions. For many agriculturally important crops such as wheat and rice, SE offers the only efficient means of plant regeneration and propagation, and is therefore of fundamental importance to modern biotechnology in the generation and maintenance of transgenic plants. Furthermore, many crop species and elite cultivars remain recalcitrant to tissue culture. Therefore, an understanding of the underlying regulatory mechanisms responsible for SE will not only significantly improve our scientific knowledge, but also facilitate the development of new and efficient tissue culture and plant propagation techniques which could be applied to aid the generation of commercial GM crops.

1.2. Wheat Agriculture and Biotechnology

Cereals are amongst the oldest plants to be domesticated. The unique properties of wheat flours and the wide geographical distribution of wheat make it one of the three most important crops in the world (Rakszegi *et al.,* 2001). The first steps in wheat existence began around 10,000 years ago when natural hybridisation between wild grasses occurred. The name 'wheat' is given to many species within the genus Triticum. Cultivated species within this genus are grouped by number of chromosomes, 14 (diploid), 28 (tetraploid) or 42 (hexaploid). Currently, the most abundant wheat is the hexaploid bread wheat, *Triticum aestivum var. aestivum.* All modern wheat varieties have been through several phases of artificial selection by man. First, there would have been subconscious selection by early growers followed

by primitive farmers deliberately selecting plants with advantageous phenotypes. More recently there have been scientifically planned breeding programs which have led to the production of high-yielding hybrid seed varieties and resulted in dramatic increases in crop yields. This form of selective breeding known as the 'green revolution' which took place between the 1950's and 1980's requires a number of generations and is normally slow. In addition, many traits of interest are controlled by multiple genes, each having a small effect, and undesirable genes are usually selected along with the gene of interest, making selective breeding unpredictable and imprecise. Since the 1970's, tissue culture and transformation techniques have rapidly developed making it possible to produce wheat through genetic modification. Genetic engineering technology allows a single gene or cluster of genes to be transferred to an organism directly in a much quicker and more precise manner than conventional breeding. Genetic engineering also allows the introduction of foreign genes (those from another organism) and therefore offers the potential to dramatically improve the quality of wheat, for example, by increasing grain size, herbicide tolerance and insect or disease resistance. Improving the efficiency of generating genetically modified wheat will allow for cheaper production costs and the ability to use more difficult elite varieties. Although there are no commercial GM varieties of wheat currently growing in the UK due to the strong controversy surrounding GM crops, ministers have declared that such crops can, on certain conditions, be allowed to be grown in the UK.

1.3. The History of Somatic Embryogenesis

Somatic embryogenesis (SE) exemplifies the concept of cell totipotency; a fundamental theory that can be traced back as early as the 19th century when Gottlieb Haberlandt (1854-1945), a German plant physiologist, put forward his radical idea that individual cells of higher plants contained all the information necessary to produce a fully formed adult. This 'vision' has accredited Haberlandt as the grandfather of SE. However, it was not until the pioneering work of F.C. Steward and J. Reinert in the late 1950s, who independently demonstrated that somatic cells of cultivated carrots could produce embryo-like structures in aseptic culture, that the process of SE was truly discovered (Reinert, 1958; Steward *et al.*, 1958). Early studies of SE were mainly carried out using the two systems: carrot and alfalfa suspension cells. However, the number of plant species in which SE can be induced

has greatly increased in the past 10-15 years, and investigations are now conducted in a whole range of agriculturally and horticulturally important plants (Zimmerman, 1993).

1.4. Plant Embryogenesis

Embryogenesis is a fundamental process of life that has been studied extensively in both animals and plants. In animals, the mature embryo is a miniature of the adult with a clearly established body plan. Unlike animals, development in plants is largely post-embryonic and adult organs are often not present in the plant embryo but are formed later in development from the juvenile seedlings' root and shoot meristems. The first systematic observation of embryo development in dicotyledons (dicots) and monocotyledons (monocots) was carried out more than 100 years ago by (Hanstein, 1870), who followed the formation of quadrants, octants and even the establishment of the three germinal layers and hypophysis. Such kind of histological analysis gives a general description of fertilization and embryo pattern formation. Embryo development begins with fertilisation of the egg and zygote formation. In plants there is a double fertilisation process in which two sperm cells from the pollen fuse with two nuclei in the embryo sac. One sperm cell nucleus fuses with the egg cell nucleus and becomes a zygote. The other sperm cell nucleus fuses with two polar nuclei in the central cell and gives rise to the triploid endosperm, which functions to nourish the developing zygote.

1.4.1. Embryogenesis in Dicotyledonous Plants

Arabidopsis is a useful model for plant embryogenesis not only because of its small genome size and short life cycle but largely because it has a highly invariant pattern of cell division, allowing detailed analysis of pattern formation. The development of polarity is thought to be one of the earliest and most fundamental steps in the establishment of the body plan. The embryo is determined early as the zygote elongates along the presumptive embryo axis before the first division. The first two divisions are transverse producing a two then four cell filament. The first division is asymmetric and establishes the apical-basal (A-B) polarity of the embryo by giving rise to a small apical cell and a larger basal cell (Fig.1A) that acquire different cell fates. Differential gene expression can also be observed in the two daughter cells as with the *Arabidopsis thaliana MERISTEM LAYER 1* (*AtML1*) homeobox gene that is

expressed in the apical but not in the basal cell (Lu *et al.*, 1996). The subsequent longitudinal and transverse divisions give rise to the octant stage embryo with the apical cell forming the pro-embryo and the basal cell forming the suspensor (Fig.1B). Although no organs form at this stage there are clear signs of partitioning of the embryo into 3 main regions (Fig.1C). The upper tier of the pro-embryo gives rise to the apical region of the embryo from which the shoot meristem and part of the cotyledons originate and the lower tier generates the central region of the embryo from which the remainder of the cotyledons, the hypocotyl, the root and upper tier of the root meristem originate. The basal region produces the remainder of the root meristem, the quiescent centre and the lower tier of stem cells (Fig 1.C-E). Although clearly defined the early embryo regions do not bear any clonal relationship with the A-B patterning of the seedling, highlighting the importance of cell-cell communication in patterning.

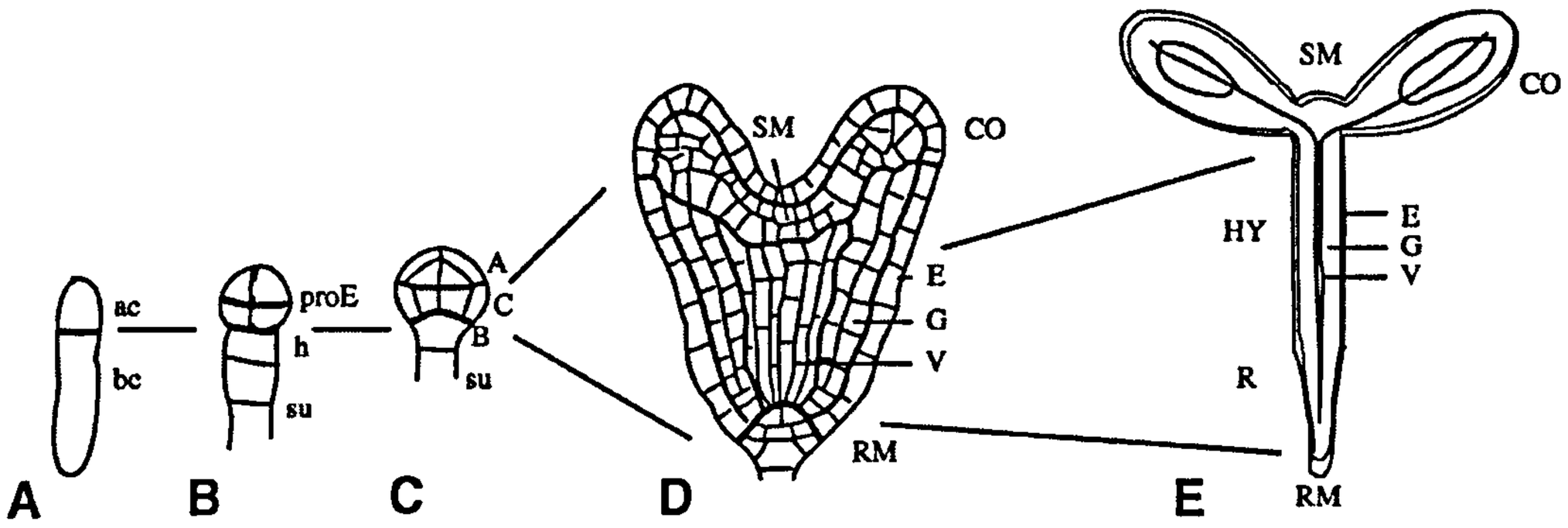


Figure. 1 Development of the apical–basal pattern during *Arabidopsis* embryogenesis. (Taken from (Jürgens, 2001) **(A)** One-cell stage. The zygote has divided asymmetrically into an apical (ac) and a basal (bc) daughter cell. **(B)** Octant stage. The proembryo (proE) derived from the apical cell consists of two tiers each of four cells. The basal cell has produced a file of cells, including the hypophysis (h) and the suspensor (su). **(C)** Dermatogen stage. Three embryo regions are indicated: A, apical; C, central; B, basal. **(D)** Heart stage. The basic body organization is in place. SM, shoot meristem; CO, cotyledon primordia; RM, root meristem; E, epidermis (L1); G, ground tissue (L2); V, vascular primordium (L3). **(E)** Seedling. HY, hypocotyl; R, root. Lines indicate the origin of seedling structures from early embryo regions.

1.4.2. Embryogenesis in Monocotyledonous Plants

Dicots and monocots represent two prototypes of plant body organization that are reflected in their seedlings. The dicot seedling is regarded as the primitive type of flowering-plant seedling (Jürgens *et al.*, 1994). Two cotyledons are symmetrically positioned across the body axis, and the meristems of the shoot and the root occupy opposite ends of the axis (Fig. 1E). The monocot seedling differs in two respects: Firstly, there is only one cotyledon on the axis in which the shoot apex situates

laterally; secondly, the embryonic root does not bear a functional root meristem, so the roots are derived from the lower portion of the hypocotyl (Jürgens *et al.,* 1994). The grass seedling represents a highly specialized variant of the monocot seedling and the various parts of the grass seedling have therefore been given special names, such as scutellum and coleoptile, as shown in Fig. 2A of the rice embryo. In contrast to the *Arabidopsis* embryo there is no set pattern of division after the first asymmetric division. However, apical cell derivatives divide actively while basal cell derivatives divide infrequently (Jürgens *et al.,* 1994). Like dicots, the protoderm divides anticlinally and represents the first evidence of histodifferentiation (Fig. 2D-E). At transition stage, a small wedge of cells becomes densely cytoplasmic and begins rapidly dividing. These cells continue dividing and form a protuberance that becomes the shoot apical meristem (SAM) (black arrow Fig. 2E). Internal to and basal to the developing apical meristem, another group of cytoplasmically dense, actively dividing cells becomes the root apical meristem (RAM) (white arrowhead Fig. 2E). Subsequently, procambial strands differentiate to connect the shoot and root primordia and this defines the plant axis (Fig. 2F-G). A collar bulges up around the apical meristem and becomes the coleoptile which ensheathes the shoot, protecting it as it pushes through the soil during germination. As the embryo continues to grow, the SAM initiates several leaf primordia (Fig. 2G). This is in contrast to dicots which form no leaf primordia prior to germination.

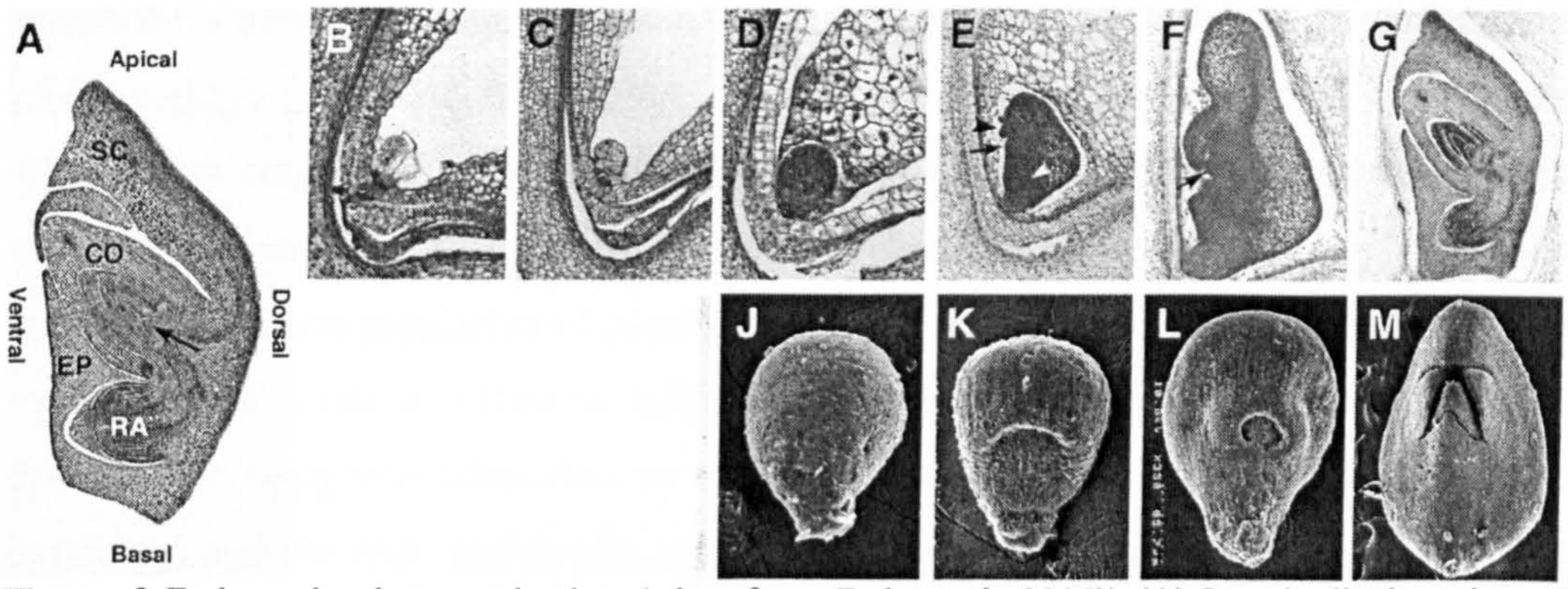


Figure. 2 Embryo development in rice. (taken from (Itoh *et al.,* 2005)) (A) Longitudinal section of mature embryo. Arrow indicates the SAM. SC, scutellum; CO, coleoptile; EP, epiblast; RA, radicle. (B) 1 DAP embryo. (C) 2 DAP embryo. (D) 3 DAP embryo. (E) 4 DAP embryo differentiating coleoptile primordium (black arrowhead), SAM (arrow) and radicle primordium (white arrowhead). (F) 5 DAP embryo differentiating the first leaf primordium (arrow). (G) Morphologically completed 10 DAP embryo. (J–M) SEM images of 3, 4, 5 and 10 DAP embryos, respectively.

Although dicot and monocot seedlings appear to represent different morphological types, their basic body organization is very similar. Two patterns are superimposed, one along the apical-basal axis of polarity and another, radial pattern perpendicular to the axis. The apical-basal pattern consists of the following elements: epicotyl (including the SAM), cotyledon(s), hypocotyl, embryonic root (radicle), RAM, and root cap. The radial pattern, which is most clearly displayed at the level of the hypocotyl, comprises the primary tissue types such as the epidermis (L1), the underlying ground tissue (L2), and the centrally located vascular tissue (L3). In an attempt to isolate early patterning mutants Jürgens and his co-workers, using *Arabidopsis*, analysed seedlings with alterations in apical-basal or radial patterning and traced the phenotypes back to changes in division patterns in the embryo (Mayer *et al.*, 1991).

1.4.3. Apical-Basal Patterning in *Arabidopsis*

Four classes of apical-basal deletion phenotypes have been isolated in Jürgens' lab: apical, basal, central and terminal, and shoot apex. The apical deletion mutant *gurke* (*gk*) corresponds to the deletion of shoot apex and cotyledons. Cotyledons are also derived from the central domain and therefore *gk* may control formation of the central and apical domains (Torres-Ruiz *et al.*, 1996). The *GK* gene, has been recently cloned and found to encode ACC1, an acetyl-CoA carboxylase which catalyzes malonyl-CoA synthesis (Baud *et al.*, 2004). Furthermore, it was found that metabolites derived from malonyl-CoA are required for partitioning of the apical part of the embryo (Kajiwara *et al.*, 2004). *Fackel* (*fk*) mutants represent central deletions which have cotyledons attached directly to the root. The *FK* gene has been cloned and found to encode a sterol C-14 reductase (Schrick *et al.*, 2000). Sterols have been shown to be active regulators of plant development and gene expression and affect the expression of genes involved in cell expansion and cell division (He *et al.*, 2003). *Monopteros* (*mp*) was identified as a basal deletion mutant, which lacks both the hypocotyl and the root, and is almost the complete converse of the *gk* seedling. The mutant embryo shows abnormalities as early as the octant stage with all cells in the proembryo behaving as upper tier cells (Berleth and Jürgens, 1993). *MP* has been cloned and shown to encode ARF5, a transcription factor of the auxin response factor (ARF) family that binds to auxin response-elements in the promoters of auxin-responsive target genes (Hardtke and Berleth, 1998). Interestingly, a second gene

BODENLOS (*BDL*) was identified in which mutants exhibited the same phenotype as *mp*. The *BDL* gene was shown to encode IAA12, a member of the IAA family of proteins that form complexes with ARF transcription factors, from which the latter are released in response to auxin. *Gnom* (*gn*) represents the terminal deletion class of mutants, with the shoot apex, cotyledons and root reduced or eliminated. Mutant *gn* embryos are almost round, with histological analysis indicating that the *gnom* allele may fail to promote the first asymmetric cell division which is necessary for establishing A-B polarity of the embryo (Mayer *et al.*, 1993). Using T-DNA tagging, the *GN* gene has been cloned from *Arabidopsis* (Shevell *et al.*, 1994). Unexpectedly, the predicted amino acid sequence of *GN* is similar to a yeast protein (Sec7p) that functions as a guanine nucleotide exchange factor involved in vesicle transport and is expressed in all tissues throughout development. Embryos resembling mutations in the *GN* gene have been observed using auxin antagonists and polar auxin transport inhibitors. Furthermore, the induced phenotypes at later stages of development resembled mutations in the auxin efflux carrier *PIN-FORMED 1* (*PIN1*) (Hadfi *et al.*, 1998). During embryogenesis the localisation of *PIN1* becomes polarised from mid-globular stage onwards which becomes disrupted in *gn* embryos (Steinmann *et al.*, 1999). Thus *GN* appears to play a specific role in auxin transport during embryogenesis. Finally, mutant deletions of the shoot apex are represented by a mutant called *shoot meristemless* (*stm*) (Barton and Poethig, 1993). Mutation of the *STM* gene completely blocks the initiation of the SAM, and is therefore required to maintain the proliferation of cells in the SAM. *STM* encodes a homeobox transcription factor of the *Knotted1*-like homeobox (*KNOX*) gene family (Long *et al.*, 1996) which accumulates in both the central and peripheral zone of the meristem but is repressed in organ primordia, in accordance with a role in maintaining cells in an undifferentiated state.

1.4.4. Radial Patterning in *Arabidopsis*

The radial pattern of a mature embryo consists of 3 primary tissues: epidermis, ground tissue and vascular bundles (L1-L3), which are originated at the globular stage and are maintained during postembryonic development (Laux and Jürgens, 1994). However, few mutants have been described with defects in the radial axis. Two such mutants, *keule*; and *knolle* (*kn*), identified in *Arabidopsis*, are defective in the establishment of epidermis (L1). Globular *keule* embryos show enlarged cells in the outer layer that

normally will give rise to the epidermis, while the inner cells appear normal. By contrast, in *knolle* the enlarged cells are not restricted to the outer cell layer but also take place in the inner tissues (Jürgens *et al.*, 1994). *KN* encodes a syntaxin, a vesicular trafficking protein that functions specifically in cytokinesis (Lukowitz *et al.*, 1996). *KEU* encodes the yeast Sec1 homologue, a key regulator of vesicle trafficking, that binds the cytokinesis-specific syntaxin *KN* (Assaad *et al.*, 2001). Two genes required for radial pattern formation of the ground tissue in the root are *SHORTROOT* (*SHR*) and *SCARECROW* (*SCR*). *SCR* is essential for the asymmetric division that normally generates cortex and endodermis (Di Laurenzio *et al.*, 1996) whereas *SHR* plays a role in both the cell division and the specification of the endodermis (Benfey *et al.*, 1993). *SHR* and *SCR* encode transcription factors of the GRAS family that play important roles in the physiology and development of higher plants (Pysh *et al.*, 1999). *SCR* is expressed in the quiescent centre (QC), cortex and endodermis whereas *SHR* is unexpectedly expressed in the vascular tissue but then appears to move into the endodermal cells where its action is required (Nakajima *et al.*, 2001). Recently, it has been shown that SCR functions to in the QC to maintain the stem cell identity of root meristem initials (Sabatini *et al.*, 2003). The scarcity of mutants defective in radial pattern could be due to either that such kinds of mutants are difficult to be recognized, or that deletion of a primary tissue would cause embryo lethality.

1.4.5. Patterning in Monocots

Although monocot embryos display no regular cell division patterns, develop only one cotyledon and consist of many more cells at maturity, there may be differences in detail rather than in overall patterning processes. It seems that it is possible, in some cases, to extrapolate the molecular mechanisms of embryogenesis from *Arabidopsis* to more distantly related plant species, such as the monocots rice and maize. For example, the epidermal cell layer in maize is established much later than in *Arabidopsis* embryogenesis. Nonetheless, a putative homologue of *AtML1*, the maize *OCL1* gene, is not only related by sequence, but also expressed in a similar manner to *AtML1* (Ingram *et al.*, 1999). Similarly, the maize *KNOTTED* gene, a putative homologue of *STM*, is expressed in a complementary domain to *rough sheath2*, the homologue of the *Arabidopsis* leaf-initiation gene *AS1*, which is negatively regulated by the shoot meristem-promoting *STM* protein (Tsiantis *et al.*, 1999). Although these

only offer a few examples, additional similarities are likely to be discovered as more genes are being analysed in maize and rice.

1.5. Somatic Embryogenesis as a Model for Zygotic Embryogenesis

A major limitation in using zygotic embryos for the study of embryogenesis is that they are small and inaccessible in maternal tissues, particularly in their early developmental stages (Zimmerman, 1993). Somatic embryos (SEs), on the other hand, are readily accessible and can be maintained at a relatively uniform level of development. Furthermore, SEs exhibit very similar properties to zygotic embryos in their early development, both undergoing similar developmental stages, i.e. globular, heart and torpedo stages. Studies suggest that the extra-embryonic tissues of maternal or non-maternal origin surrounding a zygotic embryo may not be vital for the embryo pattern formation but rather may mainly function by nourishing the developing embryo. Somatic embryogenesis (SE), therefore, provides an alternative system for studying the early events in embryogenesis at the genetic and molecular levels. Although studies have identified common features of somatic and zygotic embryogenesis, there exist a number of obvious differences between these two processes. Most notably is that unlike the zygote, somatic cells require a de-differentiated stage prior to embryo development. The mechanisms involved in this acquisition of embryogenic competence offer an additional dimension to the study of SE and are of particular interest in this study.

1.6. Molecular Understanding of Somatic Embryogenesis

The initiation of somatic embryogenesis (SE) requires substantial reprogramming of gene expression in order to abolish the differentiated cellular status and to regain the totipotency of the cells. Such reprogramming could occur at either the transcriptional or post-transcriptional level. For example, it has been observed that the transition from unorganised cell proliferation to SE requires active RNA synthesis (Fujimura and Komamine, 1980). The identification of the genes involved in such reprogramming has been considered as the key to understanding SE.

1.6.1. Genes Identified in Somatic Embryogenesis

1.6.1.1. *LEAFY COTYLEDON* (*LEC*)

Surprisingly, attempts to identify genes that control somatic embryogenesis (SE) have been largely disappointing with the majority of the genes identified so far appearing to control basic developmental processes that function throughout the life cycle of the plant (Kaplan and Cooke, 1997). However, a few of such genes appear to play a more specific role in embryo development. For example, the *LEAFY COTYLEDON 1* (*LEC1*) and *LEAFY COTYLEDON 2* (*LEC2*) genes, which were originally identified in *Arabidopsis* through the analysis of mutants with defects in embryonic maturation and cotyledon identity (West *et al.,* 1994), have recently been shown to encode seed-specifically expressed transcription factors. *LEC1* contains the HAP3 subunit of a CCAAT box-binding factor (Lotan *et al.,* 1998), a transcription factor domain unique to plant transcription factors and *LEC2* contains a B3 DNA binding domain with close similarity to the B3 domain transcription factors *VIVIPAROUS 1,* and *FUSCA3* (Stone *et al.,* 2001). Furthermore, the ectopic expression of *LEC1* in vegetative cells was shown to induce the expression of embryo-specific genes and the formation of embryo-like structures and the ectopic expression of *LEC2* could induce the formation of somatic embryos or other organ-like structures, suggesting that *LEC1* and *LEC2* genes are important regulators of embryo development by controlling the expression of genes required for embryogenesis and cellular differentiation. Another gene known to be associated with the induction of SE through the LEC pathway is the *Arabidopsis PICKLE* (*PKL*), which was identified through the analyses of a recessive mutant defective in repressing embryonic differentiation characteristics after germination. The *PKL* gene encodes a CHD3 chromatin remodelling factor which is necessary for the repression of *LEC1*. In the *pkl* mutant seeds, *LEC1* is de-repressed and some vegetative tissues have the capacity to produce somatic embryos spontaneously (Ogas *et al.,* 1999). It is therefore thought that the expression of *PKL* is required for preventing the early embryo development and for promoting the action of LEC during seed germination.

1.6.1.2. *SOMATIC EMBRYOGENESIS RECEPTOR LIKE KINASE – (SERK)*

Post-translational modification of proteins may also be involved in the regulation of somatic embryogenesis (SE). For example, (Murray and Key, 1978) showed that the application of 2,4-D resulted in a substantial increase in the number of phosphorylated proteins within 24 hours. In alfalfa suspension cells, the increase in cell division and the frequencies of SE were associated with an increase in the activity of a calcium dependent protein kinase (CDPK). This CDPK not only has the ability to auto-phosphorylate itself, but can also phosphorylate histone H3 (Bogre *et al.*, 1988).

One of the earliest known markers for SE is the putative receptor-like protein kinase named Somatic Embryogenesis Receptor-like Kinase (SERK). Analysis using a SERK promoter-luciferase reporter indicated that this gene was expressed in most of the cell clusters giving rise to somatic embryos (SEs) and in SEs up to the globular stage (Schmidt *et al.*, 1997). The *Arabidopsis* SERK homolog (*At*SERK1) was also shown to be expressed during embryogenic cell formation in culture and during early zygotic embryogenesis up to the heart stage. Although no observable phenotype was produced by the ectopic over-expression of *At*SERK1, seedlings exhibited a 3-4 fold increase in the efficiency for SE (Hecht *et al.*, 2001). Interestingly, the *At*SERK 1 protein was found to be located in the plasma membrane and to interact, in-vitro, with another membrane localised protein known as the Kinase-associated protein phosphatase (KAPP) (Shah *et al.*, 2002). Furthermore, It was proposed that the physical interaction between KAPP and *At*SERK1 is required for the internalisation of receptors and the initiation and maintenance of embryogenically competent cells (Shah *et al.*, 2002). The association of this important gene in SE with a phosphatase illustrates their ability to regulate key cellular processes such as SE.

1.6.1.3. *BABY BOOM (BBM)*

Another gene associated with somatic embryogenesis (SE) is the *BABY BOOM (BBM)* gene, which was identified using subtractive hybridisation of the up-regulated genes during microspore embryogenesis of pollen grains from *Brassica napus* (Boutilier *et al.*, 2002). *BBM* encodes an AP2/ERF domain transcription factor and is preferentially expressed in developing embryos and seeds. Furthermore, the ectopic expression of BBM, under the cauliflower mosaic virus (CaMV) 35S promoter, in

both *Brassica* and *Arabidopsis* led to the spontaneous formation of somatic embryos on seedlings in the absence of exogenously-added cytokinin and auxin, suggesting a role of this gene in promoting cell proliferation and morphogenesis during embryogenesis (Boutilier *et al.*, 2002). Although the mechanism for BBM's function is unknown, the regulation of a large number of the AP2/ERF domain genes by plant hormones at the transcriptional level and/or their functions in hormone signalling pathways suggest that the function of BBM may be also linked to hormones and hormonal signalling.

1.6.2. The Plant Cell Wall and Cell wall proteins

In plants, the cell wall is an important structure that determines cell shape, provides tensile strength and acts as a barrier against pathogens. At the same time the cell wall also shows remarkable plasticity and elasticity as the cell grows. The primary wall is composed of a network of cellulose microfibrils embedded in a matrix of hemi-celluloses and pectins (Varner and Lin, 1989). Cell wall expansion occurs by the slippage or rearrangement of matrix polymers that hold the microfibrils in place and was thought to occur primarily by hydrolysis of matrix polysaccharides. However it was discovered that if a cell wall protein, called expansin, was added to heat denatured cell walls it restored their ability to extend (McQueen-Mason *et al.*, 1992). The expression of expansin has subsequently been demonstrated in a number of growth related processes including auxin induced root formation (Hutchison *et al.*, 1999) and the emergence of primordia in the meristem where the expression of expansin precedes primordia outgrowth (Fleming *et al.*, 1997) and is therefore important for early developmental processes such as SE. Expansin appears to affect the plant cell wall by the rapid induction of wall extension and through stimulation of stress relaxation but does not involve hydrolytic activity (McQueen-Mason *et al.*, 1992) although it has been shown to enhance the hydrolysis of cellulose by cellulases. Expansins have been proposed to function by weakening non-covalent binding between wall polysaccharides, allowing turgor driven polymer-creep (McQueen-Mason and Cosgrove, 1994).

Proteins secreted into the culture medium such as endochitinase and arabinogalactan proteins (AGPs) have been shown to be required for somatic embryogenesis (SE) (Van Hengel *et al.*, 2001). Using a temperature sensitive mutant *ts11* in carrot

(*Daucus carota*) it was demonstrated that the addition of a carrot class IV endochitinase (EP3) allowed the arrested somatic embryos to develop into plantlets (De Jong *et al.*, 1992). Endochitinases function by hydrolysing beta linkages between adjacent N-acetyl-D-glucosaminyl (GlcNAc) residues in chitin polymers and partially deacylated chitin (chitosan). However, chitin polymers or chitisan are not present in plant cell walls, suggesting that there may be other factors acting as the targets for endochitinase activity. AGPs are proteoglycans that can occur attached to cell membranes or in cell walls and their role in plant development was initially proposed based on the temporal and spatial localisation pattern as visualised by the use of monoclonal antibodies (Knox *et al.*, 1989; Knox *et al.*, 1991). It has also been shown that the addition of AGPs from mature carrot seeds to a non-embryogenic line induced the initiation of embryogenic clusters (Kreuger and van Holst, 1995). It has been recently demonstrated that secreted AGPs from embryogenic cell cultures, but not from non-embryogenic cell cultures, contain glucosamine and GlcNAc which are sensitive to endochitinase cleavage (Van Hengel *et al.*, 2001). Furthermore, these authors developed assays that demonstrated the AGPs derived from embryogenic cultures and immature seeds could restore and even increase SE and that a treatment with an AGP chitinase further enhances their embryo-promoting activity. The same group also showed that the composition of carrot immature seed AGPs changes during seed development and that the somatic embryo-promoting effect of these AGPs is affected by this compositional change (Van Hengel *et al.*, 2002). The (developmentally-regulated) presence of an endochitinase cleavage site in AGPs during the same developmental window as EP3 suggests that the hydrolytic cleavage of AGPs by EP3 somehow affects embryo development either by changing the cells identity or creating new signalling molecules (Van Hengel *et al.*, 2001).

1.6.3. DNA Methylation and RNA Silencing

DNA methylation may also play an important part in the transcriptional regulation of SE. It has been reported that Auxin (2,4-D) treatment, a well-known trigger for somatic embryogenesis (SE), inhibits methylation in wheat seedlings (Kirnos *et al.*, 1986). (Morrish and Vasil, 1989) also found a significant hyper-methylation following the application of auxin and rapid de-methylation on the removal of auxin in Napier grass callus. The role of methylation and heterochromatin silencing has received much attention of late due to the discovery that RNA silencing has been

shown to function in heterochromatin silencing through the targeting of repeat sequence transcripts present throughout the genome.

In plants there are at least three RNA silencing pathways that have been revealed by genetic and molecular analysis (Baulcombe, 2004). All of these pathways involve the cleavage of double stranded RNA into short 21-26 nucleotide RNA duplexes, known as short interfering RNAs (siRNAs) and micro RNAs (miRNAs), by an enzyme called Dicer, a member of the RNAse III family of double stranded RNA-specific endonucleases (Ketting *et al.*, 2001). These siRNAs or miRNAs are incorporated into an RNA-induced silencing complex (RISC) that serves to target complementary transcripts for specific cleavage and degradation. The first of these pathways is cytoplasmic siRNA silencing (Hamilton and Baulcombe, 1999) which is an important defence mechanism against invading viruses, and responsible for the phenomenon of transgene silencing (Waterhouse *et al.*, 2001). The second pathway involves the silencing of endogenous mRNAs by miRNAs which act to negatively regulate gene expression through the cleavage or arrest of protein translation of complementary RNA sequences via the RNAi machinery (Bartel, 2004). The third pathway is involved in DNA methylation and suppression of transcription, discovered by the fact that transgene viral RNAs guide DNA methylation to specific sequences and by observations that RNA silencing is linked to histone modification and histone chromatin formation (Wessenger *et al.*, 1994). This unexpected role of RNAi in heterochromatin gene silencing is thought to be involved with protecting the genome against instability and damage caused by transposons. Furthermore 90-95% of endogenous siRNAs in *Arabidopsis thaliana* correspond to transposons or repeats whose histones and DNA are heavily methylated whereas the coding regions are practically free of this modification in plants (Lippman *et al.*, 2004).

miRNAs in plants mainly function as siRNAs that guide specific base pairing and cleavage and were originally identified from the *Arabidopsis* siRNA population that contained similarity to the *C. elegans*, *LET-7* and *LIN-4* miRNAs (Bartel, 2004). miRNAs are derived from specific genes that encode inverted repeat precursor RNA that contains double stranded RNA regions which target complementary single stranded mRNA. In *Arabidopsis* there are four putative dicer like proteins (DCL1-DCL4), three of which are thought to be involved in processing dsRNA from different

sources (Xie *et al.*, 2004). DCL1 processes miRNA precursors and is required for miRNA biogenesis. DCL2 has been implicated in the production of siRNAs from viruses. DCL3 is involved in the production of retro-elements and transposons RNA and is required for chromatin silencing. The function of DCL4 is currently unknown. The siRNA and miRNA duplex-containing ribonucleoprotein particles are subsequently unwound and assembled into effector complexes, including the RNAi induced silencing (RISC), which mediates mRNA target degradation (Hammond *et al.*, 2000), the miRNA containing effector complex (miRNP), which guides translational repression of target mRNAs (Mourelatos *et al.*, 2002), and the RNA-induced transcriptional silencing complex (RITS), that guides the condensation of heterochromatin (Verdel *et al.*, 2004). The ARGONAUT proteins (AGO) have been implicated in all three pathways (Baulcombe, 2004), appearing to play an essential role in the effector complexes either by cleavage or translational repression. Another essential component in the silencing pathway are the RNA dependent RNA polymerases (RDRs) which are required for cytoplasmic and chromatin silencing pathways in plants (Xie *et al.*, 2004). RDRs are capable of generating many dsRNAs from either a single aberrant mRNA species or from a primary siRNA molecule that primes the synthesis of normal RNA to be targeted via the RNAi pathway, thereby providing the amplification step necessary to ensure the silencing of actively-replicating viral RNAs or introduced transgene. Similarly, these amplification steps would ensure that only a few molecules of transposon RNA would be sufficient to suppress all copies of a transposable element.

1.6.3.1. miRNAs and their Role in Development

Non messenger RNAs are diverse molecules with structural, enzymatic and regulatory functions (Ambros, 2004). Among those with regulatory function are miRNAs, which have been implicated in the regulation of the expression of several genes. In plants, miRNAs appear to function, in the main, through cleavage and degradation of mRNA targets (Rhoades *et al.*, 2002), whereas animal miRNAs act in a less sequence-specific manner and inhibit protein translation of target transcripts (Olsen and Ambros, 1999). Most miRNAs are solitary, self-regulated molecules, however some genes encoding miRNAs have been found to be arranged in clusters and co-regulated (Houbaviy *et al.*, 2003). Recent efforts have been made to identify miRNA on a genomic scale through the use of cDNA sequencing, bioinformatics and functional

genomic approaches, such as miRNA gene expression microarrays and their functions in gene regulation and development (Zheng *et al.*, 2004). There are now extensive lists of plant miRNAs, several of which have been experimentally demonstrated to regulate gene expression. These include expression of a myb transcription factor involved in plant hormone response (Palatnik *et al.*, 2003) and an Apetela 2 (AP2) transcription factor involved in floral development (Chen, 2004). Many other putative miRNAs have been identified through bioinformatics approaches and correspond to transcription factors and other developmental regulators, many of which appear conserved between cereals, mosses and flowering plants (Floyd and Bowman, 2004). Some miRNAs appear to play a role in the response to environmental stresses, such as cold or drought, or to hormonal (auxin) stress. In fact gibberellic acid (GA) stimulus increases the production of both myb 33, which promotes the initiation of flowering, and a miRNA, miR159, which counteracts myb 33 RNA and therefore dampens the flowering response (Baulcombe, 2004). Another interesting aspect of RNA silencing involves secondary miRNA targets which are tolerant of limited sequence mismatches (Bartel, 2004). Mismatch results in translational repression of the target as occurs for AP2 rather than the specific cleavage and degradation of the mRNA (Chen, 2004). Although much has been discovered about the roles of RNA silencing, from antiviral defence to the regulation of gene expression and chromatin structure, there is still much to be learnt about the full extent of this functional diversity. For example, in mammals, it was found that six miRNA genes were organised in tandem and were expressed specifically in embryonic stem cells (ESCs), suggesting a role of these miRNAs in the maintenance of stem cell potency (Houbaviy *et al.*, 2003). The finding that miRNAs are involved in the maintenance of stem cell potency opens interesting questions into the possible role of miRNAs in SE. It would be interesting to investigate the extent to which miRNAs are involved in the initiation and development of SEs.

1.7. The Role of Auxin in Somatic Embryogenesis

The addition and removal of auxin are key factors in the initiation and development of somatic embryos (SEs). Auxin is required for the initiation of somatic embryogenesis (SE) in almost all tested plant species. However, significant variations in the responsiveness to auxin are observed among different plant species, tissue sources or cells. For example, cells of alfalfa require only a short pulse of auxin before being

able to initiate SE in hormone free medium (Zimmerman, 1993), whereas species such as wheat require several days exposure in order to develop mature embryos. Early work by (Halperin and Wetherell, 1964) using carrot suspension cells also showed that auxin requirement varied at different developmental stages of the SEs, observing that the continued presence of auxin prevented embryos to develop beyond the globular stage. This suggests that the transition from the globular to the heart stage requires new gene products expressed only in the absence of auxin. Furthermore, treatment of SEs with auxin transport inhibitors at the globular stage blocked the progression to the next stage (Schiavone and Cook, 1987). However, mature embryos were unsusceptible to the auxin transport inhibitors (Liu *et al.*, 1997). A difference in auxin response is also observed between embryonic and non-embryonic cells. For example, 2,4-D can induce a marked increase of protein phosphorylation in embryogenic cells, but not in non-embryogenic cells. The differences in auxin requirement among different plant species, tissues, cells or developmental stages possibly reflect differences in auxin sensitivity.

1.7.1. The Molecular Mechanisms of Auxin Action in Plants

The mechanisms by which auxin regulates plant growth and development in general is of particular relevance to understanding the molecular mechanisms involved in somatic embryogenesis (SE). The major active auxin found in plants is indole-3-acetic acid (IAA) which plays a key role in a wide variety of growth and developmental processes. The aerial parts of the plant, in particular developing leaves, are believed to function as an important source of IAA for the rest of the plant but there are also other parts of the plant such as cotyledons, expanding leaves and root tissues that have the capacity to synthesize IAA de novo (Ljung *et al.*, 2001).

1.7.1.1. Auxin Signal Transduction

Advances in our understanding of the role of auxin have been brought about by recent biochemical and genetic data. Upon application of auxin there is a rapid expression of a large number of plant genes, some responding within minutes that do not require *de novo* protein synthesis for their induction (Abel *et al.*, 1994). These auxin induced genes consist of several multi-gene families expressed in a broad range of plant tissues, such as auxin response factors (ARFs) and auxin/indole-3-acetic acid (AUX/IAA) proteins. Analysis of the GH3 gene family identified auxin responsive

elements (AuxRE) within the promoter sequence that is required for auxin induction (Liu *et al.*, 1994). Ulmasov et al. (1997) used synthetic AuxREs to screen an *Arabidopsis* cDNA library and identified an auxin response factor (ARF1) that specifically binds with the TGTCTC element and contains an amino terminal DNA binding domain also present in the maize transactivator *VIVIPAROUS1* (*VP1*) (Ulmasov *et al.*, 1997). In addition, the carboxy terminus of the ARF1 protein contains a domain related to the AUX/IAA protein family suggesting that the two proteins may interact through these domains (Ulmasov *et al.*, 1997). Furthermore, it was shown that the AUX/IAA proteins repressed the expression of reporter genes controlled by the synthetic auxin response elements (Ulmasov *et al.*, 1997), probably by preventing ARF binding. Genetic studies have shown that ARF3 is identical to a gene called *ETTIN* that has been demonstrated to be involved in floral patterning and possibly mediates auxin signalling during patterning of the floral meristem (Sessions *et al.*, 1997). IAA24 has also been shown to encode the *MONOPTEROS* (*MP*) protein with mutants showing defects in the initiation of the embryo axis (Hardtke and Berleth, 1998).

The cloning of several genes over the past few years has also elucidated that auxin promotes the degradation of the transcriptional repressor AUX/IAA proteins by stimulating interaction with an SCF-type (SKP1-CULLIN/Cdc53-F-box) E3 ubiquitin-ligase protein complex (Gray *et al.*, 1999). Ubiquitin-ligases tag proteins with a polyubiquitin polypeptide thereby marking them for degradation by the 26S proteasome. More recently the f-box protein TIR1 was shown to bind auxin directly, implicating it as an auxin receptor, and promote the interaction between AUX/IAA and the SCF complex (Dharmasiri *et al.*, 2005). This degradation of the AUX/IAA inhibitors causes the interacting ARF transcription factors to be released to regulate target genes (Weijers and Jurgens, 2004). Because there are many ARFs and AUX/IAA proteins identified, it is believed that the cell-specific combinations of the two proteins may determine the specific developmental response to auxin and help to explain the wide variety of responses generated by auxin. It is therefore possible that specific ARF-AUX/IAA protein pairings may also be involved in some way in the initiation of SE in response to the application of exogenous auxin (2,4-D) in tissue culture. Here the degradation of AUX/IAA induced by exogenous auxin would release ARFs that target genes responsible for the initiation of SE.

1.7.1.2. Auxin Distribution and Transport

The transduction of gene expression by auxin as described above is essential for interpreting the pattern of auxin distribution across the plant. However, the mechanisms responsible for the distribution of auxin are also essential for establishing that pattern in the first place. Plants use polar auxin transport as a means of conveying IAA from the site of synthesis to its target tissue and use a specialised delivery system that is mediated by influx and efflux carriers. Mutational studies using *Arabidopsis* have led to the characterisation of the influx (AUX/LAX proteins) and efflux (PIN proteins) components which belong to distinct gene families. The *AtPIN1* gene was one of the first efflux carriers identified and encodes a protein with a number of transmembrane segments with homology to prokaryotic and eukaryotic transporters (Gälweiler *et al.*, 1998). *Atpin1* mutants have decreased rate of polar auxin transport within inflorescence tissue and the pin-formed phenotype which can be phenocopied in wild type plants grown in presence of the auxin efflux carrier inhibitor, naphthylphthalamic acid (Okada *et al.*, 1991). The asymmetric localisation of the AtPIN family of proteins within the plasma membrane of plant cells was demonstrated by immunlocalisation studies. AtPIN 1 was shown to be localised to the basal membrane of inflorescence cells where they facilitate polar auxin transport, providing a mechanism for the characteristic polar transport of auxin (Gälweiler *et al.*, 1998; Müller *et al.*, 1998). Recent work by Jürgens and his co workers has shown that PIN1 not only localises to the plasma membrane but also to recycling vesicles and endosomes and that auxin transport inhibitors act by inhibiting the recycling of PIN1 (Geldner *et al.*, 2001; Geldner *et al.*, 2003). This has led to the model that PIN1 loads auxin from the cytoplasm into vesicles and endosomes and secretes auxin via the plasma membrane before being retrieved back to the cytoplasm. The *Arabidopsis AUX1* gene was the first influx carrier to be identified and was found to encode a hydrophobic polypeptide, containing up to 11 transmembrane segments (Bennett *et al.*, 1996) that shares a common ancestry with a family of plant amino acid permeases (Frommer *et al.*, 1993). It was proposed that AUX1 facilitates the movement of the amino acid like molecule IAA (Bennett *et al.*, 1996). AUX1 is most closely related to three Arabidopsis sequences termed LAX1, LAX2, and LAX3 (Like AUX1) (Parry *et al.*, 2001). AUX1 shares between 73% and 82% identity at the amino acid level with other LAX sequences, indicative of a conservation of transport function. Protein localisation studies have recently revealed AUX1 appears asymmetrically localised in

root protophloem pole cells. Furthermore, double-labelling experiments have revealed that AUX1 and AtPIN1 are preferentially targeted to the upper and lower protophloem PM domains respectively (Friml *et al.*, 2002), suggesting that influx and efflux carriers act in concert to facilitate the polar transport of auxin.

1.7.1.3. Auxin Flow and Pattern Formation

The role of influx and efflux carriers in determining auxin flow and pattern formation was also demonstrated by several studies showing that auxin transport plays a key role during embryonic, root and shoot organogenic processes (Benkova *et al.*, 2003; Friml *et al.*, 2003; Reinhardt *et al.*, 2003). All three processes require efflux dependent accumulation of auxin from polarised cells to create a localised response. Furthermore, the root and shoot tissues employ related mechanisms to initiate organogenesis through the creation of auxin maxima (Benkova *et al.*, 2003; Reinhardt *et al.*, 2003) highlighting common patterning mechanisms employed during plant development. Using a green fluorescent protein (GFP) reporter for auxin responsive gene expression (DR5::GFP) it was shown that following the first asymmetric division in the zygote, the apical but not the basal cell displays an auxin response (Friml *et al.*, 2003). Furthermore, this continues up to the globular stage where its cessation coincides with the expression in the hypophysis, thereby demonstrating the role of auxin in both the specification of the apical cell and the hypophysis (Friml *et al.*, 2003). Interestingly, the subcellular localisation of PIN1 and PIN7 is indicative of their role in directing auxin flow to create the differential responses observed using DR5::GFP. That is, PIN7 localises to the apical membrane of the suspensor and PIN1 localised to the proembryo cells in a non polar fashion. At the globular stage PIN1 is recruited to the basal membrane of proembryo cells and PIN7 polarity is reversed so that it faces the basal membrane of suspensor cells. It is believed that the cycling of the efflux components is important for the rapid changes in PIN polarity that have been observed between the different cell faces in the early *Arabidopsis* embryo. These studies in *Arabidopsis* have led to the model that PIN proteins together with AUX/LAX proteins create patterns of auxin gradients that, in turn, create patterns of gene expression and morphogenesis.

1.7.1.4. Auxin Treatment and Somatic Embryogenesis

How auxin triggers SE is largely unknown. However, treatment with auxin and its subsequent withdrawal is clearly a primary trigger for the onset of SE. It is likely that

the integration of auxin polar transport along with the auxin signalling mechanisms discussed will play a significant role in the initiation and development of SEs. Considering the high concentrations of auxin required for the initiation of SE, it is also possible that auxin acts as a stress signal, which in turn switches on the SE process. Consistent with such a hypothesis is the finding that heat shock proteins (HSPs), which are induced by high temperature stress, are induced during SE (Pitto *et al.,* 1998). HSPs have also been known to play a role during cell proliferation, differentiation and embryogenesis in a variety of eukaryotic organisms and have been shown to be transiently expressed during early embryogenesis (Bond and Schlesinger, 1987). Furthermore, a number of these proteins display characteristic differences at different developmental stages during carrot SE (Pitto *et al.,* 1998). Interestingly, a 3'element of an auxin regulated gene cluster, known as small auxin up RNAs (SAURs), exhibits high homology with a motif located downstream of the soybean heat shock gene (McClure *et al.,* 1989) suggesting a link between auxin action and the heat shock response. The relationship between the initiation of totipotency in somatic cells and the activation of heat shock genes opens interesting questions about the role of stress in plant development. Such as what are the molecular mechanisms of the heat shock response? What target genes other than HSP genes are recognised by heat shock transcription factors? What is the functional role of these genes in embryogenesis and SE?

1.8. Somatic Embryogenesis and Cell Division

One of the characteristic features in the initiation of SE is a reactivation of the cell cycle in the differentiated somatic cells or tissues upon the application of exogenous auxin. These artificially induced cell divisions pave the way for the transitional switch of cells from a differentiated somatic status to an embryogenic one. It has been established that the basic cell cycle regulatory machinery is well conserved in all eukaryotic organisms. At the centre of this machinery lie the CDC2-like protein kinases and the cyclins as components of the maturation-promoting factor (MPF). Northern analysis has revealed that 2,4-D significantly increases the expression of the CDC2 gene in alfalfa (Hirt *et al.,* 1991) and rice suspension cultures (Hata, 1991). Furthermore the auxin-induced expression of the CDC2 genes in these suspension cultures was independent of the cell cycle phase, suggesting a loss of cell cycle control under *in vitro* conditions (Magyar *et al.,* 1993). (Nuti-Ronchi *et al.,* 1992)

suggested that because meiotic events existed in culture during SE and resulted in the reduction of the DNA content of cells that somatic meiosis could be part of SE and that cleavage division and the formation of haploid nuclei were essential to the restoration of cellular totipotency. However, further studies are needed to elucidate the role of somatic meiosis-like reduction and its relationship with cellular dedifferentiation. Auxin-induced cell proliferation does not necessarily lead to the formation of embryogenic cells, as it is frequently observed that non-differentiated and unorganised callus tissue is formed. Clearly, cell cycle regulation is only part of the regulatory mechanisms for SE. At present there is only limited information available about the link between cell cycle control and embryogenesis. Analysis of the effect of auxin on various sections of young wheat leaves has indicated a gradient in cell cycle activation in response to auxin treatment (Wernicke and Milkovits, 1987), and a link between the accelerated activation of the cell cycle and embryogenic initiation of somatic cells.

1.9. Somatic Embryogenesis and Specification of Stem Cells

The initiation and regeneration of plant organs from somatic embryos indicate the importance of the correct establishment of the meristem for SE. Therefore an understanding of the molecular components involved in the establishment of the shoot and root meristem is of particular relevance. Stem cells in the shoot apical meristem (SAM) provide cells for continuous formation of organs such as leaves, stems, and flowers during post embryonic growth. Stem cells give rise to 3 clonally distinct cell layers (L1-L3) and the SAM can be subdivided into layers and zones. The cells of L1 layer divide anticlinally and therefore remain in this layer and eventually differentiate into epidermis. Cells in the L2 form a sub-epidermal cell layer and gametes. The third layer (L3) gives rise to the vascular system (Figure 3*A*). Stem cell identity is specified from the organising centre (OC) which expresses the homeodomain transcription factor *WUSCHEL* (*WUS*) (Mayer *et al.*, 1998). Loss of function of *WUS* leads to differentiation of stem cells and meristem termination (Laux *et al.*, 1996) whereas ectopic *WUS* expression in vegetative organ primordia induces ectopic stem cell identity (Schoof *et al.*, 2000). This indicates that tightly controlled WUS expression is necessary for stem cell maintenance. In turn stem cells signal back via the *CLAVATA* (*CLV*) signalling pathway to restrict the *WUS* expression domain. The

clavata (*clv*) mutants accumulate too many cells in the central zone of the SAM (Clark *et al.*, 1993; , 1995) caused by expansion of the WUS expression domain into

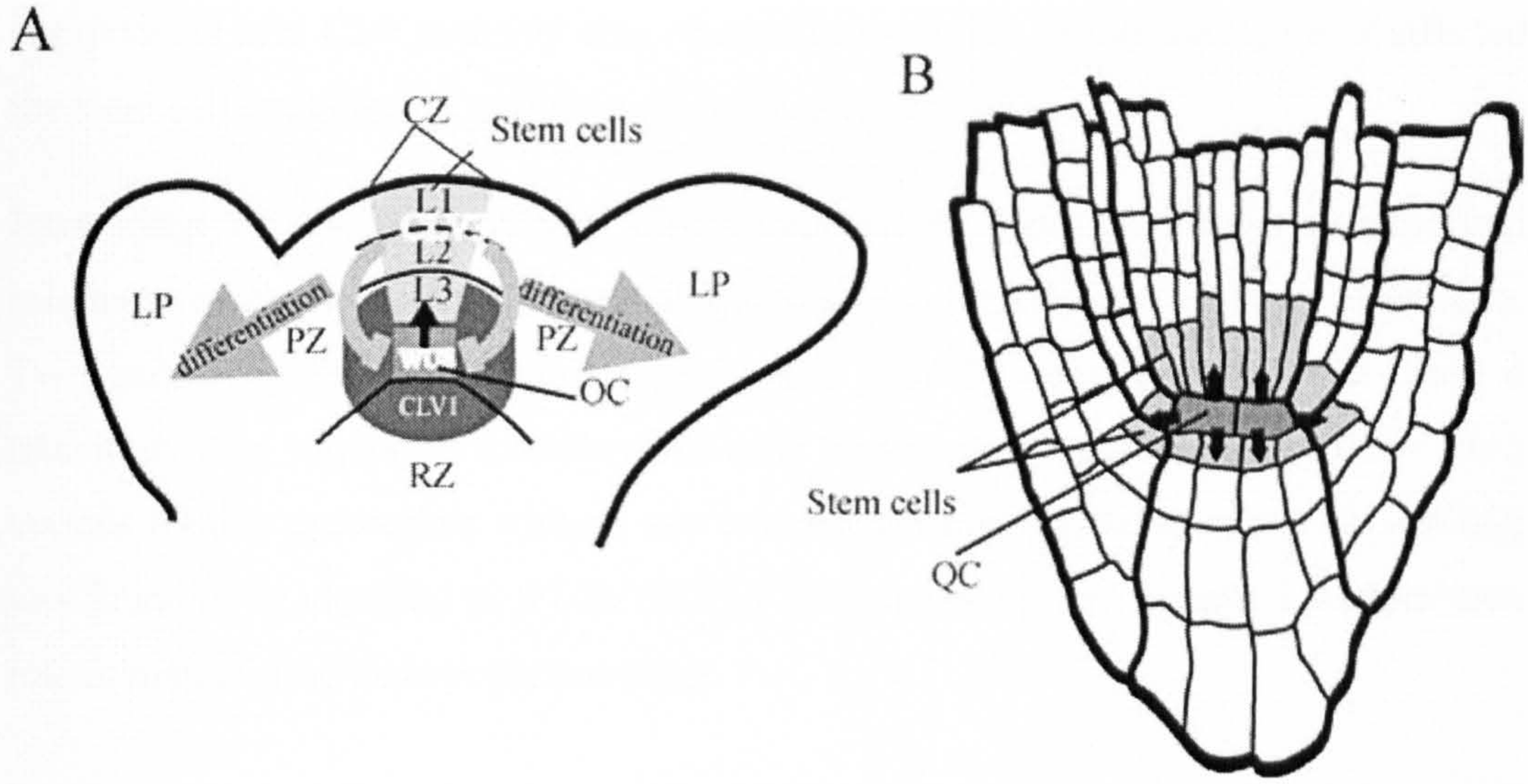


Figure 3. Organisation and maintenance of the SAM. (Taken from (Willemsen and Scheres, 2004) *A*. Schematic view of the different domains in the SAM. The central zone (CZ) contains slowly dividing cells (*in light grey*) that include the apical stem cells. Initiation of organ primordia takes place in the peripheral zone (PZ). Differentiation of central pith tissue is initiated in the rib zone (RZ). Different layers of the SAM are indicated with L1, L2, and L3. The mRNA expression domains of *CLV1*, *CLV3*, and *WUS* are depicted in dark, light-, and mid-grey, respectively. Model for shoot meristem maintenance: WUS expression in the OC promotes an as yet unidentified signal to specify stem cells. The stem cells restrict the range of WUS expression via CLV3 signalling. Cells that have passed the boundary defined by the *CLV* function establish organ founder cell populations. *B*. In analogy, the root meristem contains the QC, which promotes stem-cell identity of the surrounding cells.

more apical and peripheral cells (Schoof *et al.*, 2000). *CLV1* and *CLV2* encode leucine-rich-repeat *trans*-membrane proteins, which can associate to form a heterodimeric receptor molecule (Clark *et al.*, 1997; Jeong *et al.*, 1999). *CLV3* encodes a small secreted polypeptide whose expression domain overlaps with the stem cell region (Fletcher *et al.*, 1999) and forms the ligand for the heterodimeric receptor. *CLV1* is expressed in the L3 layer of the central zone whereas *CLV3* is expressed in a central region of the L1 and L2 layers (Figure 3*A*). It has been observed that *WUS* induces *CLV3* expression to promote stem-cell identity in the upper meristem layers and that *CLV3* acts as a negative regulator of *WUS* expression (Schoof *et al.*, 2000). Expression of *CLV1* in the OC sequesters *CLV3* thereby preventing it from entering the OC and repressing WUS transcription (Lenhard and Laux, 2003). Thus, the size of a cell population in the SAM is controlled by a positive-negative feedback loop which allows for a constant stem cell population despite transient fluctuations such as cell division. The root meristem also contains an

organising centre, the quiescent centre (QC), which was first discovered during laser ablation experiments (Van den Berg *et al.*, 1997) (Figure 3*B*). Overexpression of the *CLV3* homolog *CLE19* resulted in a restriction of root meristem size, suggesting that components of a *CLV* pathway also operate in roots, but in this case *CLE19* affected the stem cell daughters (Casamitjana-Martinez *et al.*, 2003).

Interestingly, it has been recently discovered that *WUS* not only plays an important role in meristem development but that it also acts to maintain embryonic competence. Two gain of function mutations (*pga6-1* and *pga6-2*) were identified that cause a transition from vegetative to embryonic cells leading to the development of SEs from various tissues and organs without any external hormones (Zuo *et al.*, 2002). PGA6 was found to be identical to *WUSCHEL* (*WUS*) suggesting that it plays a predominant role in maintaining embryonic potential.

Furthermore, a systematic expression analysis of a family of Homeodomain (HD) transcription factors that are homologous to the *WUS* gene, known as *WOX* (*WUSCHEL related homeobox*), have been demonstrated to be involved in early partitioning of the embryo (Haeker *et al.*, 2004). *WOX2* and *WOX8* are co-expressed in the egg cell and zygote and become confined to the apical and basal daughter cells of the zygote, respectively, by its asymmetric division (Haeker *et al.*, 2004). *WOX2* not only marks apical descendants of the zygote, but is also functionally required for their correct development, suggesting that the asymmetric division of the plant zygote separates determinants of apical and basal cell fates (Haeker *et al.*, 2004). *WOX9* expression is initiated in the basal daughter cell of the zygote and subsequently shifts into the descendants of the apical daughter apparently in response to signalling from the embryo proper. Expression of *WOX5* shows that identity of the QC is initiated very early in the hypophyseal cell and highlights molecular and developmental similarities between the stem cell niches of root and shoot meristems (Haeker *et al.*, 2004). Thus, the different cell division patterns of the zygote and its progeny are preceded by differences in the expression of HD transcription factor genes. It will be interesting to see whether the expressions of such genes are also involved in the early patterning and establishment of SEs.

1.10. Gene Expression Microarrays (GEMs) - A Powerful Tool for Large Scale Gene Identification

A number of genes have been identified from somatic embryogenesis (SE) by a variety of methods using strategies that rely on the differential expression of genes in somatic embryo cells and callus cells. For example, the differential display method that works by systematically amplifying the 3' terminal portions of mRNAs and resolving these fragments on a sequencing gel, allowing direct side-by-side comparison of mRNAs between differential cell types, has been used with some success. (Schmidt *et al.*, 1997) used this method to isolate a cDNA (SERK) that was differentially expressed in embryogenic cells and encodes a putative receptor-like protein kinase. Analysis using promoter-marker lines indicated that this gene was expressed in most of the cell clusters giving rise to somatic embryos (SEs) or SEs up to the globular stage.

However, studies by (Sung and Okimoto, 1981) using 2-D PAGE to look for patterns of gene expression during SE in carrot suspension cultures, found very few differences in protein patterns between SEs and proliferating cells. Similarly, (Wilde *et al.*, 1988) examined changes in gene expression between proembryonic masses (PEMs) and embryos and also found few differences in 2-D PAGE patterns from PEMs and torpedo stage mRNAs.

There has been some success using subtractive differential screening of mRNA libraries to isolate unique sequences involved in tissue specific expression in plants. Differential screening is a variant of the nucleic acid hybridisation method that is particularly suitable for isolating tissue specific or developmentally regulated cDNA sequences, or clones, derived from mRNAs induced by particular treatment. In order to increase the efficiency of differential screening cDNA libraries can be enriched with the desired sequences by subtractive hybridisation that removes those cDNA sequences that are ubiquitous or not differentially expressed (Old and Primrose, 1994).

cDNA microarray is a relatively new technology platform that allows for the simultaneous monitoring of expression levels of many different genes. cDNA

microarrays consist of thousands of individual DNA sequences printed in a high density array on a glass microscope slide by a robotic arrayer. The relative abundance of the spotted DNA sequences in two RNA samples can be assessed by monitoring the differential hybridisation of the two samples to the sequences on the array. For mRNA samples, the two samples are reverse transcribed into cDNA, labelled using different fluorescent dyes (Cy3 and Cy5), mixed in equal proportions and then hybridised with the array DNA sequences. Following this competitive hybridisation, the slides are imaged using a scanner and fluorescent measurements are made for each dye at each spot on the array. The ratio of the red and green fluorescence intensities for each spot indicates the relative abundance of the corresponding DNA sequence in the two nucleic acid target samples.

mRNA levels are highly informative about cellular states and the activity of genes and for most genes mRNA abundance is related to protein abundance. Microarrays can be thought of as simply a more powerful substitute for more conventional methods for evaluating mRNA abundance. Although some success has been achieved using alternative methods for analysing the differential expression of genes or their protein products, a cDNA microarray based approach offers the potential to greatly improve our understanding of SE. By combining the utilisation of the vast amounts of sequence data generated by public and private sequencing initiatives with the sensitivity of detecting gene differential expression, this method offers an unprecedented opportunity for large-scale gene discovery.

1.11. Current Methods for Expressional Analysis: Northern analysis and In situ Hybridisation

Two methods are commonly used to study gene expression at the transcriptional level: Northern analysis and *in situ* hybridisation (ISH). In Northern analysis, RNA populations are isolated from a large group of cells, separated in agarose gels, blotted to a membrane and hybridised with labelled probes. Northern analysis offers a relatively quick and convenient mean to assess the expression of a specific gene at tissue, organ or whole plant levels. However, it does not provide spatial and temporal information of the expression patterns at cellular level and cannot use multiple probes for analysis. On the other hand, the ISH technique involves complicated and time-

consuming cytological techniques such as fixing, embedding and sectioning. However, it can reveal the expression patterns in individual cells and tissues. In the ISH technique, the detection of nucleic acids (mRNA or DNA) is achieved by taking advantage of the nature of nucleic acids to undergo complementary 'Watson and Crick' base-pairing. This allows for any nucleic acid sequence to be studied by constructing an 'antisense' labelled probe to the sequence of interest, enabling rapid analysis of the spatial and temporal expression of RNA or DNA in tissues. This is particularly important in the field of developmental biology as the temporal and spatial distribution of gene expression is fundamental to developmental processes and understanding gene function.

1.12. Experimental Design

Taking a microarray based approach this project aims to utilise this technology in order to identify genes that are up- or down-regulated during somatic embryogenesis (SE) in wheat. The basic experimental design consists of:

1. Subtractive cDNA microarray analysis to identify differentially expressed genes specifically associated with SE.
2. Characterisation of genes of interest by:

   Bioinformatic sequence analysis

   ISH

   DNA sequence analysis.
3. Functional analysis of selected genes of interest by reverse genetics, RNAi technology and genetic overexpression in plant tissue.

1.13. Research Aims and Objectives

The primary purpose of this project is to increase our understanding about the mechanisms underlying somatic embryogenesis (SE) and cell totipotency. It is hypothesised that there are essentially two critical steps, in culture, that contribute to the phenomenon of SE in wheat. Firstly, the treatment of somatic tissue with the auxin, 2,4-D, results in substantial genetic re-programming, cellular de-differentiation and embryogenic callus formation. Secondly, on removal of 2,4-D the tissue undergoes further genetic events that result in the re-differentiation and regeneration of embryogenic tissue. It can also be hypothesised that auxin concentration gradients are responsible for these observations. However, it is clear that the program of events

during SE that follow induction by experimental manipulations is too elaborate and complex to be controlled by a single hormone. The effects caused by the addition of the hormone 2,4-D must trigger a chain of molecular events, probably involving a genetic signalling cascade that results in the genetic reprogramming of the cell and altered gene expression, leading to the observed formation of embryogenic calli. The subsequent removal of the same auxin causes a second molecular response that result in the re-differentiation of the embryogenic calli into adult structures. Therefore, the aim of this project is to identify and characterise the genetic components responsible for triggering the initiation and development of SE in wheat that result from the addition and removal of 2,4-D from the tissue culture medium.

Chapter Two

Experimental Testing of Wheat Donor Explants

2 Chapter Two – Experimental Testing of Wheat Donor Explants

2.1 Introduction

Plant tissue culture is essentially a technique in which isolated plant cells are grown on an artificial nutritive medium. Although these techniques were largely developed during the 1950s they have their origins in plant propagation, i.e. taking cuttings, which have been applied by horticulturists for centuries and are based on the remarkable plasticity of plant development. Typically a sterile sample of young actively growing tissue is taken and placed onto a flask or Petri dish containing a nutrient solution that contains macro nutrients (i.e. Ca and K), micronutrients (i.e. Fe and Zn), vitamins (i.e. thiamine and riboflavin), sugars (i.e. sucrose and glucose), sugar-alcohols (i.e. myo-inositol) and may also include amino acids, nucleic acids and other organic molecules. If the medium has no phytohormones or no growth regulators it is referred to as a basal medium, such as the widely employed Murashige and Skoog (MS) medium (Murashige and Skoog, 1962). Phytohormones and/or plant growth regulators can also be added to the basal medium to stimulate growth and development of the explant. Auxin and cytokinins are the most commonly used phytohormones and the ratio of the two can be specifically altered to regulate growth and development and optimize the regeneration process.

2.1.1. Somatic Embryogenesis in Carrot (*Daucus carota*)

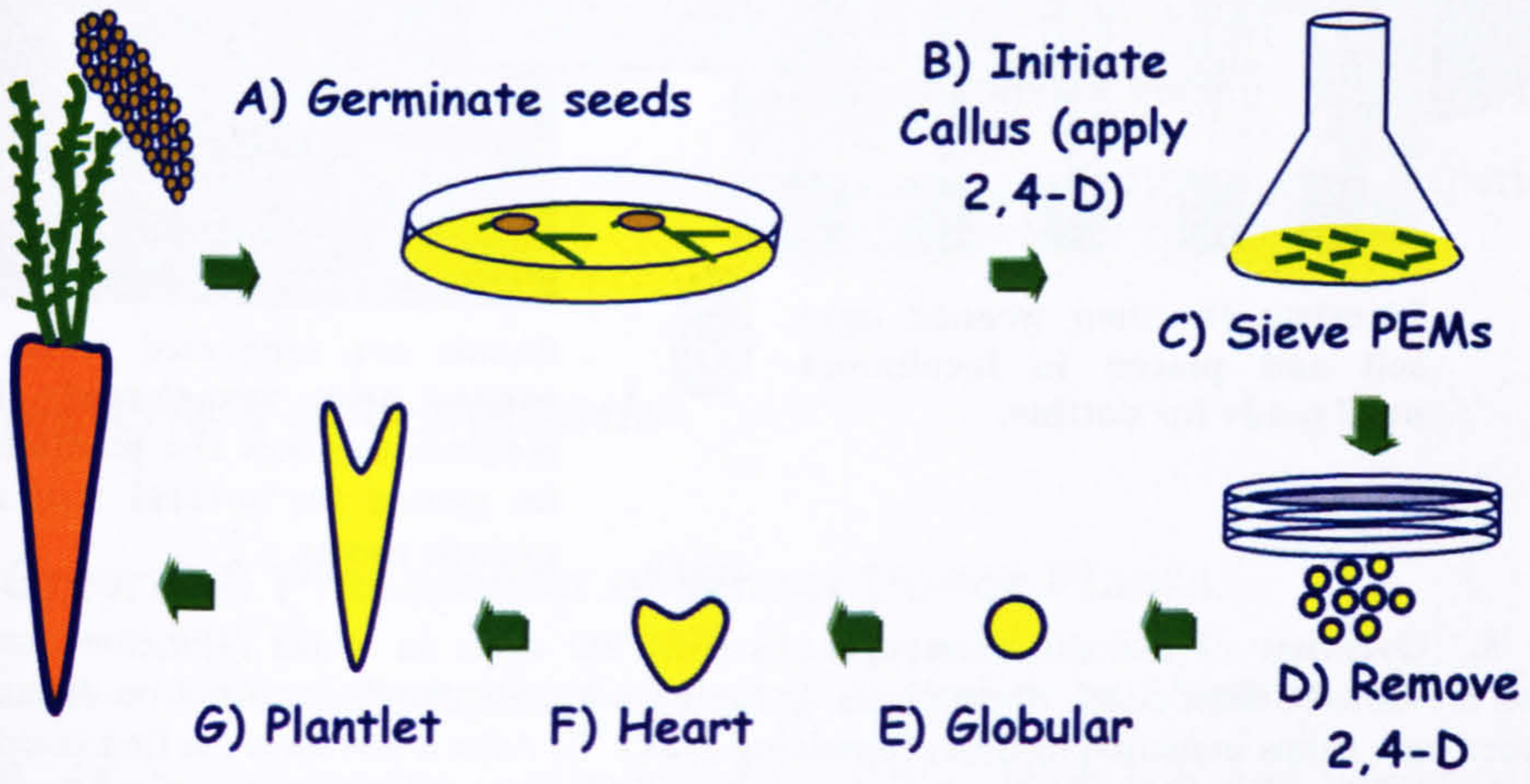


Figure 4. Overview of Somatic Embryogenesis (SE) life cycle in carrot (*Daucus carota*). Calli are initiated from hypocotyl cultures using the synthetic auxin, 2,4-D (A+B). The Pro-Embryogenic Masses (PEMs) are selected on the basis of size and removed from 2,4-D (C+D) where they differentiate through several stages of embryogenesis (E-G) into fully developed adult plants.

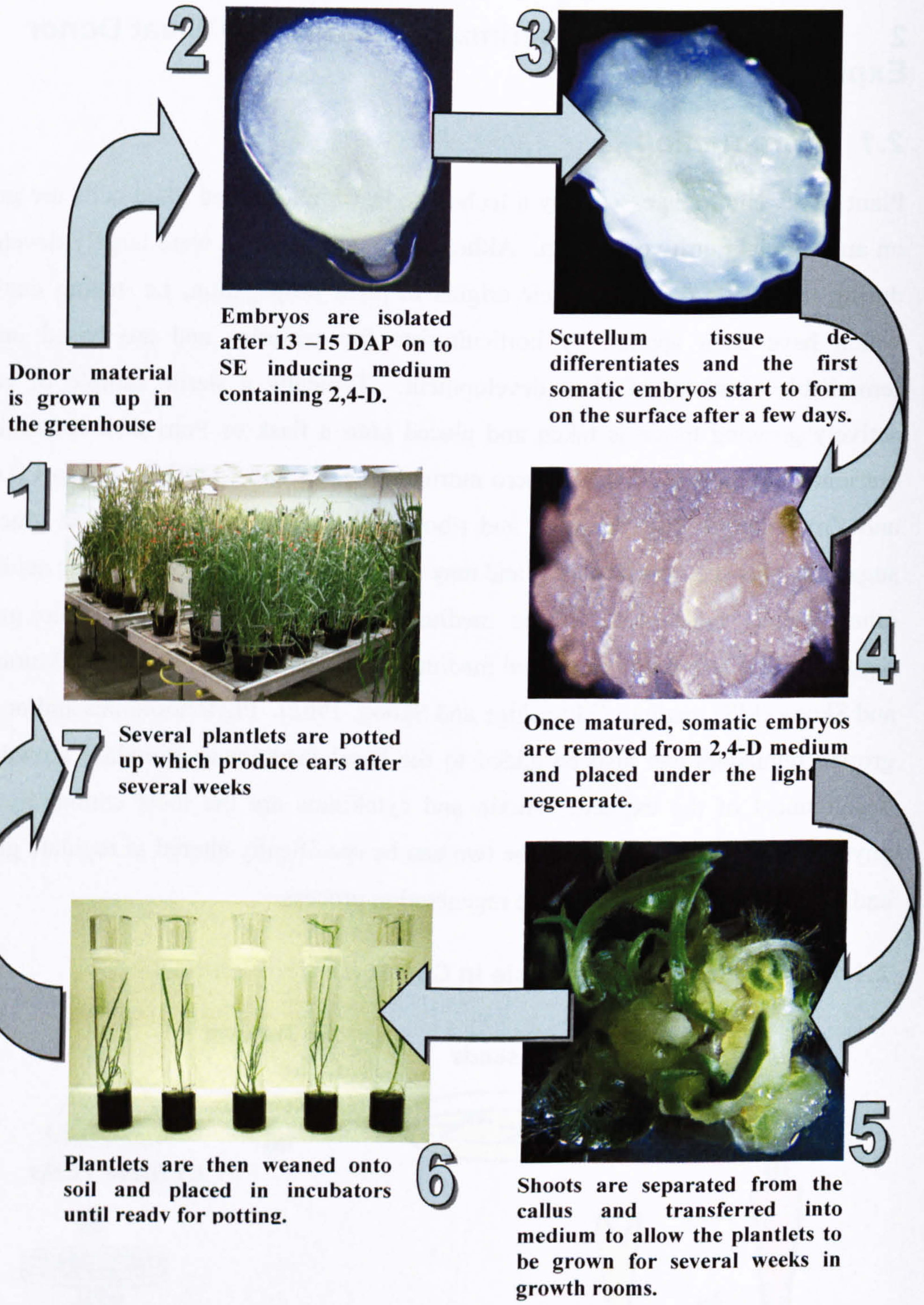


Figure 5. Overview of Somatic embryogenesis (SE) life cycle in wheat (*Triticum aestivum)* 1; Caryopsi are isolated from donor wheat plants. 2; Embryos are aseptically excised from their caryopsis and placed onto callus induction medium containing 2,4-D. 3; After a few days the first morphological signs of somatic embryo formation can be observed on the edges of the scutellum. 4; Mature SEs are removed from 2,4-D medium and placed under the light to regenerate into adult organs. 5; Emerging shoots are separated from the calli and allowed to establish into plantlets. 6; Plantlets are then weaned onto soil in incubators. 7; Plantlets are grown up in the greenhouse and are able to produce ears after several weeks from which a second generation of caryopses can be isolated.

The pioneering work of Reinert and Steward in the 1950s (Reinert, 1958; Steward *et al.*, 1958) using carrots employed a suspension culture system which consisted of hypocotyl derived single cells or cell clumps called pro embryogenic masses (PEMs) that were initiated using 2,4-D. These cells or PEMs were then separated on the basis of size, removed from 2,4-D and suspended in aerated liquid medium from which adult plants could be generated (Fig. 4.)

2.1.2. Somatic Embryogenesis in Wheat (*Triticum aestivum*)

In contrast to the carrot suspension cultures, wheat tissue culture explants are grown on a semi-solid medium to form callus tissue from which the regeneration of plantlets is carried out (Fig. 5).

2.1.3. Establishment of Suitable Experimental Cultivars

In order for the aims of the project to be met and the hybridisation experiments to be effectively carried out, it was first necessary to establish and compare two suitable experimental systems from which sample populations could be collected. That is, an embryogenic, wheat cultivar against a non-embryogenic, wheat control. The spring wheat 'Bobwhite' (BW) was selected as the embryogenic cultivar because it is one of the most responsive and frequently used varieties in wheat tissue culture. Conversely, the winter wheat 'Savannah' (SAV) was chosen as the experimental control because of its characteristically poor embryogenic response to tissue culture conditions. In order to evaluate the two wheat varieties as experimental systems for use in this study, BW and SAV immature embryos (IEs) were tested in culture to assess their responsiveness to the synthetic auxin, 2,4-dichlorophenoxyacetic acid (2,4-D) and their respective effectiveness as embryogenic and non-embryogenic experimental samples.

2.2. Materials and Methods

2.2.1. Growth & Preparation of Wheat Donor Plants.

Savannah is a winter wheat, therefore, seeds needed to be vernalised for a 6-week period at 4°C before sowing (not required for BW). Both BW and vernalised SAV seeds were sown into 3-litre pots containing a mixture (50/50, v/v) of John Innes No. 3 and high grade peat. Osmocote controlled release fertiliser (containing N:P:K in the

ration of 16:8:12) was added to the compost mixture at 40g/25 litres. The pots (each containing 7 seeds) were placed on capillary matting in growth cabinets with a 16/8 hours light/dark cycle and temperature and humidity set at 18°C and 70% in the light and 12°C and 80% in the dark respectively. The light intensity during the photoperiod was approximately 350μmols at bench height. Once germinated, seedlings were thinned down to 4 per pot. Ears were produced approximately 8 weeks after sowing and were tagged, at the first sign of anthesis, with a small label to mark the date as a reference for donor material collection.

2.2.2. Isolation & Culture of Wheat Immature Embryos

Wheat spikes were cut at 13 and 14 days after pollination (DAP) and the caryopses isolated from the ears by hand. Isolated caryopses were surface-sterilised by briefly rinsing in 70% ethanol (EtOH), followed by 15 minutes incubation in 5% Sodium Hypochlorite (NaOCl) plus 0.1% (v/v) Tween-20 on a horizontal shaker at 120 rpm. Caryopses were then rinsed four times with autoclaved water in a sterile flow hood. Immature embryos (IEs) were aseptically isolated from each caryopsis in the flow hood, under a dissection microscope, using a sterile scalpel and forceps and placed onto the callus induction medium (2.2.3.2) with the scutellum facing upwards and the embryonic axis in contact with the medium (approximately 50 IEs/Petri dish).

2.2.3. Wheat Tissue Culture Media

All tissue culture media were prepared under sterile conditions and poured into sterile 10cm Petri dishes (Corning) in a flow hood cabinet. Once set, media dishes were colour coded and placed in sealed sterile bags at RT in the dark until required.

2.2.3.1. Murashige and Skoog (MS) Medium

Murashige and Skoog (MS) medium contains the full strength of the macro and micro salts and vitamins as described by (Murashige and Skoog, 1962) and was made using 4.4g/L of MS salts and vitamins mixture (Invitrogen, 10632), 30g/L sucrose, 300mg/L Glutamine and 150mg/L Asparagine. The pH was adjusted to 5.8 with KOH and 7g/L Phytagel (Sigma, P8169) was added before autoclaving for 30 minutes at 115^0C.

2.2.3.2. Callus Induction (3MS3S) Medium

The callus induction medium contained all the chemical ingredients of the MS medium (2.2.3.1.) plus 3 mg/L of 2,4-D, which was added before autoclaving.

2.2.3.3. Regeneration (NG) Medium.

The regeneration medium was MS medium (2.2.3.1.) with the addition of filter sterilised 1 mg/L gibberellic acid (GA3) and 1 mg/L naptheleneacetic acid (NAA) after autoclaving.

2.2.4. Callus Induction & Regeneration Frequencies

Donor material was prepared and collected, as described in 2.2.1. and IEs were aseptically isolated, as described in 2.2.2. For each experiment, one hundred BW and SAV IEs were placed on the callus induction medium (2.2.3.2.) and incubated at 25^0C/40% humidity in the dark for 2 weeks. Callus tissue that developed during the two-week initiation period was then transferred onto the regeneration medium (2.2.3.3.) and cultured at 25^0C/40% humidity under continuous light for one further week. Regeneration (as evidenced by the emergence of green shoot structures) was assessed under a dissecting microscope and the regeneration efficiencies calculated using the following formula:

$$\textbf{Regeneration Frequency} = \frac{\textbf{Number of calli producing shoots}}{\textbf{Total number of calli}}$$

To assess the effects of sample variation, the experiment was repeated a total of three times.

2.3 Results

2.3.1. Morphological Response of Donor Explants

As previously mentioned, the experimental strategy of the project requires two donor plant systems, an embryogenic cultivar responsive to SE induction, and a non-embryogenic cultivar which is unresponsive to SE induction. BW and SAV IEs were tested in culture, to assess their responsiveness to 2,4-D and their respective effectiveness as embryogenic and non-embryogenic experimental samples. Over three separate experiments one hundred IEs were isolated for callus induction and regeneration under light (Materials and Methods 2.2.4). It was observed that nearly

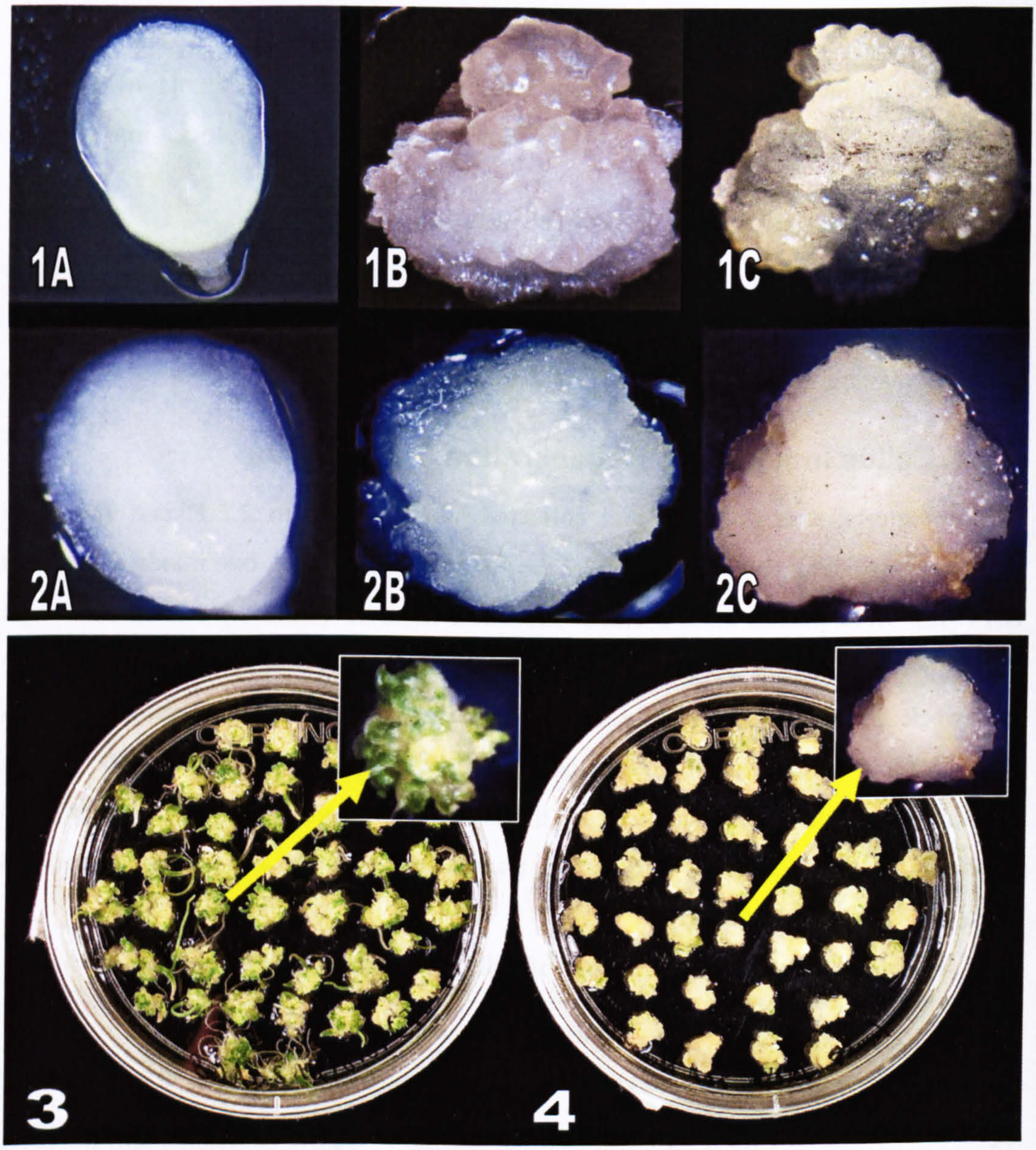


Figure 6. The differing responses between the two experimental cultivars BW (l & 3) and SAV (2 &4) under tissue culture conditions; 1A and 2A = freshly excised embryos. 1B and 2B = 1 week incubation on callus induction medium. 1C and 2C = 2 weeks incubation on callus induction medium. 3 & 4 = 2 weeks callus induction medium and 2 week incubation on regeneration (NG) medium under light. 3 illustrates the large number of regenerating shoots observed on BW calli that have developed from somatic embryo structures, whereas in 4 the SAV calli have little or no regenerating shoots forming at all.

all of the BW IEs exhibited a strong embryogenic response characterised by the swelling of scutellar tissue and formation of smooth, glossy and organised embryogenic clusters on the surface of each callus. In contrast, the majority of SAV IEs produced a weak embryogenic response characterised by enlarged, watery and disorganised, non-embryogenic calli from which there was little structured growth. The differences in the observed responses between the two cultivars when subjected to SE-inductive tissue culture conditions during these experiments can be seen opposite in Figure 6.

2.3.2. Regeneration Frequencies of Bobwhite and Savannah

For each experiment, the numbers of developing shoots that were scored under a light microscope and the regeneration frequencies from each cultivar that were calculated from the formula given in materials and methods 2.2.4. are illustrated in Table 1. These results highlight the difference in regeneration between the embryogenic cultivar BW, which has a mean regeneration frequency of 16.82 ± 3.01 shoots per calli and the non-embryogenic SAV Calli which have a mean regeneration frequency of 2.97 ± 0.78 shoots per calli. This clearly establishes the suitability of the two experimental systems as useful donor explant samples in evaluating the difference in gene expression between embryogenic and non-embryogenic calli.

Table 1: Regeneration frequencies of Bobwhite and Savannah cultivars.

Genotype	Experiment No.	No. of Shoots	No. of Calli	Regeneration Frequency
Bobwhite	1	1860	100	18.6
Bobwhite	2	1334	100	13.34
Bobwhite	3	1852	100	18.52
Bobwhite	**Total**	**5046**	**300**	**16.82±3.01**
Savannah	1	225	100	2.25
Savannah	2	285	100	2.85
Savannah	3	380	100	3.8
Savannah	**Total**	**890**	**300**	**2.97±0.78**

2.4. Discussion

The success of the subtractive microarray experiments depends on the use of an effective non-embryogenic control. Besides tissue culture conditions there are a number of factors, such as explant type and genotype, which affect the induction of the embryogenic state. A different explant type could be used for the non-embryogenic control e.g. BW leaf, however as the starting material is so different from the immature embryo this would make microarray analysis less straightforward.

The explant has not only to exhibit a poor embryogenic response to the same tissue culture conditions but also preferably be of identical or similar origin as this would allow for a more simplified and powerful analysis of the generated microarray data. It was apparent from wheat tissue culture studies that cultivars differ in their embryogenic response to tissue culture (Redway *et al.*, 1990). Individual genotypes within a given species vary greatly in their embryogenic capacity and many genotypes appear recalcitrant to tissue culture. There have been a number of reports on the identification of callus types that are responsive to tissue culture conditions (Weeks *et al.*, 1993). There are fewer reports, however, on non-regenerative cultivars. Using the experience of the Syngenta wheat team who identified SAV as having a poor response to tissue culture conditions this was chosen to fully evaluate its response.

The differing response to tissue culture conditions of the two experimental cultivars can be seen in figure 6 in the results section. Firstly, the initial comparison of the early morphological response between BW and SAV clearly demonstrates that the BW genotype exhibits a strong embryogenic response. This is in stark contrast to the weak embryogenic response exhibited by the SAV. Furthermore, the photographs in figure 6 taken of the plates for each cultivar following growth under light, also clearly demonstrate the high efficiency of the embryogenic BW cultivar to regenerate adult root and shoot structures. In contrast the non-embryogenic SAV cultivar failed to form any real regeneration of root and shoot structures whatsoever. Although there was nodulation on some of the calli, this did not take on the smooth, glossy appearance as in BW and resulted in fewer, weaker shoots regenerating from small, localised embryogenic clusters that appeared on a only few of the calli.

These morphological examples are further supported by an analysis of the regeneration frequency between these two cultivars (Table 1) which clearly demonstrates, in a replicable manner, the superior regeneration frequency exhibited by the BW cultivar: BW has an average regeneration frequency of 16.82 ± 3.01shoots per calli confirming previous reports that BW is a highly regenerable cultivar (Pellegrineschi *et al.*, 2002), whilst SAV has an average of 2.97 ± 3.01 shoots per calli. The fact that there is the regeneration of a small number of embryo and shoot structures in the SAV cultivar mean that there will be some expression of the genes involved in the initiation and development of SE. However, it is likely that due to the

majority of the tissue being non-embryogenic the expression will be diluted to the extent that it is undetectable in the microarray hybridisations. These results therefore, clearly establish the suitability of the two experimental systems as useful donor explant samples for evaluating the differential gene expression associated with SE. By identifying SAV as a non-regenerable control comparisons can be made with closely related starting material that will allow for a more simplified analysis of microarray data.

Chapter Three

Isolation of RNAs for Microarray Analyses

3 Chapter Three: Isolation of RNAs for Microarray Analyses

3.1. Introduction

Obtaining high quality, intact RNA is the first and often the most critical step in performing many fundamental molecular biology experiments, including cDNA microarray hybridisations. To be successful however, the RNA hybridisation procedure should include some important steps before, during and after the RNA purification. Firstly, the treatment and handling of cells or tissues prior to the isolation of RNA is typically when RNA integrity is at risk, and correct procedures must be ensured when handling tissues. Secondly, thorough cellular disruption is critical for high RNA quality and yield. RNA that is trapped in intact cells is often removed with cellular debris and is therefore unavailable for subsequent isolation. Therefore it is crucial to choose a disruption method best suited to the tissue and organism to maximise yield. In plants this can be achieved by freezing the tissues in liquid nitrogen and grinding the samples with a mortar and pestle to a fine powder which is added to the denaturant. The total RNA samples are mainly comprised of ribosomal RNA and mRNA quality has historically been assessed by electrophoresis. While crisp rRNA bands are indicative of intact RNA it does not actually reflect the quality of the underlying mRNA. An improved analytical tool for total RNA analysis is the Agilent 2100 bioanalyser which uses a combination of microfluidics, electrophoresis and fluorescents to evaluate both RNA concentration and integrity. Another advantage is that it requires very small inputs allowing the assessment of RNA quality in limiting samples. Poly (A) RNA makes up between 1-5% of total RNA and is used for the synthesis of probes for array analysis to increase the detection of rare transcripts. Poly (A) isolation allows high yields of mRNA from total RNA without the co-isolation of rRNA. Therefore the isolation and quantification of high quality RNA from each of the selected timepoints can be seen as critical to the success of the hybridisation strategy outlined in figure 7. This hybridisation strategy has been designed to test the hypothesis that there are essentially three critical steps, in culture, that contribute to the phenomenon of somatic embryogenesis (SE). The first step, is that the treatment of somatic tissue

with the auxin 2,4-D, induces substantial genetic re-programming, resulting in cellular de-differentiation and acquisition of embryogenic competence. The second step is the development and formation of somatic embryos (SEs) and embryogenic tissue from these competent cells or cell clusters. Finally, the removal of 2,4-D from the medium induces further differentiation and regeneration of SEs into proper plantlets.

3.2. Materials and Methods

3.2.1. Microarray Hybridisation Strategy

Microarray experiments generate large and complex data sets and the greatest difficulty lies in the analysis rather than the generation of such large amounts of data. For successful interpretation of the microarray data, it is important to design the experiments in such a way that makes the analysis as simple as possible. For that purpose, the SE induction process was divided into 5 different timepoints, as outlined below in figure 7, for sample collection. These time points were chosen to represent: 1) the de-differentiation of somatic cells towards embryogenic competence (Timepoint 1); 2) the initiation and development of SEs (Timepoints 2 & 3) and 3) the response of SEs to the removal of 2,4-D (Timepoint 4).

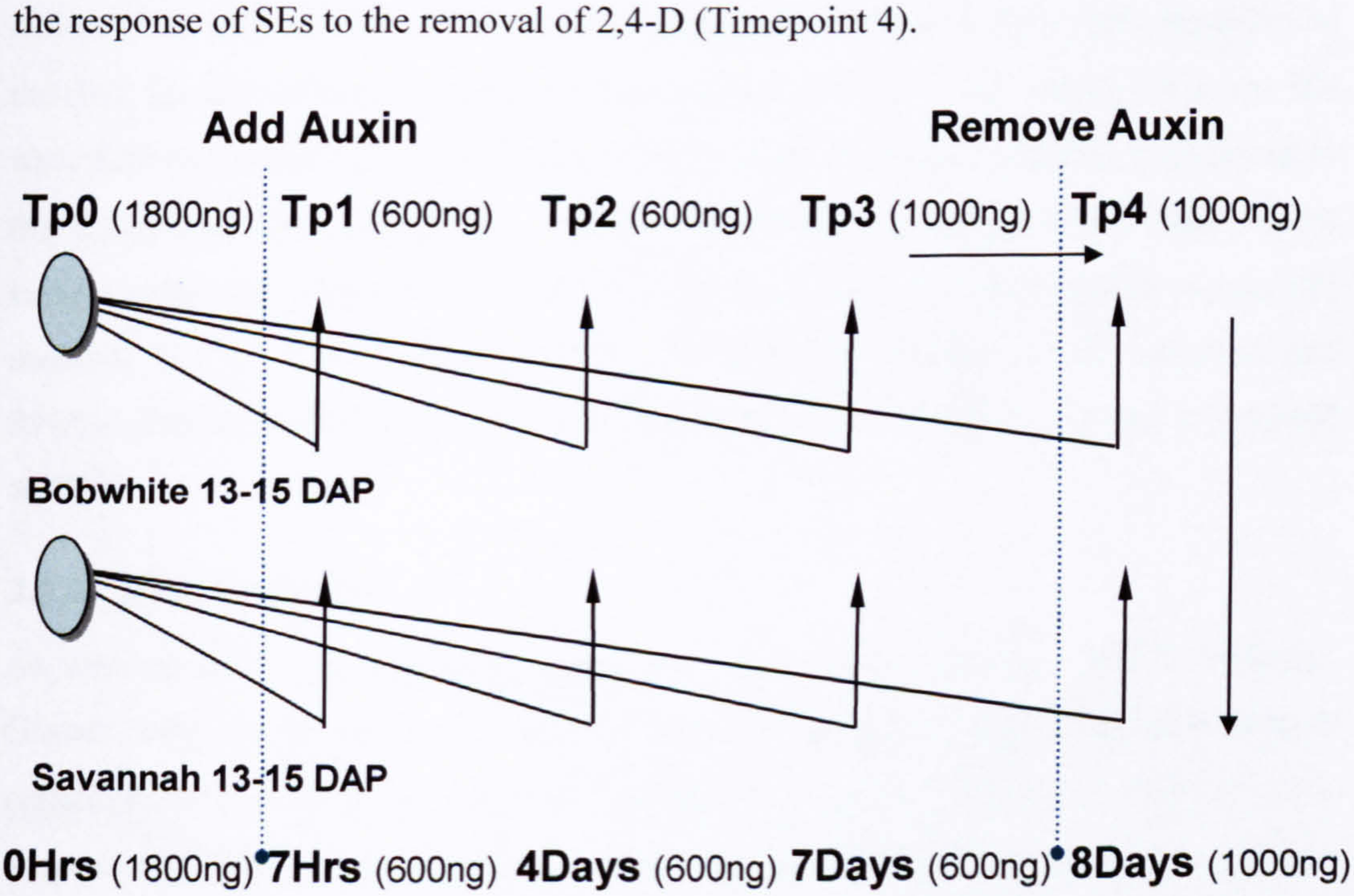


Figure 7. Overview of the different timepoints for sample collection. Tp0: Timepoint 0 (0 hrs on 2,4-D); Tp1: Timepoint 1 (7 Hrs on 2,4-D); Tp2: Timepoint 2 (4 Days on 2,4-D); Tp3: Timepoint 3 (7 Days on 2,4-D); Tp4: Timepoint 4 (Day 8, 24 Hrs after 2,4-D removal).

The microarray hybridisation strategy is to compare the 2,4-D treated samples (timepoints 1, 2, 3 & 4) with the untreated standard reference (timepoint 0) within the same cultivar to identify genes that are differentially expressed during various stages of the induction process in Bobwhite (BW) and Savannah (SAV) calli separately. The differentially expressed genes from the SAV results will then be used to subtract common elements from the BW hybridisations to identify differentially expressed genes specifically associated with SE. Such comparisons will help eliminate the 'background noise' of genes not associated with the embryogenic response.

3.2.2. Collection of Donor Material

For RNA extraction, wheat donor material was grown and prepared as described in (2.2.1.) and immature embryos (IEs) isolated as described in (2.2.2.). For BW, we used a morphological marker in the selection of the IEs in order to maintain sample uniformity by selecting only those embryos with a translucent appearance and approximately between 0.75 and 1.25mm in diameter. The SAV IEs isolated at the same DAP were slightly larger than the BW IEs (at approximately 1-1.5mm in diameter) and more opaque in colour. The IEs for Timepoints 1, 2, 3, & 4 were isolated and then incubated on the callus induction medium at 25°C/40% humidity in the dark for the indicated lengths of time outlined in Fig. 7. For Timepoint 4, the IEs were first incubated on the callus induction medium for 7 days and then transferred to the 2,4-D-free MS medium for a further 24 hours' incubation under light before sample collection. IEs for the control (Timepoint 0) were isolated directly on the MS medium (2.2.3.1.) and harvested without incubation. Samples were collected into RNAse-free Eppendorf tubes, weighed, labelled, frozen in liquid nitrogen and stored at -80^0C .

3.2.3. RNA Isolation

As with all RNA work, standard sterilisation procedures were followed throughout. Gloves were worn when handling tissues and equipment and to remove RNase contamination, pestles, mortars and spatulas were soaked overnight in Virkon®, pre-treated with RNase ZAP® (Fluka, 83930) and rinsed several times with Diethyl pyrocarbonate – (DEPC) (Sigma, D5758) treated water 0.1%(v/v). Solutions were incubated with DEPC (0.1% v/v) overnight and then autoclaved at 115^0C for 30 minutes to inactivate the DEPC.

3.2.3.1. Total RNA Extraction from Wheat Tissues

The protocol used to extract total RNA was adapted from (Chang *et al.*, 1993). Briefly, samples were ground to a fine powder in liquid nitrogen using a pestle and mortar and transferred to a pre-cooled 50ml Falcon tube on dry ice. Extraction buffer (2% CTAB, 2% PVP 40,000, 100mM Tris-HCl, 25mM EDTA, 2M NaCl, 0.5g/L spermidine) with 2% beta-mercaptoethanol (ß-ME) was heated to 65°C and added to each sample at the ratio of 2ml (buffer)/100mg (plant tissue). Samples were briefly vortexed, incubated at 65°C for 5 minutes and then spun in a Sorvall centrifuge at 10,000 rpm for 20 minutes at RT to pellet out any starch and debris. The supernatants were transferred to fresh 50 ml Falcon tubes and extracted with an equal volume of chloroform: isoamyl alcohol (IAA) 24:1 (v:v) (Sigma, C0549) and spun in a bench-top centrifuge at 3,000 rpm for 10 minutes at RT. This extraction step was repeated until a clear aqueous phase was obtained. RNA was precipitated overnight at 4°C with a $^1/_4$ volume of 12M LiCl and harvested by centrifugation at 10,000 rpm for 20 minutes at 4°C in a Sorvall centrifuge. The pellet was dissolved in 500 µl pre-warmed (65°C) SSTE buffer (1.0M NaCl, 0.5% SDS, 10mM Tris-HCl pH 8.0, 1mM EDTA pH 8.0), extracted with equal volume of phenol:chloroform:IAA 25:24:1 (v:v:v), precipitated with 2 volumes of ice-cold EtOH for one hour at -80°C and finally dissolved in 20-50 µl DEPC treated water and stored at -80°C.

3.2.3.2. Poly (A) Messenger RNA Isolation

Poly (A) messenger RNA (mRNA) was isolated from total RNA using a commercial mRNA purification kit (QIAGEN, 70042) according to the manufacturer recommended protocol. Each total RNA sample was first adjusted to 250µl volume with DEPC-treated H_2O and then mixed with 250µl of OBB binding buffer and 15 µl Oligotex resin. The mixture was incubated at 65°C for 3 minutes to denature the RNA and then at 37°C for 20 minutes with shaking (100rpm) to allow the binding of mRNA to the resin. The Oligotex resin was pelleted by spinning for 2 minutes at 12,000rpm in a microcentrifuge and the supernatant discarded. The resin was then washed twice with 500µl 0.5 x OW2 solution (10mM Tris-Cl ph 7.5, 150mM NaCl, 2mM EDTA, 0.2% SDS). To elute the mRNA, the resin was resuspended in 30µl pre-warmed (70°C) RNase-free TE buffer (10mM Tris-Cl pH 7.5, 1mM EDTA), transferred to a spin column and spun at 12,000rpm for 1 minute in a microcentrifuge.

To achieve maximum yield, the elution step was repeated using 20µl TE. The mRNA samples were stored in siliconised Eppendorf tubes (Ambion, 12450) at –80°C.

3.2.4. RNA Quantification

Three methods were used in assessing the integrity and quality of RNA samples: the Agilent 2100 Bioanalyser (for assessing the integrity of RNA samples); OD reading (for quantifying the total RNA samples) and Ribogreen™ assay (for quantifying mRNA).

3.2.4.1. Agilent 2100 Bioanalyser

The Agilent 2100 Bioanalyser, 'Lab-on-a-Chip' technology uses electrokinetic forces to move, mix and separate nucleic acid samples. Electrodes make contact with the samples to create electrokinetic forces that drive fluids through selected pathways in the chip and then send the information through to a computer. The miniaturised fluid pathways shorten running times and the strong electrokinetic forces improve analytical resolution. This allows for increased detail and accuracy of analysis while simultaneously reducing the sample size required. Samples were run according to the Agilent 2100 Bioanalyser handbook as follows:

An electrode-cleaning chip containing 350µl of RNase ZAP® (Fluka, 83930) was placed in the Agilent 2100 for 1 minute to decontaminate the electrodes. This was replaced for 10 seconds by another electrode-cleaning chip containing 350 µl DEPC-treated water, 0.1% (v/v) to rinse off the RNase Zap® and the electrodes allowed to dry. Nine microlitres of gel dye were then added to a new RNA 6000 Nano LabChip® (Agilent, 5065-4476) followed by the addition of 5µl of buffer and 1µl of RNA sample to each sample well on the chip. The RNA 6000 Nano LabChip® was then vortexed for 1 minute to mix the contents and remove any air bubbles before placing it in the Agilent 2100 bioanalyser and running the RNA assay from within the computer software program.

3.2.4.2. OD Reading

Total RNA concentration and purity was determined using optical density (OD) readings obtained from a Cecil CE 8020 spectrophotometer. Typically, 1µl RNA sample was added to 500µl of TE in a quartz cuvette and the absorbance at OD_{230} (polysaccharide) OD_{260} (RNA) and OD_{280} (protein) was measured in the

spectrophotometer. The RNA concentration was calculated using the following equation:

RNA concentration = OD_{260} x 40 x dilution factor (500) (µg/ml)

3.2.4.3. Ribogreen™ Assay

The most accurate quantification of RNA is determined spectrophotometrically by obtaining an OD_{260} value. Availability of tissue and hence mRNA yield often precludes the use of OD_{260} readings for assessment of mRNA concentration. Furthermore, although the Agilent 2100 Bioanalyser provides a reasonable estimate, it is not sufficiently accurate in determining RNA concentrations. The Ribogreen™ assay however, is a simple yet highly accurate fluorescent-based nucleic acid detection assay. It works by using a fluorescent dye (Ribogreen™) that intercalates with nucleic acids and is accurately detected by a fluorescent plate reader. The RiboGreen™ reagent is non-fluorescent when free in solution but upon binding RNA, the fluorescence of the RiboGreen™ reagent increases more than 1000-fold. The RNA bound RiboGreen® reagent has an excitation maximum of ~500nm and an emission maximum of ~525nm similar to Fluorescein which can be accurately detected by a fluorescent plate reader. The procedure for the Ribogreen™ assay is described below: A standard dilution series was set up using a 200ng/µl RNA stock solution (Sigma, R1253) diluted in TE (10 mM Tris-Cl pH 7.5, 1 mM EDTA) to give standards of 4000ng/µl, 2000ng/µl, 1000ng/µl, 500ng/µl, 250ng/µl, 125ng/µl and finally a zero standard consisting of just TE. For the controls, the RNA stock solution was used directly at a concentration of 200ng/µl for the first control and diluted in TE to give 50ng/µl for the second control. Each of the standards and controls were loaded, in duplicate, into a 96-well black bottomed microtitre plate (Greiner Bio-One) and mixed with a (1:100) Ribogreen™ solution. The plate was wrapped in foil, incubated at room temperature for 5 minutes and scanned with a Tecan fluorescence plate reader (SPECTRAFlourPlus, Tecan). Using the XFluor software, the results were presented in an excel spreadsheet and used to generate a standard curve. The fluorescence assay was then repeated with the RNA samples to give values which were used to calculate the RNA concentrations from the standard curve.

3.2.4.4. Glycogen Precipitation

For a successful hybridisation Incyte GEMs™ require 600ng of mRNA for each probe at a concentration of 50ng/µl. The concentrations of the mRNA samples were often

below this threshold, so glycogen precipitation was used, where necessary, to concentrate the RNA. Samples, suspended in 40µl TE, were mixed with 1µl 5 mg/ml glycogen (Sigma, G1508), 10µl 3M NaOAc (pH 5.2) and 250µl 100% EtOH. The mixture was left at -80°C for 30 minutes and then spun at 12,000rpm in a microcentrifuge for 20 minutes at 4°C. Each pellet was washed with 250µl of 75% EtOH, air-dried for 10 minutes at RT and then dissolved in 20µl of 1x TE. RNA was re-quantified using the Agilent 2100 Bioanalyser and Ribogreen™ fluorescence assays (3.2.4.1, 3.2.4.3.) and the volume adjusted accordingly to give a final concentration of 50ng/µl.

3.3. Results

3.3.1. Embryo Harvesting and Storage

Following verification of the two experimental explant systems, large scale embryo isolation experiments were set up to collect the relevant plant materials required for RNA extraction. Starting material was collected over several different isolation experiments as described in 3.2.1., so that a sufficient quantity could be accumulated for RNA extraction. In order to minimise sample variations between individual experiments, samples were pooled prior to RNA extraction. The total quantity of starting material (approx) for each sample can be seen from the table below:

Table 2: Net weight of starting materials collected for RNA extraction.

Timepoint	Bobwhite Fresh weight (mg)	Savannah Fresh weight (mg)
0	500	500
1	200	250
2	250	200
3	300	350
4	500	500

3.3.2. Morphology of Somatic Embryos at Harvested Timepoints

IEs were isolated and harvested for each cultivar at the 5 time-points as described in materials and methods. The morphological characteristics for the embryogenic cultivar BW at these 5 developmental timepoints (0-4) can be seen in figure 8, A-E respectively. Although five developmental timepoints were chosen to analyse the key stages of SE in wheat over an 8-day period, it should be remembered that the events that characterise SE occur as a continuous developmental process.

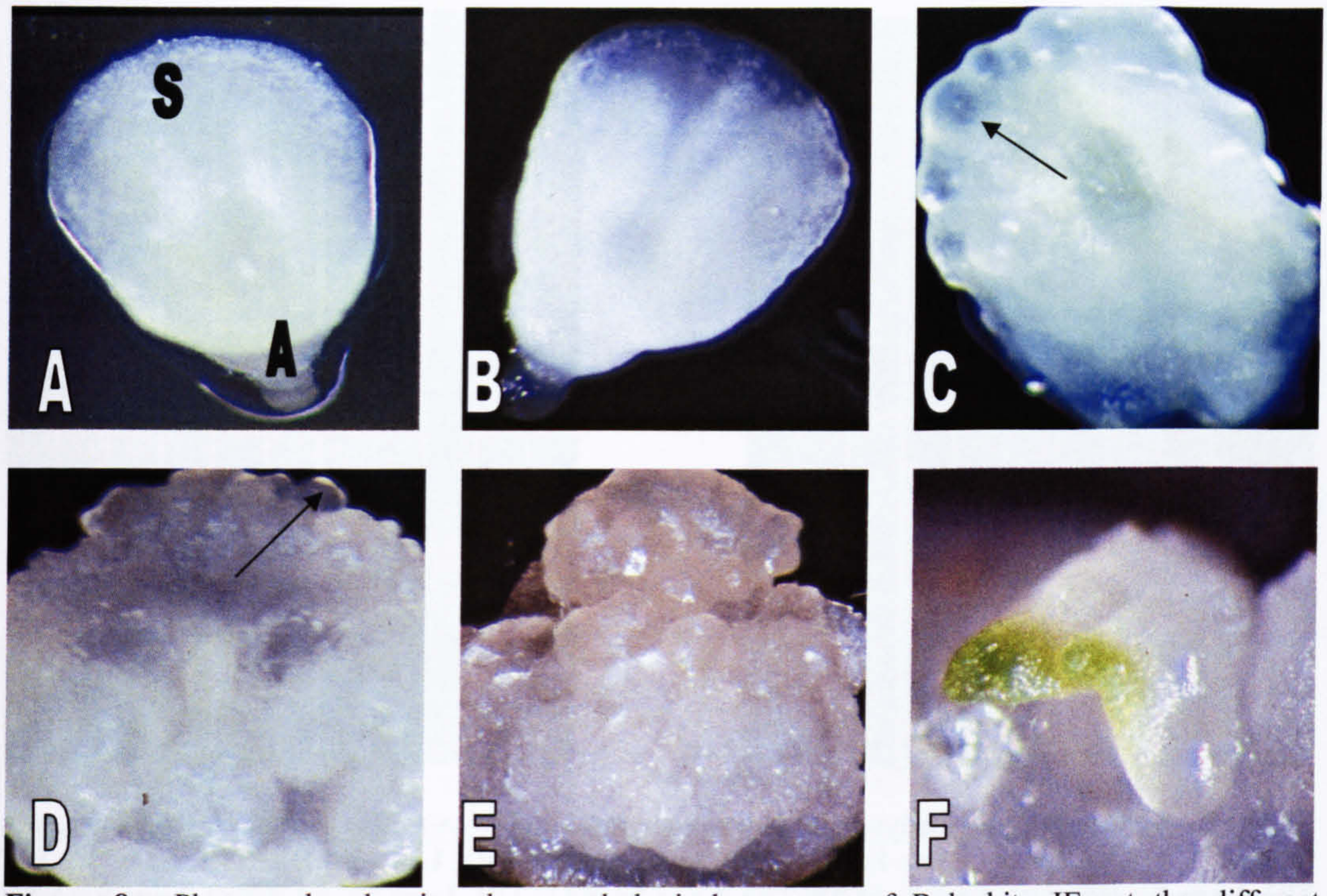


Figure 8. Photographs showing the morphological response of Bobwhite IEs at the different timepoints. **NB**. Figures not to scale **(A) Timepoint Zero:** A typical, freshly excised embryo. The embryo is excised onto medium with the scutellum, **S**, facing away from the medium and the axis, **A**, facing into the medium. **(B) Timepoint One:** After 7 hrs treatment with 2,4-D there is little or no observable change in the morphology of the embryo. **(C) Timepoint Two:** After four days treatment with 2,4-D SE has been undoubtedly initiated and globular, embryo-like structures can be observed from the edges of the scutellum (black arrow). **(D) Timepoint Three:** After 7 days treatment with 2,4-D the whole area of the scutellum has callused and multiple, somatic embryos can be observed (black arrow). **(E) Timepoint Four:** There is no real observable difference 24 hrs after the 7-day callus is taken off 2,4,-D medium but if the calli are placed under light, differentiating, photosynthesising shoot structures can be seen to emerge from the somatic embryos. (F)

3.3.3. Agilent 2100 Bioanalyser

Total and mRNA was isolated for each timepoint for both BW and SAV and run on the Agilent 2100 Bioanalyser as described in Materials and Methods (3.2.3. & 3.2.4.) The analysis of RNA quality using the Agilent 2100 Bioanalyser produced two outputs: one in a gel image format and the other in a fluorescence-scan image format.

3.3.3.1. Total and Messenger RNA Gel Image Formats

The gel image formats for each of the pooled total RNA and mRNA samples can be seen in figure 9 and an example of the fluorescence-scan images taken for each total and mRNA sample is shown in Figs. 10 & 11

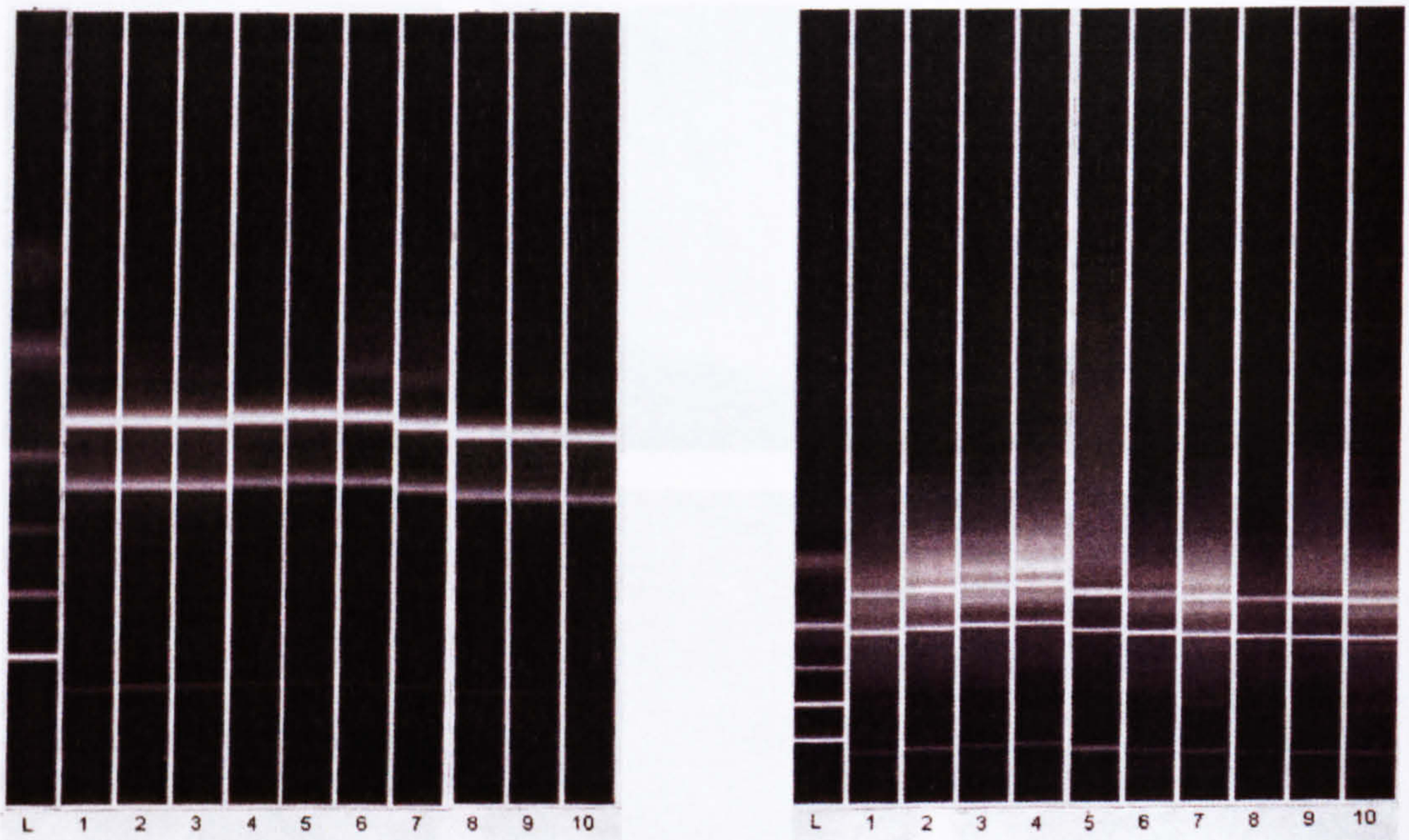
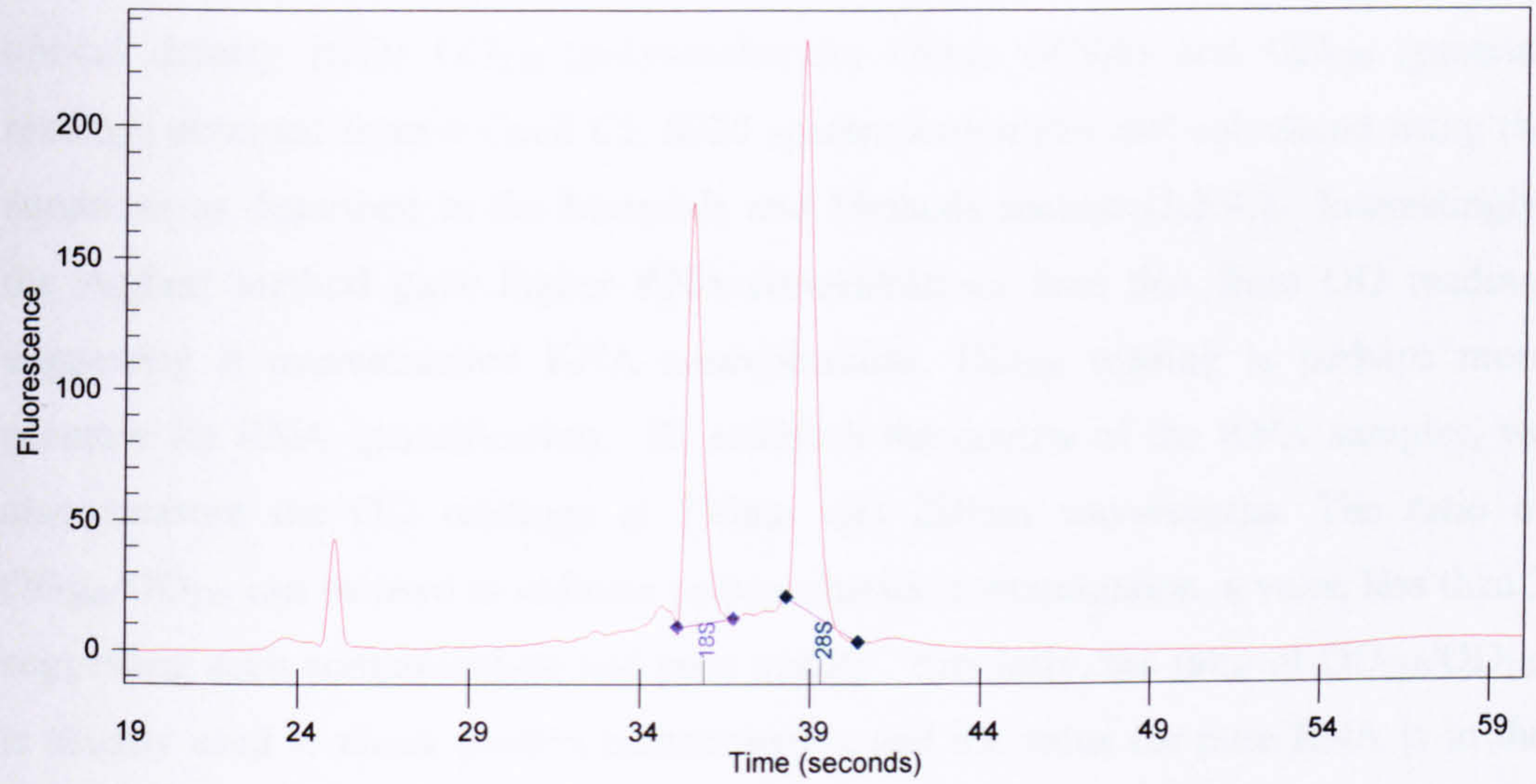


Figure 9. Agilent 2100 Bioanalyser gel format showing Total RNA (left) and mRNA (right) for each sample. **L** = RNA ladder, **Lane 1** = BW TP0, **Lane 2** = BW TP1, **Lane 3** = BW TP2, **Lane 4** = BW TP3, **Lane 5** = BW TP4, **Lane 6** = SAV TP0, **Lane 7** = SAV TP1, **Lane 8** = SAV TP2, **Lane 9** = SAV TP3, **Lane 10** =SAV TP4. The sharpness of the ribosomal RNA bands indicate that the isolated RNA is intact. The mRNA preparations also have RNAs of various molecular weight which appear as bands in the mRNA gel image due to the enrichment of the mRNA populations.

3.3.3.2. Total RNA Fluorescent Scan Images

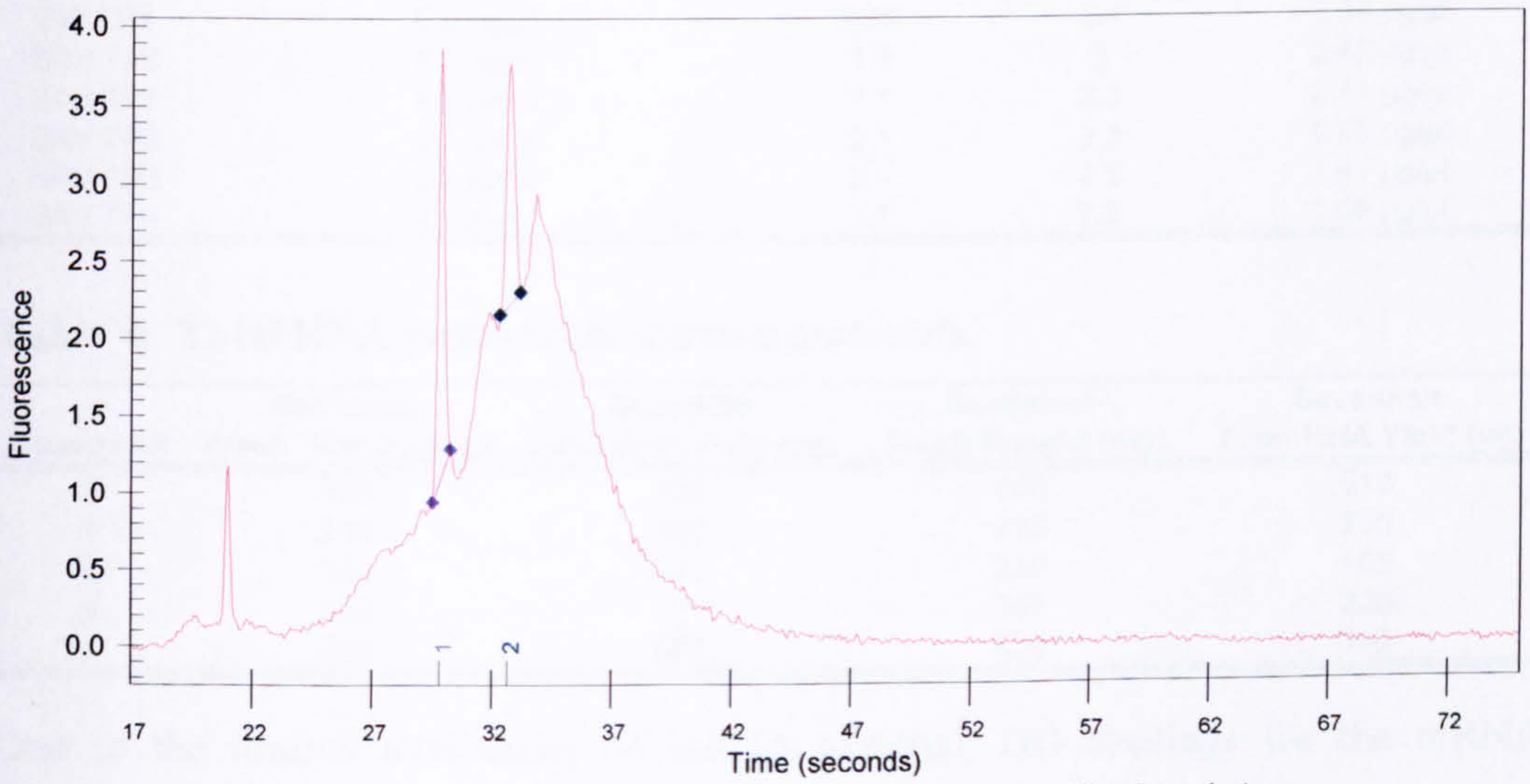


Fragment	Name	Start Time (secs)	End Time (secs)	Area	% of total Area
1	18S	35.10	36.75	239.90	25.64
2	26S	38.30	40.40	303.91	32.48

Corrected RNA Area	935.62
RNA Concentration (ng/µl)	2,233.99
rRNA Ratio [26S / 18S]	1.27

Figure 10. Agilent 2100 Bioanalyser trace showing Total RNA fluorescence Vs time (seconds) for Bobwhite TP2 (4Days 2,4-D). Ribosomal peaks are underscored in blue (18S) and green (26S).

3.3.3.3. mRNA Fluorescent Scan Images



Fragment	Name	Start_Time(secs)	End_Time(secs)	Area	%_of_total_Area
1	rRNA	29.45	30.20	2.86	4.29
2	rRNA	32.25	33.10	1.96	2.94

Corrected RNA Area	66.60
RNA Concentration (ng/µl)	273.41
rRNA Contamination (%)	7.22

Figure 11. Agilent 2100 Bioanalyser trace showing mRNA fluorescence Vs time (seconds) for Bobwhite TP1 (7Hrs 2,4-D). Ribosomal peaks are underscored in blue (1) (18S) and green (2) (26S).

3.3.4. OD Reading

Total RNA concentration and purity values for each sample were determined using optical density (OD) OD_{230} (polysaccharide) OD_{260} (RNA) and OD_{280} (protein) readings obtained from a Cecil CE 8020 spectrophotometer and calculated using the equations as described in the Materials and Methods section (3.2.4.). Interestingly, the Agilent method gave higher RNA concentrations than that from OD reading, suggesting it overestimated RNA concentrations. OD_{260} reading is perhaps more accurate for RNA quantification. To establish the quality of the RNA samples, we also measure the OD readings at 230nm and 280nm wavelengths. The ratio of OD_{260}/OD_{230} can be used to indicate polysaccharide contamination, a value less than 2 suggesting such contamination and poor quality. Similarly, the ratio of OD_{260}/OD_{280} is usually used to check protein contamination and the value for pure RNA is in the region of 1.9-2.3. Table 3 shows that all the samples are good quality, with the OD_{260}/OD_{280} ratio above 2 and OD_{260}/OD_{280} ratio between 1.9-2.3.

Table 3: Comparison of OD and Agilent 2100 Bioanalyser readings.

Time Point	Total RNA Concentration (OD 260)	OD260/OD280	OD260/OD230	Bioanalyser Results
BW TP0	5.5 µg/µl	2	2.2	2.91 µg/µl
BW TP1	5.2 µg/µl	2.1	2.3	2.74 µg/µl
BW TP2	4.9 µg/µl	2.1	2.3	2.23 µg/µl
BW TP3	2.7 µg/µl	2.1	2.3	1.55 µg/µl
BW TP4	6.9 µg/µl	2.25	2.4	3.34 µg/µl
SAV TP0	5.1 µg/µl	1.9	2	2.46 µg/µl
SAV TP1	4.6 µg/µl	2.1	2.3	2.11 µg/µl
SAV TP2	2.1 µg/µl	2.1	2.3	0.85 µg/µl
SAV TP3	3.2 µg/µl	2.1	2.3	1.61 µg/µl
SAV TP4	3.6 µg/µl	2.1	2.3	1.66 µg/µl

Table 4: Total RNA yields from starting materials.

Timepoint	Bobwhite Fresh Weight (mg)	Bobwhite Total RNA Yield (µg)	Savannah Fresh Weight (mg)	Savannah Total RNA Yield (µg)
0	500	550	500	510
1	200	260	250	230
2	250	245	200	105
3	300	270	350	320
4	500	680	500	360

Due to the limited availability of mRNA material, OD readings for the mRNA samples was not a feasible option for assessing mRNA concentration and quality. Although the Agilent 2100 Bioanalyser provides a good indication of mRNA quality and a reasonable estimate of concentration, it is not sufficiently accurate to determine the RNA concentrations required for GEM hybridisations as mRNA needs to be supplied at a consistent and accurate concentration of 50ng/µl for each sample.

3.3.5. Ribogreen Assay – mRNA Concentration

The Ribogreen fluorescence assay was used to accurately quantify mRNA concentration as described in Materials and Methods 3.2.4.3. The Tecan fluorescence scanning results of the dilution series and standards were presented in an excel spreadsheet and the readings used to generate an individual standard curve from which sample concentrations could be calculated (Fig. 12.).

	1	2	3	4	5	6	7	8	9
A	37820	18983	8618	3554	1467	810	71	25769	3780
B	36666	19085	8620	3655	1746	861	77	24325	4475

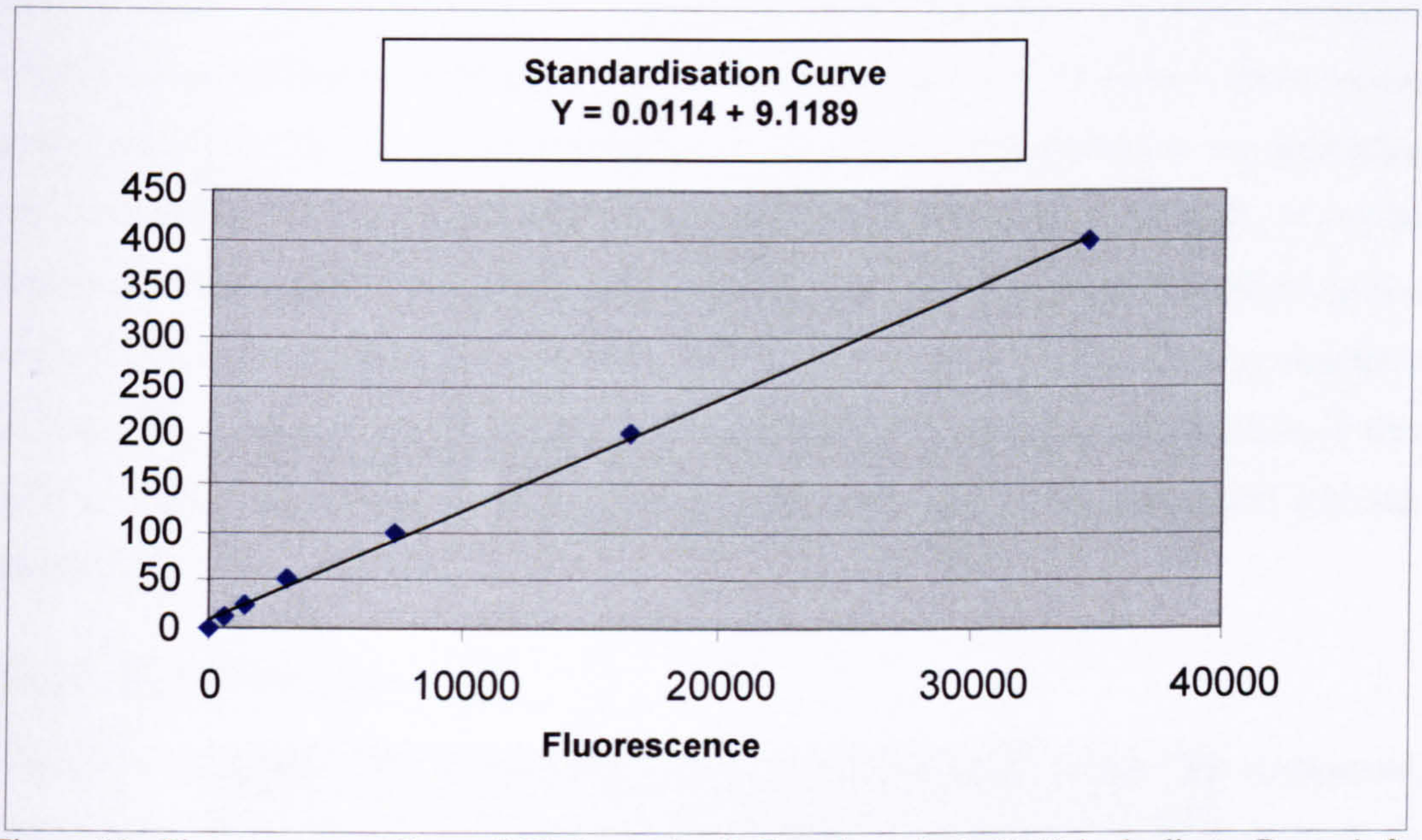


Figure 12. Fluorescence readings of dilution series (1-7) and standards (8+9) in duplicate: Rows A & B and the standardisation curve derived from the RNA dilution series and standards plotting RNA concentration (ng) against fluorescence readings which were used to calculate mRNA concentration.

A reading of each sample was then carried out in duplicate and an average taken to calculate mRNA concentration using the equation derived from the standardisation curve. For example, SAV TP0 was calculated from Ribogreen™ readings as follows:

3474 + 3810/2 = 3642 x 0.0114 + 9.1189 = 50.64 µg/µl

Each sample was submitted at 50ng/µl following concentration via glycogen precipitation and/or dilution with the appropriate amount of TE buffer.

3.4. Discussion

3.4.1. Microarray Hybridisation Strategy

The microarray hybridisation strategy was designed to identify genes involved in 1) the de-differentiation of somatic cells in response to the application of exogenous auxin (2,4-D), 2) the initiation and development of somatic embryos (SEs) as a result of continued high auxin exposure and 3) the early differentiation response of SEs on removal of exogenous 2,4-D. It can be hypothesised that auxin concentration gradients are responsible for these observations. That is that the shock treatment of somatic tissue with high auxin, or the perception of high auxin levels, somehow results in a signal being transduced via cellular signalling mechanisms to the nucleus. This then results in the genetic reprogramming of the cell and altered gene expression that causes the response observed as callus and somatic embryo formation. It can be argued that high auxin levels are responsible for not only the initial de-differentiation and development response associated with SE but also for arresting the development of the early SEs as seen in carrots and wheat. Only on removal of the auxin is this second response initiated, resulting in the re-differentiation of SEs associated with the later part of SE.

3.4.2. Morphology of SEs

Figure 8. illustrates the morphological appearance of each of the developmental timepoints that were chosen in order to test these hypotheses. It is important to clearly define each of the stages in terms of both the length of time cultured and the morphology of the developing IEs and calli. This is because the samples at each timepoint as well as the samples between the two experimental cultivars at each respective timepoint need to be as consistent and uniform as possible in order to generate a set of RNA populations which can reliably demonstrate the differences in differential expression between each other.

3.4.3. RNA Quality and Concentration

Figures 9, 10 and 11 illustrate that the quality of the extracted RNA from each of the sample populations is of an appropriate standard for successful hybridisations. The RNA trace exhibits a typically good quality total RNA profile. The sharpness of the ribosomal RNA bands (Fig. 9) and the pronounced ribosomal peaks (Fig. 10.) indicate

that there is no observable degradation of the isolated RNA. It can also be seen that the mRNA preparations have, in addition to ribosomal RNA, RNAs of various molecular weight. This can be seen from the mRNA scan images (Fig. 11) where a gentle peak, between approximately 25 and 40 secs, runs alongside the steep ribosomal RNA peaks and appear as bands in the mRNA gel image due to the enrichment of the mRNA populations. Furthermore, OD readings taken of the total RNA samples show that RNA purity falls within the expected range indicating that there is little or no contamination of each of the samples. In addition the Ribogreen™ assay allowed for the accurate quantification of mRNA concentration, thereby ensuring the provision of consistently concentrated and pure RNA samples for the microarray hybridisations.

Chapter Four

Incyte Genomics Gene Expression Microarray (GEM™) Hybridisation Experiments

4 Chapter Four - Incyte Genomics Gene Expression Microarray (GEM™) Hybridisation Experiments

4.1. Introduction

The ever increasing rate at which genomes are being sequenced and their genetic sequences are being deposited in international repositories has attracted a lot of attention to the field of functional genomics, the assigning of biological function to DNA sequences (Yang and Speed, 2002). However, the biological functions of most of these genes remain largely unknown or have been predicted through their homology with genes of better known function. There are a number of ways in which the determination of the function of these genes can be carried out including protein prediction, homology searching and expression analysis. Since their conception in the mid-1990s high density oligonucleotide or cDNA microarrays have emerged as one of the most powerful and versatile tools for genomics and their use has spread rapidly through the research community (Lockhart and Winzeler, 2000). One of the most important applications for microarrays is the monitoring of gene expression. The collection of genes that are transcribed from genomic DNA, sometimes referred to as the "transcriptome" is a major determinant of cellular phenotype and function. Unlike the genome, the transcriptome is a highly dynamic environment that can change rapidly and dramatically in response to environmental conditions or during events such as cell division (Lockhart and Winzeler, 2000). Knowing when and where a gene is expressed and to what extent, provides an essential understanding of the function of genes and the biological role and activity of their encoded protein. Microarrays allow researchers to ask big questions and the results often generate large quantities of data. Examining individual arrays and identifying which changes are significant provides one of the main difficulties associated with the application of this technology.

4.2. Materials and Methods

4.2.1. Incyte Genomics cDNA Library Construction (Phytoseq)

The cDNA libraries used in the construction of the wheat GEM slides were generated using proprietary methods developed by Incyte Genomics not all of which have been

fully disclosed. The principles, however, were developed from the Soares method (Soares and Bonaldo, 1998) with some modifications. Briefly, the Incyte cDNA clones, Phytoseq, were synthesized from pooled mRNAs from a number of different developmental stages and tissues. Expressed sequence tags (ESTs) selected for the Phytoseq libraries were all 3' long read ESTs to ensure that ESTs derived from a common mRNA always overlapped and were allocated to the same "gene bin". The cDNA clones were cloned into the pSport1 (Invitrogen™) vector via EcoR1-Not1 direct cloning. The libraries were normalised using a method modified from Bento Soares (Soares and Bonaldo, 1998). Selected clones were sequenced to monitor the level of elimination of the abundant clones after each round of subtractive hybridisation. Further hybridisation cycles were performed until the required level of depletion was achieved. An outline of this strategy can be seen in Fig.13.

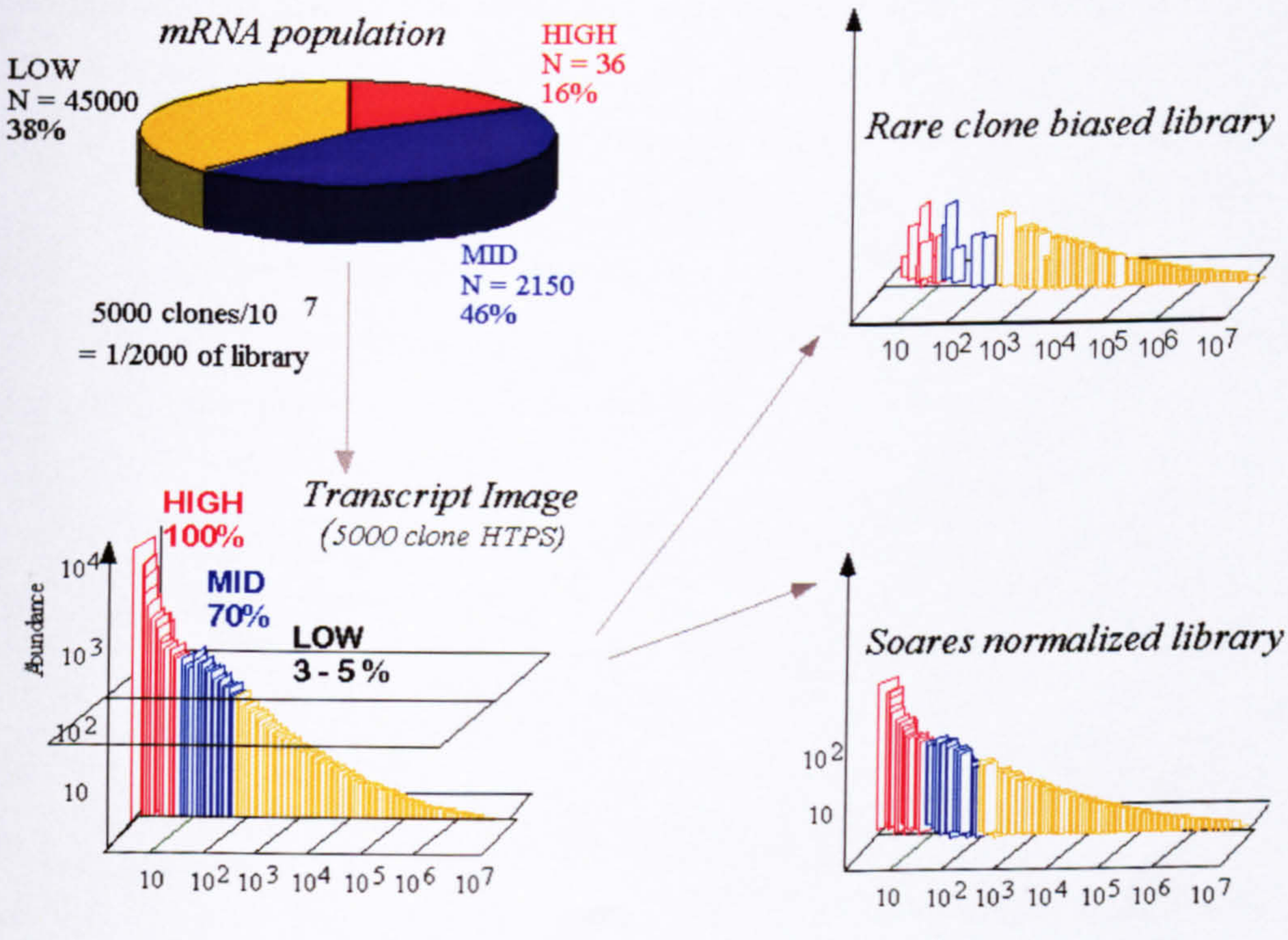


Figure 13. The relative distribution of high, middle and low abundance genes in a typical cell and the subsequent distribution of transcripts following Soares normalization. (Diagram extracted from Incyte's literature)

Approximately 10,000 wheat gene sequences are present on the wheat GEM2 chip that are derived from pooled RNA libraries biased towards tissue sources representing fertilisation and embryogenesis such as pollen, embryo sacs, pericarp, whole

inflorescences, whole caryopses, germination and floral development, as well as from vegetative tissues such as stems and leaves. Wheat consists of 3 sets of genomes and is thought to contain between 75,000 and 150,000 genes. Therefore, a vast majority of wheat genes are not represented on the wheat GEM slide. However, by using libraries relevant to embryogenesis that have been normalised to allow for rare transcripts, the chances of omitting a key gene involved in SE are markedly reduced.

4.2.2. GEM™ Design and Construction

Following cDNA library construction and normalisation, clone inserts between 500 and 5,000 bases in length were deposited onto a glass substrate in an array format. This was performed using a robotic device that draws the PCR product from a set of 96 well plates and spots each sample on to a glass slide. The PCR products were then chemically bonded onto the glass substrate to give the GEM™ chip. Each cDNA fragment mounted to the substrate represents a clone and each clone occupies a unique location called an element, which are spaced between 100-200μm apart. For each hybridisation, mRNA was extracted from the two cell samples to be studied and used to generate two differently coloured fluorescent cDNA probes (Cy3&Cy5) by reverse labelling (4.2.3.).

4.2.3. Labelling of RNA Probes and GEM™ Hybridisations

The labelling reaction was carried out using the Incyte fluorescent RT labelling kit, resulting in fluorescent-labelled single strand cDNA containing modified nucleotides with either the Cy3 or Cy5 fluorochromes (for each comparison, one probe was labelled with Cy3 and the other with Cy5). The unlabelled RNA is then hydrolysed and removed. The two differently labelled, comparative probes were then mixed and competitively hybridised with the GEM™ slides. Following hybridisation, the slides were washed under stringent conditions and scanned with a laser that both excites and detects the fluorescence from the Cy3 and Cy5 molecules at each element. The intensity of the fluorescence corresponds to the amount of bound probe and therefore the expression level of that particular cDNA clone, while the ratio of the intensities between the two different fluorochromes provides a quantitative measurement of the relative gene expression levels, of that clone, between the two treatments.

4.2.4. GEM™ Validation and Reproducibility

The credibility of expression data generated from the Incyte GEM™ experiments is backed up by a comprehensive statistical study (Pharmaceuticals, 1999) that measured Cy3:Cy5 ratio results to evaluate gene expression ratio reproducibility and accuracy over a range of probe and target concentrations. Using a concordance correlation coefficient (the assessment of concordance in continuous data), to evaluate the ability of an RNA sample to fall on the 45^0 line through the origin, reproducibility was calculated to be approximately 98 percent and therefore deemed highly reproducible. The limit of detection for differential expression was calculated using statistical tolerance limits for the fold change of Cy3:Cy5 ratios input at 1:1, to be plus or minus 1.74. Consequently elements with observed ratios greater than or equal to 1.8 can be deemed to be differentially expressed. To further increase confidence in the microarray results a threshold of 2 was used to identify differentially expressed genes. Another value, the limit of quantitation, was also calculated to determine the lowest differential expression value that can be detected with a high degree (99 percent) of confidence for low signal values. This was found to be approximately plus or minus 3.8 and therefore low ratios of differential expression (between 2 and 3.8) should be interpreted with caution for elements with low signal values. Finally, by comparing observed with expected differential expression ratios of RNA samples input at varying Cy3:Cy5 ratios, it was shown that differential expression levels tend to be increasingly underestimated at increasing differential expression values. Taken together the results of these analyses can help define the operational limitations of an Incyte GEM™ experiment and allow for an understanding of the accuracy and reliability of the results generated from the GEM™ experiments during data analysis.

4.2.5. Incyte GEMTools™ Element Validation

Incyte's GEMTools™ expression analysis software provides a user-friendly interface to help analyse & track the vast amounts of gene expression data that is generated from the GEM experiments. GEMTools™ can be used to quantify and graphically illustrate gene expression data in a variety of ways. The results of each GEM™ hybridisation are ensured by a set of 96 control transcripts that are included on every GEM™ microarray. These include a complex target of house keeping genes and blanks for assessing background effects to enable many useful quality control parameters, such as, background level, labelling effects, cross talk and carryover

effects, that can be used to internally validate the experimental data. The results from each of the hybridisation experiments were extracted and placed into the GEMTools database using unique barcode identifiers to search and select each hybridisation experiment (using the 'Find GEM' window) so that adjustments and a range of analyses could be carried out. Specific elements were selected based on customised selection criteria to filter unwanted elements using 'Element Query Builder' window. In order to correct for biological, chemical and system signal differences between the Cy3 and Cy5 signals, the Cy5 (probe 2) signal was multiplied by a unique balance co-efficient which was calculated for each individual GEM hybridisation in order to normalise signal data relative to the expression of a set of elements on the GEM. Once balanced, an absolute differential expression value (DEV) of plus or minus 2 (limit of detection described in 4.2.4.) was used to filter out any elements whose differential expression values (DEVs) fell inside this threshold. All of the selected elements were then viewed through 'The Element Display Window', which provides further information on each element, such as their differential expression levels, annotation, signal intensity and signal/background ratios etc. Any elements with both low signal intensities and low signal to background ratios were also removed resulting in all the elements from that particular hybridisation which could be said with a high degree of confidence to be differentially expressed. These elements were highlighted, made into a 'gene group' and displayed as graphic representations (see results, 4.3.) which were carried out for each of the 10 individual hybridisations, to generate gene lists from each developmental timepoint of both experimental cultivars.

4.2.6. Experimental Screening Strategy

The gene groups created for the non-embryogenic Savannah (SAV) control hybridisations, acted as a control set of gene 'hits' that were used to subtract differentially expressed gene elements common to both sets of hybridisations, leaving only genes that were uniquely expressed in embryogenic, Bobwhite (BW) tissue samples (Fig. 14). Of these elements, those with the highest DEVs were preferentially selected as candidates for further analysis. Additional selective criteria were applied based on each element's annotation as well as their expression profiles across the developmental time series experiments. Due to the design of the hybridisations (as discussed in 3.2.1), individual GEM hybridisations representing the different stages of somatic embryo development could also be assigned to a 'GEM

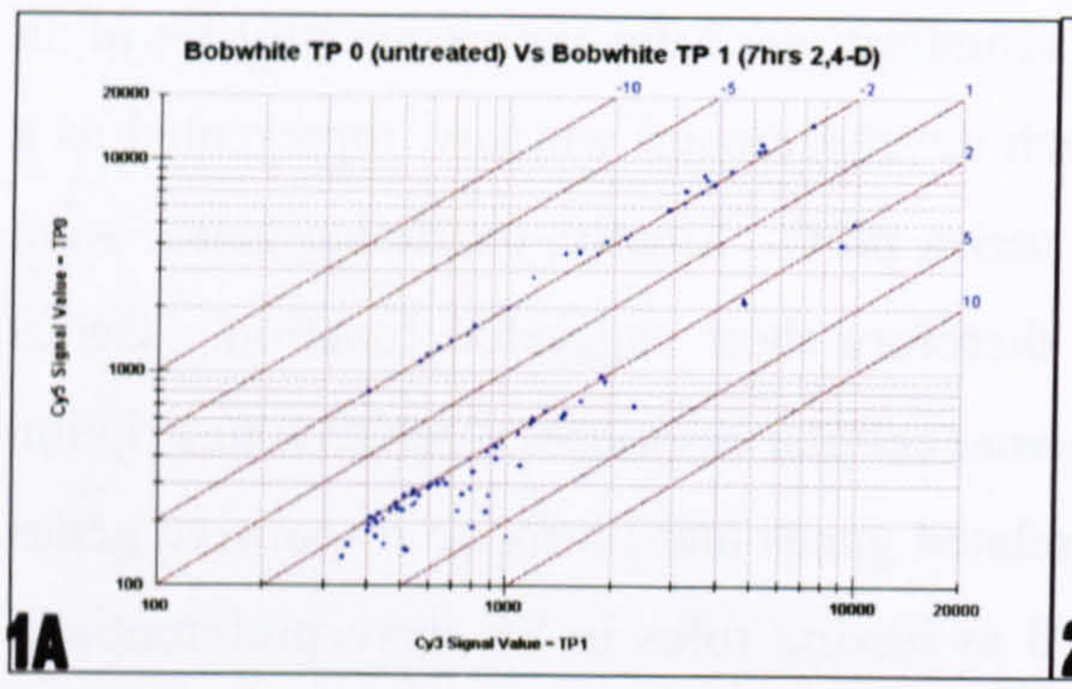
series' experiment. This allowed for the visualisation of the expression profiles of an element or group of elements through each developmental window, represented as a series of line graphs known as a 'GEM series plot'. Finally, candidate genes were selected based on their annotation and therefore their suggested function. Genes known to have important roles in fundamental cellular responses, such as transcription factors, cell signalling genes, cell cycle related genes and hormone responsive genes as well as genes previously demonstrated as having roles in SE were preferentially selected. This allowed for a shortlist of genes to be chosen as candidates for further analysis based on a diverse range of selection criteria.

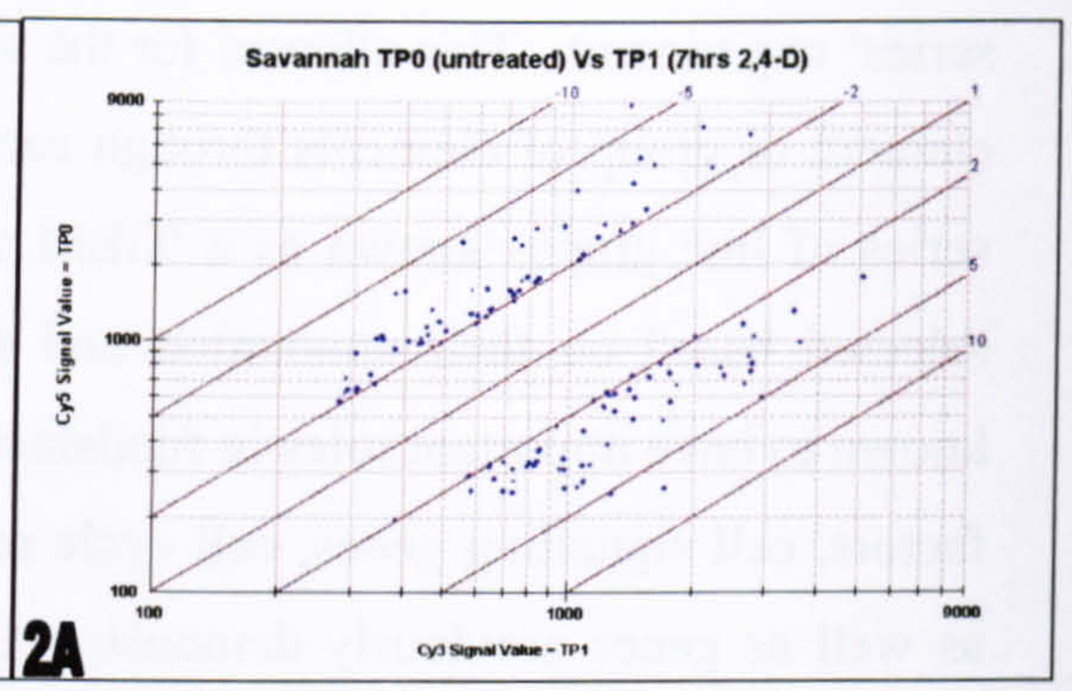
4.2.7. Bioinformatic Annotation Verification

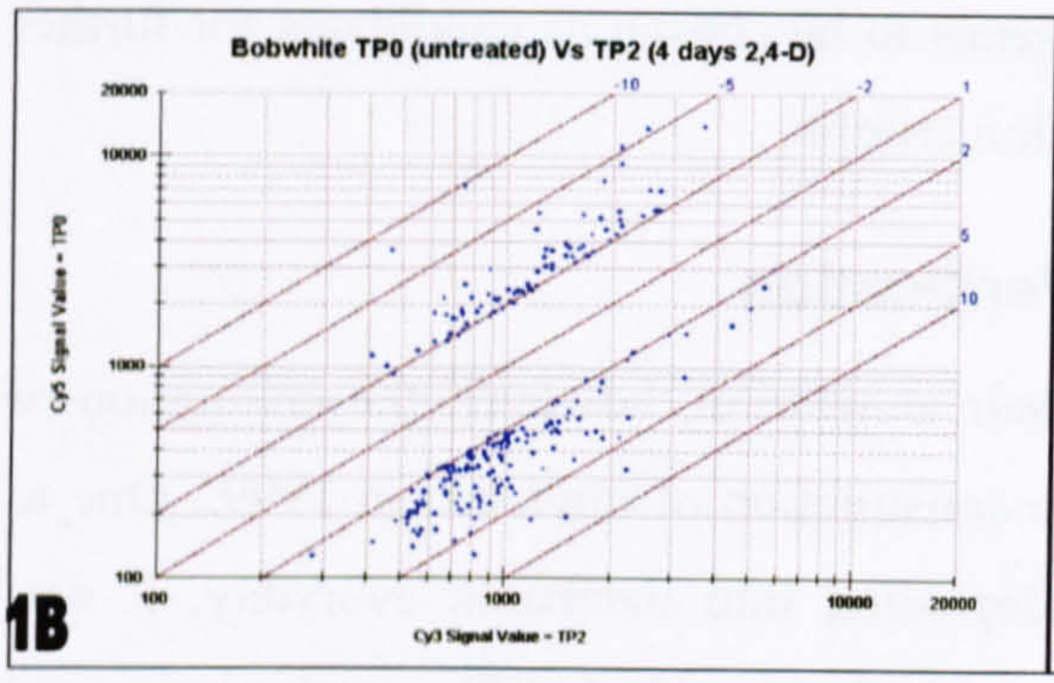
Gene elements were selected based on their annotation, however, the annotation of those elements was undertaken during the construction of Phytoseq in 1999. Due to the fact that new gene sequences are deposited into databases everyday, it was necessary to update the annotation of the positively identified, differentially expressed gene sequences. Using the elements unique identifiers to select their nucleotide sequences, each positively identified element was subject to sequence alignment searches in gbpln database using BLAST to verify whether the annotation was up to date by comparing the new top hit data with the previously annotated Incyte top hit results.

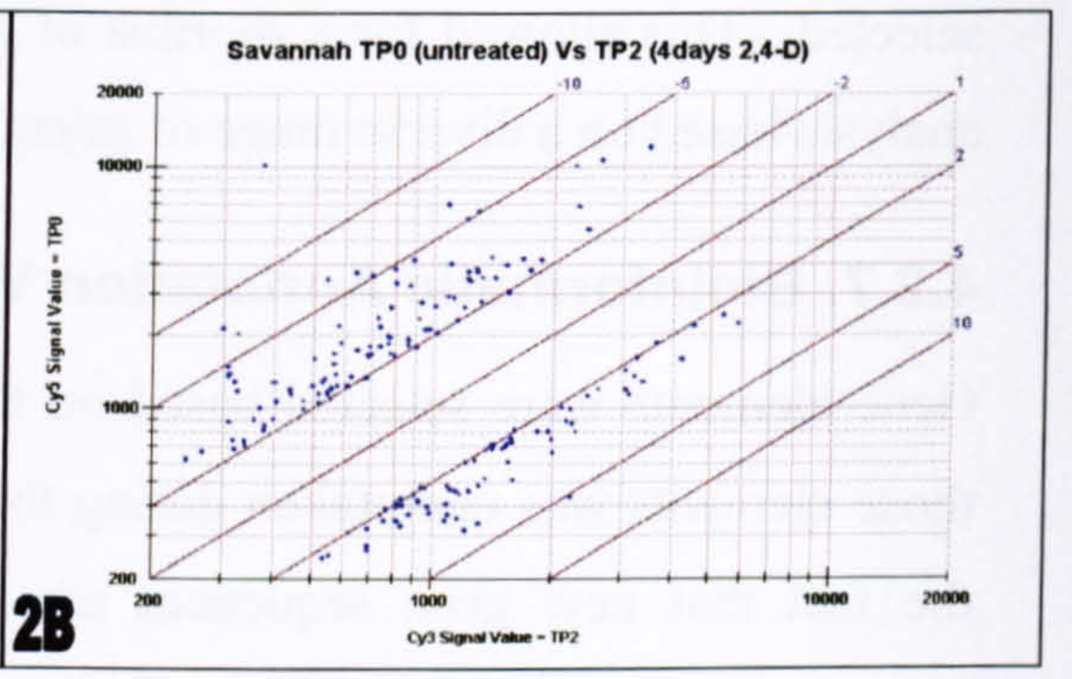
4.2.8. Clone Selection

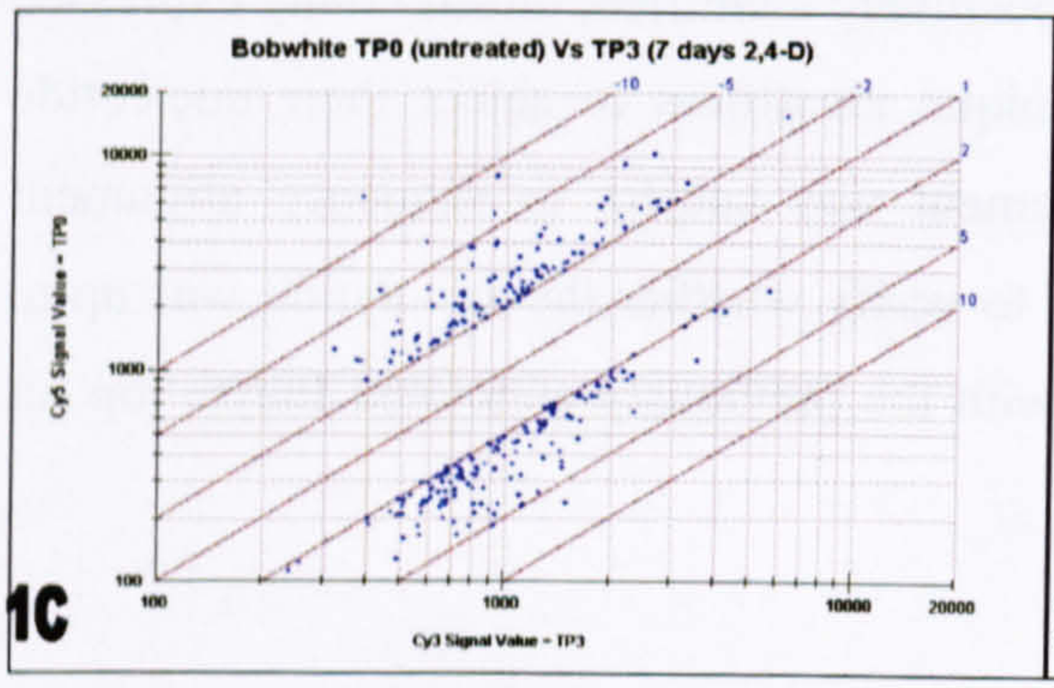
The total number of gene elements that were identified as being differentially expressed was far more than could be practically handled for subsequent analysis. Therefore, it was necessary to reduce the amount to a more manageable subset which could be further characterised by *in situ* hybridisation. Consequently, following the filtering process and selection of clones with interesting annotation and expression profiles, a shortlist of clones (Table 12) was selected for further characterisation. Once selected, the clones were picked from the library by a robotic picker so it was therefore necessary to confirm the identity of the clones through sequencing. This was carried out by preparing plasmid DNA from the supplied *E.coli* cultures using a QIAGEN liquid handling BioRobot workstation and a QIAprep 96 Turbo BioRobot kit (QIAGEN 962141) according to manufacturer's instructions. The DNA was

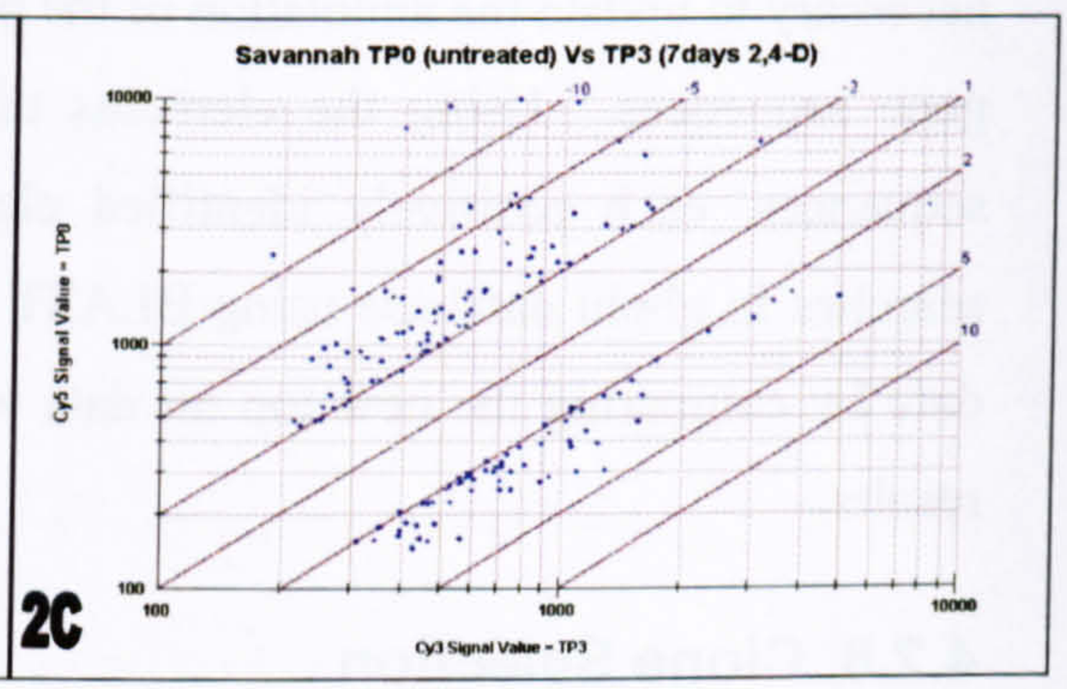


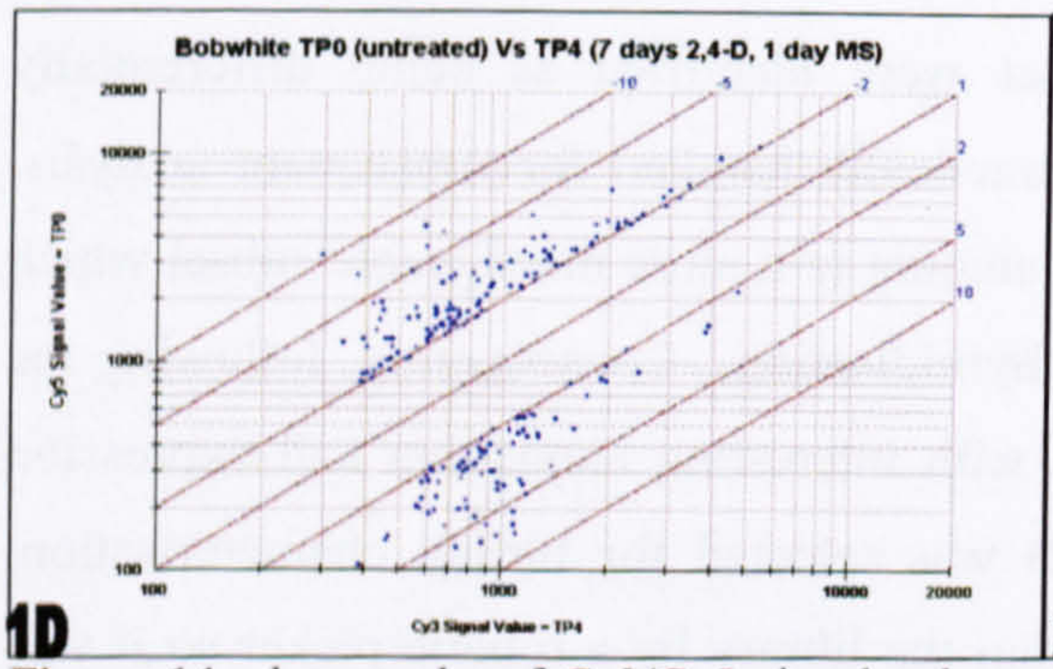












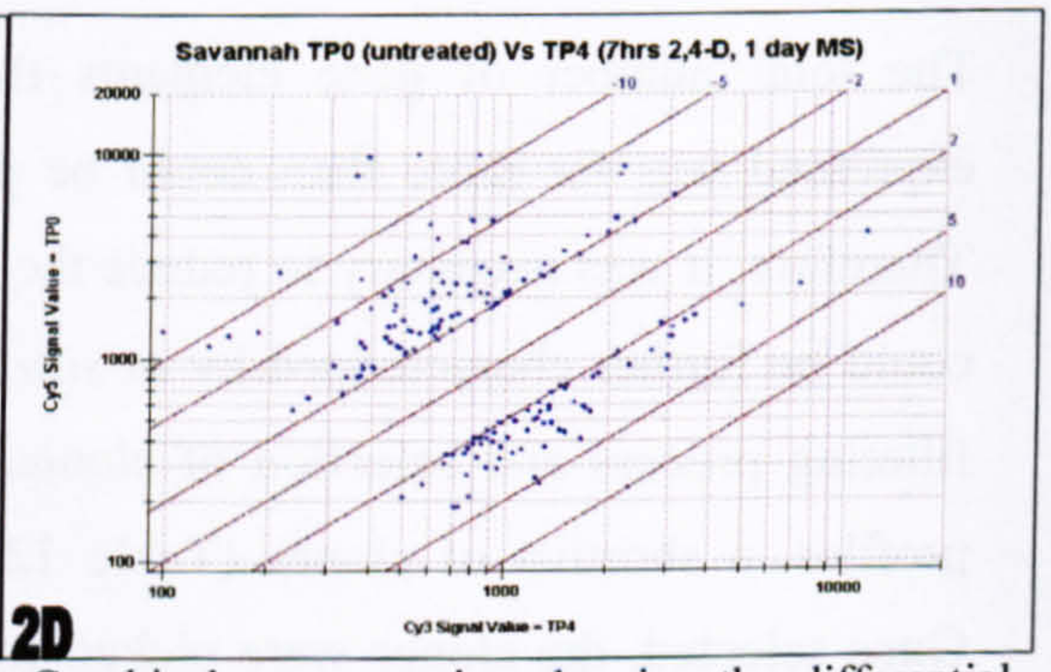


Figure 14. log graphs of Cy3/Cy5 signal values. Graphical representation showing the differential expression value results for Bobwhite (1) and Savannah (2) hybridisations. 1A=BW TP0 Vs BW TP1, 1B=BW TP0 Vs BW TP2, 1C=BW TP0 Vs BW TP3, 1D=BW TP0 Vs BW TP, 2A=SAV TP0 Vs SAV TP1, 2B=SAV TP0 Vs SAV TP2, 2C=SAV TP0 Vs SAV TP3, 2D=SAV TP0 Vs SAV TP4.

sequenced by the Syngenta Sequencing Facility at the Jealotts Hill site using an ABI PRISM377 sequencer. The resulting sequence information was aligned with the corresponding clones sequence deposited in the Phytoseq database using the software program megalign to determine whether the picked clones corresponded with the shortlisted clones that were requested.

4.3 Results

4.3.1. Incyte GEMTools™ - Cy3 Vs Cy5 Expression Plots

A series of graphical representations of the GEM hybridisation results for the embryogenic Bobwhite (BW) cultivar and the non-embryogenic Savannah (SAV) cultivar can be seen in figure 14. These representations illustrate the results of all the positively identified, differentially expressed elements for each timepoint (as described in materials and methods 4.2.5.). The red diagonal lines in each graph indicate different degrees of differential expression level values and each blue spot represents a differentially expressed element. The entire set of blue spots, or differentially expressed elements, shown in each graph make up the 'gene group' for that hybridisation. Gene groups that were made for the non-embryogenic Savannah hybridisations (Fig.14, 2A-2D) acted as a control set of gene 'hits'. These were used to subtract differentially expressed gene elements common to both sets of hybridisations, thereby leaving genes that were uniquely expressed in embryogenic, Bobwhite tissue samples. The rationale behind this filtering approach was to minimise the selection of genes that were generally associated with responses to tissue culture conditions and non-embryogenic callus induction to allow for the concentration of resources on candidate genes that were more likely to be directly associated with the onset and development of SE.

4.3.2. Incyte GEMTools™ - Subtractive Hybridisation Results

Following subtractive hybridisation analysis in excess of 500 genes were identified. It is not possible to present and discuss all of the results from these subtractive hybridisations here, therefore a selection of these gene sequences are illustrated in the following gene list representations (Tables 5-10) (for full lists see appendix CD). The genes presented in these representations were selected for based on their high DEVs and annotation of genes of potential interest.

Table 5: Selection of genes from Bobwhite Tp0 Vs Tp1 - Savannah Tp0 Vs Tp1

Incyte Clone ID	DEV	Annotated Gene Name	Accession Number	E-Value
702008465	5.9	hypothetical protein; 53156-50996	AAG50962	9.00E-13
701989150	5.9	Incyte EST	No Hit	/
701996695	4.1	Incyte EST	No Hit	/
702039485	4.3	lipid transfer protein	CAA69949	4.00E-28
702006345	4	contains similarity to DsPTP1 protein~gene_id:MQM1.1	BAB10045	3.00E-57
701965733	3.6	putative cinnamoyl-CoA reductase	AAG13987	1.00E-12
701992452	3.5	putative blue copper-binding protein	AAK16182	1.00E-17
701960043	3.4	putative NADPH-dependent oxidoreductase	AAF13742	2.00E-54
702013541	3.3	Incyte EST	No Hit	/
702001685	2.9	Incyte EST	No Hit	/
701970112	2.8	quinone oxidoreductase-like protein	CAC01713	e-111
702001259	2.8	putative nodulin protein	AAK59618	1.00E-36
701998788	2.5	Strong similarity to UDP-glucose glucosyltransferase from Arabidopsis thaliana	AAF87255	1.00E-76
701990762	2.4	HMGc2	CAA69606	7.00E-08
701961333	2.3	hypothetical protein	AAD22346	5.00E-18
702008966	2.3	unknown protein	AAF26142	5.00E-37
701967501	2.2	RNA helicase (RH16)	CAB80208	6.00E-40
702015038	2.1	Rice (O. sativa) gene for proliferating cell nuclear antigen (PCNA).	X54046	1.00E-81
702015571	2.1	putative ATP-dependent RNA helicase	AAD24375	3.00E-59
701988495	2.1	DnaJ homologue	BAB11149	3.00E-23
701998934	2.1	Hypothetical protein	AAB71963	1.00E-11
701987450	2.1	WD-40 repeat protein	AAB70241	3.00E-43
701958460	2	Nt-gh3 deduced protein	AAD32141	1.00E-11
701967005	2	protein kinase MK6	CAB82852	3.00E-12
701958329	2	contains similarity to a DNAJ-like domain	AAB61072	5.00E-06
702015586	2	serine protease-like protein	BAB10901	2.00E-34
702008794	2	Hordeum vulgare L. beta-glucosidase (BGQ60) gene, complete cds.	L41869	7.00E-23
701998659	-2	Z.mays mRNA for HMG protein.	Y08807	3.00E-64
702044675	-2	putative zinc finger transcription factor	BAB12694	1.00E-57
702002181	-2	poly(A) polymerase, putative; 41591-39333	AAG51325	2.00E-59
702003754	-2	H.vulgare mRNA for cystein proteinase inhibitor.	Y12068	4.00E-32
701992193	-2	adenine phosphoribosyltransferase	AAD55485	9.00E-09
702015062	-2	similar to wpk4 protein kinase	AAK62444	1.00E-39

BW Tp0 Vs Tp1 minus SAV Tp0 Vs Tp1 (7Hrs 2,4-D)

This subtractive hybridisation resulted in a total of 71 genes being differentially expressed between the absolute values of 5.9 and 2. The gene list representation is presented in Table 5 which was selected for based on their high DEVs and annotation of genes of potential interest. There are substantially fewer differentially expressed genes at this timepoint compared with the others, with the majority of them having a relatively low DEV at around +/-2. Furthermore, those elements that have higher DEVs also generally possess lower signal value intensities and are therefore more prone to variation. Consequently, there are only a few obvious candidate genes which can be selected for based on their high levels of differential expression. These include a hypothetical protein with a DEV of 5.9, a number of Incyte ESTs that are expressed at values of 5.9, 4.1, 3.3 and 2.9, a lipid transfer protein with a DEV of 4.3 and a dual specific phosphatase (DSPTP1) with a DEV of 4. In addition to these genes there are also differentially expressed elements present that are interesting by implication of their annotation. These genes include a high mobility group protein (HMGC2) with a DEV of 2.4, a proliferating cell nuclear antigen (PCNA) expressing at a value of 2.1 along with a DNAJ homologue and a WD-40 repeat protein that also have DEVs of 2.1. Finally there is a *Nicotiana tabacum* (Nt)-GH3 deduced protein expressed at a DEV of 2, along with a protein Kinase MK 6 and another DNA J like protein also expressed at a value of 2.

Table 6: Selection of genes from Bobwhite Tp0 Vs Tp2 - Savannah Tp0 Vs Tp2

Incyte Clone ID	DEV	Annotated Gene Name	Accession Number	E-Value
701996560	18.4	Incyte EST	No Hit	/
702040950	12.8	Incyte EST	No Hit	/
701990838	11.2	Incyte EST	No Hit	/
701988979	10.2	Incyte EST	No Hit	/
701990231	7.7	ribosomal protein-like	CAC01774	9.00E-13
701991091	6.8	Triticum aestivum cold-responsive LEA/RAB-related COR protein (Wrab19)	AF255052	5.00E-18
701990312	6.3	putative peptide transporter	AAG46154	9.00E-40
702008794	6.3	Hordeum vulgare L. beta-glucosidase (BGQ60) gene, complete cds.	L41869	7.00E-23
702001685	4.9	Incyte EST	No Hit	/
701990762	4.6	HMGc2	CAA69606	7.00E-08
702013541	4.3	Incyte EST	No Hit	/
701988082	4.3	putative ankyrin	AAK26126	8.00E-57
701991407	4.3	hypothetical protein	AAD23029	3.00E-23
701960557	4	SCARECROW transcriptional regulator-like	BAB10182	2.00E-32
701957921	4	protein phosphatase 2C	CAA05875	1.00E-06
701960043	3.9	putative NADPH-dependent oxidoreductase	AAF13742	2.00E-54
702006188	3.7	Oryza sativa embryo-specific (Ose731) mRNA, complete cds.	AF049892	0.00E+00
702006506	3.5	putative villin	CAB52460	8.00E-23
701988495	3.5	DnaJ homologue	BAB11149	3.00E-23
702014751	3.5	CRK1 protein	CAB89665	2.00E-36
701985820	3.4	CXC domain containing TSO1-like protein 1	AAF69125	7.00E-14
702015038	3.2	Rice (O. sativa) gene for proliferating cell nuclear antigen (PCNA).	X54046	1.00E-81
701992136	3.2	putative arabinose kinase	CAA74753	6.00E-39
701965681	3.1	putative cell cycle protein kinase	AAG60199	4.00E-64
701958618	2.9	Similar to DNA mismatch repair protein; T14P8.6	AAC19275	9.00E-13
702001502	2.8	Mcm6 gene product	AAF46184	3.00E-35
702012025	2.7	Oryza sativa replication protein A1 (Os-RPA1) mRNA, complete cds.	AF009179	9.00E-19
702001259	2.7	putative nodulin protein	AAK59618	1.00E-36
702006345	2.7	contains similarity to DsPTP1 protein~gene_id.MQM1.1	BAB10045	3.00E-57
701993859	2.6	abscisic acid responsive elements-binding factor	AAF27182	3.00E-49
702007940	2.6	putative MADS domain transcription factor GGM13	CAB44459	3.00E-53
701969217	2.5	mis5 protein	CAB75412	6.00E-19
702015702	2.5	Oryza sativa mRNA for C-type cyclin, complete cds.	D86925	1.00E-48
701987450	2.5	WD-40 repeat protein	AAB70241	3.00E-43
701991445	2.5	GYMNOS/PICKLE	AAF07084	3.00E-89
701958542	2.4	Zea mays tousled-like kinase 4 (MTK-4) mRNA, partial cds.	AF012889	2.00E-13
702047024	2.4	protein synthesis initiation factor eIF2 gamma	AAK29674	4.00E-64
701990064	2.3	Triticum aestivum ABA-inducible protein WRAB1 mRNA, complete cds.	AF139915	2.00E-46
702036030	2.3	chromosome assembly protein homolog	BAB11491	3.00E-35
702037288	2.3	WD-repeat protein-like protein	CAB38785	9.00E-28
701993623	2.2	Hordeum vulgare rubisco activase genes,complete cds.	M55449	2.00E-37
701997153	2.2	DNA replication factor MCM6	AAD32858	2.00E-28
701966862	2.2	structural maintenance of chromosomes (SMC)-like protein	CAB77587	8.00E-49
701995510	2.2	Dactylis glomerata somatic embryogenesis related protein mRNA	AY011122	2.00E-73
701996158	2	Ser-Thr protein kinase-like protein	BAB02869	0.00E+00
701965575	2	Zea mays eukaryotic translation initiation factor 3 large subunit mRNA	AF073329	1.00E-08
701993777	2	contains similarity to fertilization-independent seed 2 protein~gene	BAA97387	3.00E-25
702035890	-2.5	T.aestivum mRNA for peptidylprolyl isomerase.	Y07636	0.00E+00
701992258	-2.5	Hordeum vulgare cv. Haisa mRNA for MCB1 protein.	AJ303354	2.00E-17
701993242	-2.5	Saccharum officinarum pyruvate orthophosphate dikinase (PPDK) mRNA	AF194026	1.00E-43
702046792	-2.1	Rice lip19 mRNA for basic/leucine zipper protein.	X57325	4.00E-40
701702096	-2.1	Triticum aestivum gamma-type tonoplast intrinsic protein mRNA, complete cds.	U86762	1.00E-121
701997130	-2.1	Hordeum vulgare seed imbitition protein (Sip1) gene, complete cds.	M77475	8.00E-40
701997623	-2	ethylene-responsive transcriptional coactivator-like protein	BAB01997	4.00E-33

4.3.2.2. BW Tp0 Vs Tp2 minus SAV Tp0 Vs Tp2 (4Ds 2,4-D)

In contrast to the first hybridisation there are a larger number of elements with higher DEVs, which is to be expected as the starting materials are morphologically more distant. There are a total of 145 genes that are differentially expressed between the absolute values of 18.4 and 2. The gene list representation presented in Table 6, which was selected for based on their high DEVs and annotation of genes of potential interest, show that the highest DEVs are a number of Incyte ESTs unique to this timepoint with values of 18.4, 12.8, 11.2 and 10.2. Other high DEVs with interesting annotation include a late embryogenesis abundant responsive to ABA (LEA/RAB) protein and a putative peptide transporter at 6.8, a putative ankyrin at 4.3 and a SCARECROW (SCR) transcription factor and a protein phosphatase 2C (PP2C) expressed at 4. Both the DNAJ and PCNA genes are continued to be expressed at a higher DEV than during the previous hybridisation (3.5 and 3.2 respectively) along with the putative nodulin protein at 2.7 and two Incyte EST genes (IDs, 702001685 + 702013541) with DEVs of 4.9 and 4.3 that were identified in the first hybridisation with DEVs of 2.9 and 3.3 respectively. Other differentially expressed genes of interest include, a rice embryo-specific (Ose731) gene expressed at 3.7, a CRK1 protein at 3.5, a TSO1-like protein at 3.4, a putative arabinose kinase at 3.2, and an ABREBF and MADS BOX 13 gene at 2.6. There is also a C-type cyclin, a WD40 repeat and a Gymnos/Pickle gene all expressed at a value of 2.5. Genes that showed lower DEVs of interest include a tousled like kinase expressing at 2.4, an ABA inducible WRAB gene at 2.3, a SE Related protein at 2.2 and a serine/threonine protein kinase and fertilization independent seed like protein expressed at a value of 2. Finally, there is the down-regulation of a rice basic leucine zipper protein and a seed imbibition protein at a value of -2.1 and an ethylene responsive co-activator like protein at -2.

Table 7: Selection of genes from Bobwhite Tp0 Vs Tp3 - Savannah Tp0 Vs Tp3

Incyte Clone ID	DEV	Annotated Gene Name	Accession Number	E-Value
702040995	12	Incyte EST	No Hit	/
702041163	7.7	Incyte EST	No Hit	/
701993516	6.4	Festuca pratensis mRNA for beta expansin B3 (expB3 gene).	AJ295942	2.00E-24
701961028	5	Incyte EST	No Hit	/
701992432	4.9	Incyte EST	No Hit	/
701967564	4.9	Wheat alpha-Amy2/54 gene 5'-region.	X13580	3.00E-37
702004325	4.8	Incyte EST	No Hit	/
702015004	4.3	contains similarity to limonene cyclase~gene_id:MED24.9	BAB08605	6.00E-20
701967967	3.8	putative cytosine-5 DNA methyltransferase	AAC16389	2.00E-56
701990762	3.7	HMGc2	CAA69606	7.00E-08
701988082	3.4	putative ankyrin	AAK26126	8.00E-57
702001259	3.4	putative nodulin protein	AAK59618	1.00E-36
702015038	3.1	Rice (O. sativa) gene for proliferating cell nuclear antigen (PCNA).	X54046	1.00E-81
701988495	2.9	DnaJ homologue	BAB11149	3.00E-23
701987450	2.9	WD-40 repeat protein	AAB70241	3.00E-43
701972138	2.8	beta-expansin	AAF72986	5.00E-65
701961741	2.5	lipid transfer protein	CAC13149	3.00E-51
701969217	2.4	mis5 protein	CAB75412	6.00E-19
702001502	2.4	Mcm6 gene product	AAF46184	4.00E-41
701969223	2.4	protein kinase	CAA06503	3.00E-09
701965681	2.4	putative cell cycle protein kinase	AAG60199	4.00E-64
702045519	2.4	DNA polymerase alpha subunit IV (primase)-like protein	BAB10663	9.00E-26
702011085	2.4	SAR DNA binding protein	BAA31260	2.00E-15
701959894	2.4	T.aestivum (Chinese spring) chi gene for endochitinase.	X76041	6.00E-22
702015571	2.3	putative ATP-dependent RNA helicase	AAD24375	3.00E-59
701960043	2.2	putative NADPH-dependent oxidoreductase	AAF13742	2.00E-54
701959805	2.2	leucine-rich receptor-like protein kinase, putative; 84911-81624	AAG51803	9.00E-18
702006429	2.2	flowering protein CONSTANS, putative; 7571-5495	AAG52620	4.00E-05
701966508	2.2	Hordeum vulgare mRNA for MADS-box protein 9 (m9 gene).	AJ249147	3.00E-07
701957921	2.2	protein phosphatase 2C	CAA05875	1.00E-06
701992136	2.1	putative arabinose kinase	CAA74753	6.00E-39
701995104	2.1	transcription factor IIA small subunit	CAA11524	5.00E-04
702000861	2.1	contains similarity to SET-domain protein~gene_id:MAC12.7	BAB11124	8.00E-12
701992254	2.1	scarecrow-like 13	AAD24411	2.00E-66
701992420	2.1	somatic embryogenesis receptor-like kinase 3	CAC37642	2.00E-03
701993859	2	abscisic acid responsive elements-binding factor	AAF27182	3.00E-49
701958329	2	contains similarity to a DNAJ-like domain	AAB61072	5.00E-06
701965327	2	IAA-amino acid hydrolase homolog ILL3	AAC31939	1.00E-52
702015076	2	calcium dependent protein kinase	AAC49405	1.00E-11
702007715	-16.8	Hordeum vulgare myo-inositol 1-phosphate synthase mRNA	AF056325	9.00E-31
701768814	-4.8	Hordeum vulgare gamma-thionin (HTH3) mRNA, complete cds.	U22951	3.00E-43
702015062	-3.9	similar to wpk4 protein kinase	AAK62444	1.00E-39
701994986	-3.7	putative Bowman Birk tryspin inhibitor	BAB55530	3.00E-75
701997130	-3.5	Hordeum vulgare seed imbitition protein (Sip1) gene, complete cds.	M77475	8.00E-40
702014249	-3	Hordeum vulgare ABA-responsive protein mRNA, complete cds.	AF026538	2.00E-36
702046792	-2.3	Rice lip19 mRNA for basic/leucine zipper protein.	X57325	4.00E-40
702014489	-2.3	Oryza sativa alpha-expansin OsEXP7 mRNA, complete cds.	AF247164	3.00E-52

4.3.2.3. BW Tp0 Vs Tp3 minus SAV Tp0 Vs Tp3 (7Ds 2,4-D)

Again there are a much larger number of differentially expressed elements than when compared with the first hybridisation with a total of 188 genes being differentially expressed between absolute values of 16.8 and 2. The gene list representation presented in Table 7 shows that the highest differentially expressed genes from this hybridisation include a myo-inositol 1-phosphate synthase at -16.8 and several Incyte ESTs unique to this timepoint at values of 12, 7.7, 5, 4.9 and 4.8. There is also the expression of a beta-expansin gene with a DEV of 6.4, an alpha amylase with a DEV of 4.9, a limonene cyclase with a DEV of 4.3 and a cytosine-5 DNA methyltransferase expressed at a value of 3.8. The continued expression of the HMGC2, PCNA and PP2C genes are detected during this hybridisation at values of 3.7, 3.1 and 2.2 respectively. In addition, both the DNAJ homologues and WD 40 repeat protein also continue to be expressed at a value of 2.9, along with the putative arabinose kinase at a value of 2.1 and the ABREBF gene at a value of 2. Similarly, the seed imbibition and rice bZip proteins continue to be down regulated at values of -3.5 and -2.3 respectively. Other genes down regulated during this timepoint include a gamma thionin (HTH3) gene with a DEV of -4.8, a wpk4 protein kinase with a DEV of -3.9, a putative Bowman Berk trypsin inhibitor with a DEV of -3.7 and an ABA responsive protein with a DEV of -3. Of additional interest is the expression of an endochitinase at a value of 2.4, the CONSTANS flowering protein and MADS-box protein both expressed at a value of 2.2 and a Suvar3-9, Enhancer-of-zeste, Trithorax (SET) domain protein along with a SCARECROW-like protein and somatic embryogenesis receptor-like kinase (SERK) all expressed at a value of 2.1. Finally, a calcium dependent protein kinase and an IAA-amino acid hydrolase are expressed with a DEV of 2.

Table 8: Selection of genes from Bobwhite Tp0 Vs Tp4 - Savannah Tp0 Vs Tp4

Incyte Clone ID	DEV	Annotated Gene Name	Accession Number	E-Value
701996691	11.6	Incyte EST	No Hit	/
702012143	11	Incyte EST	No Hit	/
701958679	7.2	Incyte EST	No Hit	/
702002595	6.6	At2g29670/T27A16.23	AAK53044	4.00E-25
701966806	6.5	unknown protein; 12187-10624	AAG52489	5.00E-37
701990087	6.4	Similar to hypothetical protein YLR002c, gb|Z7314 from S. cerevisiae.	AAC17047	2.00E-23
701993520	5.7	Incyte EST	No Hit	/
701960451	5.4	hypothetical protein; 75067-63678	AAG51799	4.00E-56
701966508	4.6	Hordeum vulgare mRNA for MADS-box protein 9 (m9 gene).	AJ249147	3.00E-07
702014751	4	CRK1 protein	CAB89665	2.00E-36
701993516	3.8	Festuca pratensis mRNA for beta expansin B3 (expB3 gene).	AJ295942	2.00E-24
701970106	3.3	Similar to membrane-associated salt-inducible protein from Nicotiana tabacum	AAC28506	1.00E-56
701975167	3	Zea mays ZmRR6 mRNA for response regulator 6, complete cds.	AB042268	8.00E-21
701998612	2.9	Nicotiana tabacum DNA-binding protein 2 (WRKY2) mRNA, complete cds.	AF096299	5.00E-14
702043443	2.9	Similar to SEC7 protein, Saccharomyces cerevisiae, PIR2:S49764	BAA82387	8.00E-57
701990762	2.8	HMGc2	CAA69606	7.00E-08
702012268	2.8	putative polygalacturonase	CAB81300	6.00E-34
702037281	2.7	Triticum aestivum heat shock protein 101 (HSP101) mRNA, HSP101-E allele	AF174433	5.00E-71
702017431	2.7	Hordeum vulgare gene for ids2, complete cds.	D15051	7.00E-17
702006345	2.6	contains similarity to DsPTP1 protein~gene_id.MQM1.1	BAB10045	3.00E-57
701968565	2.6	nucleic acid binding protein	AAD31844	3.00E-56
702035826	2.6	Triticum aestivum heat shock protein 80 mRNA, complete cds.	U55859	9.00E-61
702006433	2.6	transporter-like protein	CAB88358	3.00E-20
701988082	2.6	putative ankyrin	AAK26126	8.00E-57
701988495	2.5	DnaJ homologue	BAB11149	3.00E-23
701958329	2.5	contains similarity to a DNAJ-like domain	AAB61072	5.00E-06
702015038	2.4	Rice (O. sativa) gene for proliferating cell nuclear antigen (PCNA).	X54046	1.00E-81
701958295	2.2	Oryza sativa mRNA for Ran, complete cds.	AB015287	1.00E-35
702037649	2.2	nonspecific lipid transfer protein	AAA82182	5.00E-92
702045519	2.2	DNA polymerase alpha subunit IV (primase)-like protein	BAB10663	9.00E-26
702015571	2.2	putative ATP-dependent RNA helicase	AAD24375	3.00E-59
701969217	2.1	mis5 protein	CAB75412	6.00E-19
701992851	2.1	putative RING zinc finger protein	AAF20226	2.00E-34
702001502	2	Mcm6 gene product	AAF46184	4.00E-41
701997153	2	DNA replication factor MCM6	AAD32858	2.00E-28
702006429	2	flowering protein CONSTANS, putative; 7571-5495	AAG52620	4.00E-05
701997130	-3.3	Hordeum vulgare seed imbitition protein (Sip1) gene, complete cds.	D38089	8.00E-40
702046792	-2.5	Rice lip19 mRNA for basic/leucine zipper protein.	AJ238697	4.00E-40
701974449	-2.4	Hordeum vulgare mRNA for GPX12Hv, cultivar Haisa.	AF026538	8.00E-10
702014489	-2.4	Oryza sativa alpha-expansin OsEXP7 mRNA, complete cds.	D37945	3.00E-52
702019247	-2	VIP2 protein	AAB94598	8.00E-29
701972996	-2	polyphosphoinositide binding protein Ssh1p	CAA07563	7.00E-68

4.3.2.4. BW Tp0 Vs Tp4 minus SAV Tp0 Vs Tp4 (7Ds 2,4-D, 1D MS)

At this hybridisation a total of 137 genes are differentially expressed between the absolute values of 11.6 and 2. The gene list representation presented in Table 8 shows that the highest differentially expressed genes in this list are a number of unknown, hypothetical or Incyte ESTs. Of particular interest is the increased expression of the MADS-box gene at a value of 4.9, the CRK1 protein at a value of 4 as well as the continued expression of the beta expansin gene at a value of 3.8. Similarly, the HMGC2, DsPTP1 and putative ankyrin, DNAJ homologue and DNAJ-like, PCNA and mis5 continue to be expressed at values of 2.8, 2.6, 2.5, 2.4 and 2.1 respectively. Furthermore, the seed imbibition protein continues to be down regulated at a value of -3.3 as well as the continued down regulation of the basic/leucine zipper protein at -2.5. There are also a number of genes of interest that are uniquely expressed for this hybridisation which include a *Zea mays* response regulator at a value of 3, a SEC7 protein at 2.9, a polygalacturonase at 2.8, a heat shock protein and nucleic acid binding protein at a value of 2.6, a ran protein and RNA helicase at 2.2 as well as the down regulation of a VIP2 protein and a polyphosphoinositide binding protein both at a value of -2.

4.3.2.5. BW Tp3 (7Ds 2,4-D) Vs BW Tp4 (7Ds 2,4-D, 1D MS)

It can be seen in Table 9 that there are only a small number of gene elements that are differentially expressed between these two time points. In total, 28 genes were differentially expressed from this hybridisation between values of 8.5 and 2, with no genes being identified that show down regulation. The majority of the top expressers are putative, hypothetical, unknown or Incyte EST genes. Interestingly, the MADS domain transcription factor that was identified from the TP0vTP2 hybridisations is also expressed here and again there is the expression of 2 DNAJ-like genes. The expression of a nucleic acid binding protein that was identified from the TP0vTP4 hybridisation is also confirmed from this hybridisation with an expression value of 2. Additional genes of interest include a starch synthase expressed at a value of 3.7, two ubiquinone related genes at values of 3.6 and 3.3 and there is a TAR RNA loop binding protein expressed at a value of 2.1.

Table 9: Selection of genes from Bobwhite Tp3 Vs Tp4

Incyte Clone ID	DEV	Annotated Gene Name	Accession Number	E-Value
701967276	8.5	Incyte EST	No Hit	/
701965407	6.3	Incyte EST	No Hit	/
701965494	5.7	putative protein	CAB96849	1.00E-71
702046770	5.3	hypothetical protein~similar to Arabidopsis thaliana F20D10.300	BAB12035	4.00E-62
702004012	5.3	Incyte EST	No Hit	/
701993439	3.7	Triticum aestivum starch synthase III mRNA, complete cds.	AF258608	2.00E-30
702003794	3.6	T.aestivum nad4 gene for NADH-ubiquinone oxidoreductase chain 4.	X57164	1.00E-14
701967387	3.3	putative ubiquinone biosynthesis protein	AAD20072	5.00E-74
701961145	2.8	Similar to M. tuberculosis gene gb|Z96072 and M. leprae gene gb|400019.	AAC17039	3.00E-35
702007940	2.6	putative MADS domain transcription factor GGM13	CAB44459	3.00E-53
701971219	2.6	Contains similarity to DnaJ gene YM8520.10 gb|825566 from from S. cerevisiae	AAC27138	5.00E-24
701958329	2.5	contains similarity to a DNAJ-like domain	AAB61072	5.00E-06
701996401	2.1	TAR RNA loop binding protein	AAC50379	9.00E-25
701968565	2	nucleic acid binding protein	AAD31844	3.00E-56
701993165	2	putative glycerol-3-phosphate dehydrogenase	AAF02807	7.00E-92

Table 10: Selection of genes from Bobwhite Tp4 Vs Savannah Tp4

Incyte Clone ID	DEV	Annotated Gene Name	Accession Number	E-Value
701967692	4.5	Incyte EST		/
702035890	11.7	T.aestivum mRNA for peptidylprolyl isomerase.	Y07636	0.00E+00
702037281	11.3	Triticum aestivum heat shock protein 101 (HSP101) mRNA, HSP101-E allele,	AF174433	5.00E-71
702007884	9.1	phosphoinositide-specific phospholipase C P13	AAB03258	2.00E-49
702041871	7.9	Triticum aestivum L. (clone 1LC1) heat shock protein (hsp16.9-1LC1) mRNA,	L37071	7.00E-61
701994767	7.4	Triticum aestivum small heat shock protein HSP17.8 mRNA, complete cds.	AF350423	5.00E-20
701991275	7	Oryza sativa secretory carrier membrane protein (SC) mRNA, complete cds.	AF225922	5.00E-19
702035387	6.7	Incyte EST		/
702002904	6.4	Triticum aestivum L. (clone 10LC1) heat shock protein (hsp16.9-10LC1) mRNA,	L37061	8.00E-46
702036830	6.1	T.aestivum mRNA for heat shock protein 17.3.	X58279	1.00E-30
701968407	5.3	AT-hook DNA-binding protein (AHP1)	AAB80677	4.00E-07
702038119	5.1	Incyte EST		/
701969736	5	putative heat shock protein	AAD25656	2.00E-12
701969838	5	putative RNA binding protein	CAB61958	6.00E-34
702010150	4.6	stress-induced protein sti1-like protein	CAB78283	7.00E-90
701967687	4.4	putative ATP-dependent CLPB protein	AAF19220	5.00E-31
702017838	4	PREG-like protein	AAC32127	3.00E-10
701986547	4	Highly similar to cullin 3	AAD14503	3.00E-24
701993587	3.5	Triticum aestivum heat shock protein 70 (HSP70) mRNA, HSP70-S allele,	AF074969	1.00E-28
701999024	3.3	membrane protein	CAB87281	6.00E-40
701997262	3.1	Zea mays MFP1 attachment factor 1 (maf1) mRNA, complete cds.	AF118114	2.00E-67
701969412	2.8	amino acid transporter	BAB10054	4.00E-85
702035826	2.8	Triticum aestivum heat shock protein 80 mRNA, complete cds.	U55859	9.00E-61
701969327	2.7	Hordeum vulgare ABA-responsive protein mRNA, complete cds.	AF026538	8.00E-54
701985982	2.6	putative transposase protein	AAD55677	9.00E-70
702015426	2.5	DnaJ-like protein	AAF28382	5.00E-18
702006433	2.4	transporter-like protein	CAB88358	3.00E-20
701996422	2.4	peptidylprolyl isomerase	CAA68913	7.00E-12
701993596	2.4	Z.mays OBF1 mRNA for ocs-element binding factor 1.	X62745	1.00E-29
702046406	2.3	putative dnaJ-like protein	AAD27555	2.00E-09
701966045	2.1	cell division related protein-like	BAA98200	3.00E-42
701993859	2.1	abscisic acid responsive elements-binding factor	AAF27182	3.00E-49
702042046	2	similar to mouse Glt3 or D. malanogaster transcription factor IIB	AAH05152	1.00E-05
701965606	2	contains similarity to kinesin heavy chain~gene_id:MNA5.20	BAB11568	2.00E-28
701702096	-2.5	Triticum aestivum gamma-type tonoplast intrinsic protein mRNA,	U86762	1.00E-121
701975337	-2.4	Rice mRNA for Phospholipid transfer protein.	D10955	6.00E-54
701685747	-2.3	Hordeum vulgare low temperature-responsive RNA-binding protein (blt801)	U49482	6.00E-19
702000029	-2.2	porin-like protein	BAB08458	4.00E-11
701973302	-2	N6-DNA-methyltransferase-like protein	BAB01750	7.00E-89
701767066	-2	Barley embryo globulin (BEG1) mRNA, complete cds.	M64372	5.00E-45
702039835	-2	auxin-responsive-like protein	BAA97524	1.00E-16

4.3.2.6. SAV Tp4 Vs BW Tp4 (7Ds 2,4-D, 1D MS)

This hybridisation perhaps offers a more useful and interesting comparison of the gene expression profile at this developmental time point as it provides a direct comparison of the differences in gene expression between the embryogenic and non embryogenic experimental systems. It can be seen in Table 10 that there are significantly more differentially expressed genes identified from this hybridisation when compared with the Bobwhite TP3 v TP4 hybridisation with a total of 132 genes being differentially expressed between the absolute values of 14.2 and 2. Effectively, the DEVs from this timepoint correspond to the differential expression of genes in BW compared to SAV. The genes that showed the highest differential expression in this list were Incyte ESTs at values of 14.2, 6.7 and 5.1. It can also be seen that there are a total of 9 heat shock proteins expressed here between the values of 11.3 and 2.8. There are also a number of genes of interest with quite high expression values including a wheat mRNA for peptidylprolyl isomerase at a value of 11.7, a phospholipase expressed at 9.1, a secretory carrier membrane protein expressed at 7, an AT-hook DNA binding protein at a value of 5.3 and a PREG-like and cullin-like protein expressed at a value of 4. In addition, there are a number of differentially expressed genes that had been previously identified during the earlier hybridisations which include an ABA responsive protein, a DNAJ like protein and an ABA responsive element binding factor at values of 2.7, 2.5 and 2.1 respectively. Finally, additional genes of interest include an OCS-element binding factor at a value of 2.4 and a cell division related protein at a value of 2.1.

4.3.2.7. Summary of Differentially Expressed Genes

Subtractive microarray analysis has identified hundreds of differentially expressed genes from each of the developmental time points which can be see in table 11.

Table 11: Summary of differentially expressed genes identified during SE

GEM Hybridisation	Gene Hits	Absolute Differential Expression Value Range
BWTP0 v BWTP1 - SAVTP0 v TP1	71	5.9 to -2.3
BWTP0 v BWTP2 - SAVTP0 v TP1	145	18.4 to -5.4
BWTP0 v BWTP3 - SAVTP0 v TP1	188	12 to -16.8
BWTP0 v BWTP4 - SAVTP0 v TP1	137	11.6 to -3.6
BWTP3 v BWTP4	28	8.5 to - 2
SAVTP4 v BWTP4	132	14.2 to -3.3
Total	**701**	**18.4 to -16.8**

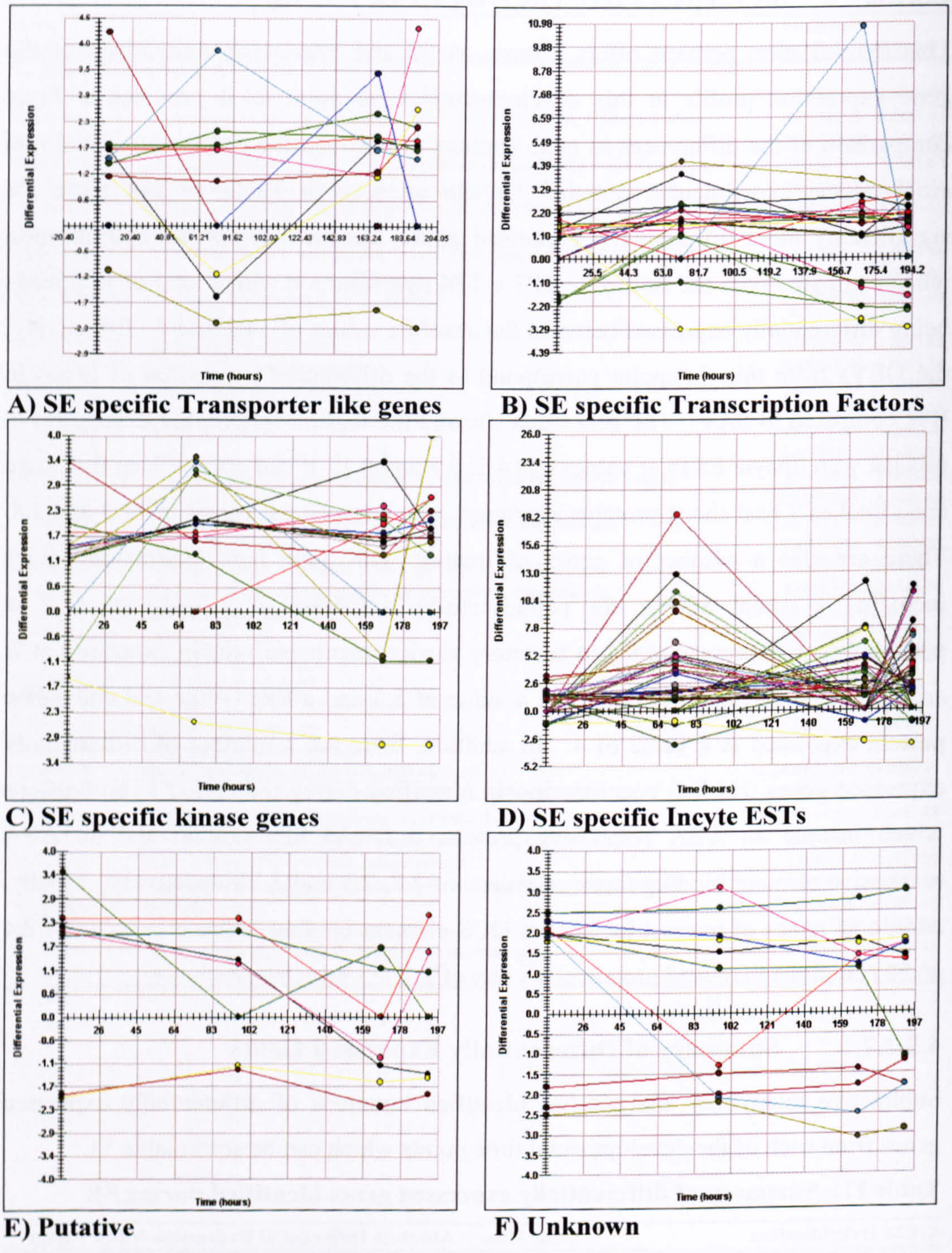


Figure 15. GEM time series plots of different gene groups. A, Clones annotated as transporter-like. B, Clones annotated as transcription factors. C, Clones annotated as kinases. D, Incyte EST clones. E, Clones annotated as putative proteins. F) Clones annotated as unknown.

4.3.3. Incyte GEMTools™ - GEM Time Series Plots

Since the number of differentially expressed clones generated is still very high, a process of further elimination is required. The most obvious way to achieve this is to select those clones that show the highest DEVs. However, this may exclude genes of particular interest expressed at lower levels such as transcription factors, or groups of genes that are known, or thought, to be involved in Somatic Embryogenesis (SE). Therefore, the expression of clones annotated with particular functions of interest such as stress response genes, cell cycle related, signalling, transcription factors, hypothetical proteins and Incyte ESTs were placed into different gene groups and analysed through a GEM time series (BW Tp0–BW Tp4) experiment (Fig. 15.).

4.4. Discussion

The genes presented in tables 5 to 8 of the results section represent a selection (full list in appendix CD) of those genes that were identified as being specifically expressed in the developing embryogenic Bobwhite (BW) cultivar and will form the basis of the discussion as candidates that are involved in the de-differentiation, initiation and development of somatic embryogenesis (SE). The genes that were differentially expressed in the non-embryogenic Savannah (SAV) cultivar were used to screen out any elements that were common to both sets of hybridisations in order to pinpoint genes that are only expressed during the embryogenic response and thus specifically involved in SE.

4.4.1. BW Tp0 Vs Tp1 minus SAV Tp0 Vs Tp1 (7Hrs 2,4-D)

Among the highly differentially expressed genes from this timepoint is a lipid transfer protein (LTP) with a DEV of 4.3. LTPs facilitate the transfer of phospholipids, glycolipids, fatty acids and hormones between membranes in mammalian cells. There are 2 types of LTPs; specific and non-specific. Specific LTPs have been implicated in response to abiotic stress and Abscisic acid (ABA) treatment and therefore may play an important signalling role in the initiation and development of SE via some form of response to high auxin stress conditions. Another highly expressed gene is the dual specificity protein phosphatase (DsPTP1), expressed at a value of 4. This gene corresponds to a cDNA identified in *Arabidopsis* which is capable of hydrolysing phosphoserine/threonine and phosphotyrosine residues of substrates (Yoo *et al.*,

2004). It was shown that calmodulin differentially regulates the activity of DsPTP1 and was proposed that the Ca_2^+ signalling pathway is mediated by cross-talk with a protein phosphorylation signal pathway in plants via protein de-phosphorylation (Yoo *et al.*, 2004). In addition, mitogen-activated protein kinases (MAPKs) play a key role in plant responses to stress and the activation and inactivation of MAPKs involve phosphorylation and de-phosphorylation on both threonine and tyrosine residues in the kinases. AtDsPTP1 was shown to de-phosphorylate and inactivate AtMPK4, a MAPK member from *Arabidopsis* (Gupta *et al.*, 1998) implicating this signalling pathway in the initiation response of SE. Another protein kinase that is expressed at this timepoint is the Protein kinase MK6 which is a member the serine threonine family of protein kinases and may play a role in signalling during the initiation of SE. A protein of particular interest is the Nt-gh3 protein expressed at a value of 2. GH3 is an early auxin-responsive cDNA isolated from *Nicotiana tabacum* by the differential display method and is a homologue of the GH3 soyabean gene which is one of the earliest molecular markers of the primary auxin response. The presence of this gene confirms the expected result that early auxin responsive genes are expressed in response to the application of 2,4-D and implicates this gene as playing an important role in the somatic embryo initiation response to the application of exogenous auxin. Also expressed at a value of 2 is the beta-glucosidase (BGQ60) gene which encodes a 60Kda protein that was purified and characterised from barley (*Hordeum vulgare L.*) seeds and appears to play a role in degrading cell wall polysaccharides during germination (Leah *et al.*, 1995). Furthermore, certain beta glucosidases are involved in regulating the biological activity of phytohormones by releasing inactive hormone-glucoside conjugates, for example a 60Kda isolated from maize (*Zea mays*) is thought to play a role in auxin-conjugate metabolism suggesting a possible role for this gene in the regulation of auxin during SE (Brzobohaty *et al.*, 1993).

The expression of a WD40 repeat protein provides a number of possible links with the initiation of SE. WD-repeat proteins are a large family found in all eukaryotes and are implicated in a variety of functions ranging from signal transduction and transcription regulation to cell cycle control and apoptosis. The underlying common function of all WD-repeat proteins is coordinating multi-protein complex assemblies, where the repeating units serve as a rigid scaffold for protein interactions. A WD40 repeat protein has been shown to regulate fungal cell differentiation and can be

functionally replaced by a mammalian homologue (Poggeler and Kuck, 2004) suggesting the existence of an evolutionarily conserved process controlling eukaryotic cell differentiation. A transcription factor that may be involved in the initiation of SE is the zinc finger transcription factor. This is part of a family of transcription factors involved in the initiation of transcription via a zinc finger motif, a DNA binding domain first recognised in TF IIIA from RNA polymerase III. Interestingly, hormone receptors in mammals have been shown to contain zinc finger motifs such as glucocorticoid which raises interesting possibilities for the role of theses genes in SE.

Expressed at values of 2.2 and 2.1 is an RNA helicase and an ATP dependent helicase. RNA helicases are proteins involved in a variety of RNA metabolic processes that burn the universal cellular fuel molecule ATP and use the energy gained from this reaction to unwind double-stranded RNA. RNA helicases have also been recently reported to be able to change the shape and composition of RNA-protein complexes without duplex unwinding and with more specificity than previously thought (Fairman *et al.*, 2004). Many of the ATP dependent RNA helicase are members of the DEAD box family which have been implicated in embryogenesis, spermatogenesis and cell growth and division in mammals. This offers an intriguing possibility for the role of this RNA helicase in the initiation and development of SE.

The expression of an HMGc2 gene at a value of 2.4 demonstrates the increase in transcriptional activity of cells after only several hours of exposure to 2,4-D. HMGc2 is a member of the chromosomal high mobility group (HMG proteins) which are small, non abundant, non histone proteins common to eukaryotes. The maize (*Zea mays*) HMGc2 is a novel plant protein containing an HMG box DNA binding domain belonging to the HMG1 protein family and is believed to be an architectural protein involved with the assembly of nucleoprotein structures associated with the transcription of plant chromatin (Grasser *et al.*, 1996). Also expressed during this timepoint are a number of genes which belong to the NADPH dependent oxidoreductase family such as a blue copper binding protein, a nodulin, cinnamoyl-CoA reductase and an NADPH dependent oxidoreductase, all of which are involved in oxidative stress and metabolism. There is also a quinone oxidoreductase and a UDP-glucose glucosyltransferase, which are involved in metabolism and ribonucleotide synthesis. The expression of these genes demonstrate the increase in

oxidative and metabolic activity of the cell which are probably as a result of their exposure to high levels of 2,4-D.

Also expressed at a value of 2.1 and 2 is a DNA J homologue and a DNA J-like domain protein. DNA J proteins are the *E.coli* homologue of the eukaryotic heatshock protein Hsp40 which bind to and regulate members of the DNA K/Hsp70 family. DNA J like proteins are important regulators of protein folding, transport, translation initiation, gene expression and growth control. The activation of heat shock proteins by 2,4-D treatment has been previously demonstrated in a number of plants and may serve as molecular chaperones during the developmental switch for the initiation of embryogenic cells.

4.4.2. BW Tp0 Vs Tp2 minus SAV Tp0 Vs Tp2 (4Ds 2,4-D)

One of the top expressers during this timepoint is a ribosomal protein expressed at a value of 7.7, which indicates the increased levels of protein translation and developmental activity of the cell during this time point. This is confirmed by the identification of a eukaryotic initiation factor (EIF3) at a value of 2. EIF3 relates to a meristem specific cDNA identified in maize that encodes for the homologue of a large subunit of the EIF3 an essential multi protein complex for the initiation of protein synthesis (Sabelli *et al.*, 1999). In addition a putative peptide transporter, most closely related to the peptide transporter family, is seen to be expressed at a value of 6.3. Peptide transporters fall into three families; 1) the ATP binding cassette family (ABC transporters); 2) the peptide transporter family (PTF); 3) the oligopeptide transporter (OPT family). It was demonstrated that the down-regulated expression of *At*PTR2, a peptide transporter, resulted in the arrest of seed development (Song *et al.*, 1997) and furthermore that a T-DNA insertion in *At*OPT3 resulted in the arrest of embryo development (Stacey *et al.*, 2002), clearly implicating the importance of the role of peptide transport during embryogenesis and plant development.

Expressed at a value of 6.8 is a cold responsive late embryo abundant protein (LEA)/ responsive to abscisic acid (RAB) gene isolated from wheat. It is a group 3 LEA/RAB protein designated as Wrab19 and was shown to be responsive to ABA and GA3 (Tsuda *et al.*, 2000). The expression of this gene suggests the possible involvement and importance of additional hormone pathways other than auxin in the

development of SE. However, its expression could also be as a result of a general stress response. Another gene of interest in this context is the expression of an ABA responsive elements-binding factor. ABA plays an important role in the environmental stress response of higher plants. One of the ABA-mediated responses is the induced expression of a large number of genes controlled *cis*-regulatory elements known as ABA-responsive elements (ABREs). Using an ABRE in a yeast 1 hybrid system, a protein named ABF (ABRE binding factor) was isolated and characterised as belonging to a subfamily of basic leucine zipper (bZip) proteins (Choi *et al.*, 2000). Expressed at a value of 6.3 is a beta glucosidase (BGQ60) that was initially expressed at a value of 2 after 7 hours of treatment with 2,4-D. The increased transcription of this gene confirms the up-regulation of this gene in the initiation and development of SE and its influence either through its role in cell wall degradation or phytohormone metabolism as previously discussed.

Several transcription factors were expressed during this timepoint and of particular interest is the expression of a Scarecrow (SCR) transcriptional regulator. SCR encodes a putative transcription factor that was first identified as regulating asymmetric cell division and is essential for generating the radial organisation of *Arabidopsis* root (Di Laurenzio *et al.*, 1996). SCR has also been shown to be required for distal-specification of the quiescent centre in *Arabidopsis* root meristem which in turn regulates stem cell fate of the immediately surrounding cells (Di Laurenzio *et al.*, 1996). This role in the regulation of stem cell fate and asymmetric division has obvious implications for its potential involvement in the regulation and development of SE with regards to meristematic activity and the initiation of asymmetry in developing SEs. Another interesting transcription factor up-regulated during this timepoint is the putative MADS domain transcription factor GGM13. In-situ hybridisation experiments revealed that the MADS GGM13 mRNA was detectable in inflorescence meristems before the differentiation of flowers was detected (Sheppard *et al.*, 2000). MADS box genes represent a large family of transcription factors with essential functions during flower development and organ differentiation processes in plants. (Heuer *et al.*, 2001) isolated a maize (*Zea mays*) MADS box cDNA (ZmMADS1) and found that its expression during SE is restricted to cells with the capacity to form somatic embryos and was not detectable in non-embryogenic suspension cultures or in zygotic embryos, confirming the expression of this MADS

gene in embryogenic cultures and its importance in the development of SE. Another gene that has been previously shown to be specifically expressed in embryogenic cultures is the *Dactylis glomerata* SE related protein (DGE) expressed at a value of 2.2. DGE was originally identified by differential display as being differentially expressed in embryogenic but not non embryogenic leaf cultures and shows strong similarity with the WRKY DNA binding protein 21 from *Arabidopsis* which contains putative nuclear localisation sequences (Alexandrova and Conger, 2002). The WRKY family of genes appear to be involved in the regulation of various physiological programmes that are unique to plants and possess a highly conserved WRKY amino acid sequence that constitutes a DNA binding domain, implicating this gene in the regulation of transcription in embryogenic cultures.

The expression of a putative ankyrin at a value of 4.3 suggests an increase in cytoskeletal activity during SE. Ankyrins are a family of membrane associated proteins responsible for the attachment of the spectrin cytoskeleton to the cell membrane and have been identified as important factors in a variety of cell responses such as the regulation of transcription and the cell cycle, cell polarity, cell adhesion, and orientation of mitotic spindles in asymmetric cell divisions (Bennett and Baines, 2001). Another cytoskeletal related gene expressed during this timepoint is a villin at a value of 3.5. Villin is a member of a protein family which regulates eukaryotic cytoskeletal actin by severing, capping and nucleating actin filaments. It is thought to have two calcium binding domains for calcium dependent active binding which is interrupted by the presence of phosphatidyl-inositol-4,5-biphosphate (Markus *et al.*, 1997) thereby implying Ca_2^+ and phosphatidyl-inositol-4,5-biphosphate in the signalling and restructuring of the cytoskeleton in the development of SE.

Also expressed at a value of 3.5 is a cytokinin regulated kinase 1 (CRK1) gene which encodes a CDC2-related protein kinase that is located in the plasma membrane. It has a single transmembrane domain followed by a highly conserved intracellular serine/threonine kinase domain. It has been proposed that CRK1 is involved in an early step of auxin signalling as abscisic acid (ABA) and auxin negatively regulate this gene in a similar way to cytokinin (Schäfer and Schmülling, 2002). The importance of this gene in cell growth is demonstrated by the observation that gene knock out experiments in *Candida albicians* decreased the growth rate of cells (Chen

et al., 2000). CRK1 therefore may be of importance in the phytohormone regulated growth rate of developing embryogenic calli. Another gene of interest in the regulation of cytokinesis is the CXC domain containing TSO1-like protein. TSO1 is a nuclear protein that has been shown to modulate cytokinesis and cell expansion in *Arabidopsis* (Hauser *et al.*, 2000). TSO1 mutants have been shown to develop callus like tissues in place of floral organs due to defects in mitosis and cytokinesis of the floral meristem (Liu *et al.*, 1997).

Expressed at a value of 3.7 is a rice embryo-specific gene (Ose731) which was originally identified from cDNA clones specifically expressed in rice embryos ten days after pollination (Chen *et al.*, 1997) although its function was unknown. In addition a maize homologue of this gene was also identified as being differentially expressed in an expression profiling experiment of dessication tolerance in maize and is thought to be involved in stress response during seed maturation and germination (Kollipara *et al.*, 2002).

During this hybridisation there are a number of cell cycle associated genes that demonstrate the increased activity in cell division. Expressed at a value of 3.2 is a proliferating cell nuclear antigen (PCNA) which has a cell cycle specific expression pattern and has been identified as a gene that plays an essential role in DNA replication in mammals. Homologues have been identified in a number of different plant species and have strong similarity with mammalian PCNA, suggesting a strongly conserved DNA replication mechanism throughout the eukaryotes (Suzuka *et al.*, 1991). The characterisation of *Os*PCNA found it to be present in meristematic tissues such as cultured cells and the shoot and root apical meristems in a cell cycle dependent distribution pattern. Furthermore, *Os*PCNA was found to be localised in the nucleus during mitotic and meiotic cell divisions (Kimura *et al.*, 2001). The expression of a replication protein (*Os*RPA1) and a DNA mismatch repair protein at 2.7 and 2.9 respectively confirmed the involvement of DNA replication and mismatch repair activity in the actively dividing embryogenic calli cells. There is also an expressed mini-chromosome maintenance deficient 6 gene (MCM6) at a value of 2.8 which is a highly conserved group of proteins essential for the initiation of eukaryotic genome replication. MCM6 Binds to chromatin during G1 and detaches from it during S phase as if it licenses the chromatin to replicate (Tsuruga *et al.*, 1997). The

MCM complex has helicase activity and may act as a DNA unwinding enzyme during the replication of chromosomal DNA. Phosphorylation by CDC2 kinase reduces the helicase activity suggesting a role in the regulation of DNA replication (Lei and Tye, 2001).

Other genes involved in the regulation of chromatin are expressed during this timepoint. At a value of 2.5 is *PICKLE* (*PKL*), a CHD3 chromatin remodelling factor that has been shown to regulate the transition from embryonic to vegetative development in *Arabidopsis* (Ogas *et al.*, 1999). *PKL* is necessary for the repression of *LEC1*, a gene implicated as a critical activator of embryo development and it has been proposed that PKL is a component of a gibberellic acid (GA) modulated pathway that functions to establish repression of embryonic identity during germination (Ogas *et al.*, 1999). The up-regulation of this gene during the development of embryogenic calli is somewhat unexpected. However, it is possible that due to the proliferative state of the embryogenic calli *PKL* acts to keep the nature of this proliferation under a certain degree of control. In addition expressed at a value of 2.4 is *Zea mays* Tousled-like kinase 4 (MTK-4). The *TOUSLED* (*TSL*) gene was identified in *Arabidopsis* and encodes a protein kinase that is required in the floral meristem for the correct initiation of floral organ primordia and development of organ primordia (Roe *et al.*, 1993). Double mutant analysis suggests that TSL interacts with ETTIN by restricting TSL to apical regions (Roe *et al.*, 1997). Tousled-like kinases (TLKs) constitute a family of SER/THR kinases conserved in plants and animals that act in a cell cycle dependent manner. TLKs act during S-phase where they phosphorylate the anti-silencing function protein (ASF1), a histone chaperone involved in replication dependent chromatin assembly. Loss of function and overexpression experiments suggest that TLK coordinates cell cycle progression through the reorganisation of chromatin dynamics (Carrera *et al.*, 2003).

Finally, the down-regulation of a seed imbibition protein SIP1 highlights the importance of the secondary action of 2,4-D on wheat immature embryos which is to suppress the germination of the embryo proper during the initiation of the embryogenic calli.

4.4.3. BW Tp0 Vs Tp3 minus SAV Tp0 Vs Tp3 (7Ds 2,4-D)

A number of genes identified from the previous hybridisations are continued to be expressed at this timepoint. For example, the continued expression of HMGC2 at a value of 3.7, PCNA at a value of 3.1, DNAJ homologue and a WD40 repeat protein at a value of 2.9 highlights the importance of these genes throughout the initiation and development of SE. The continued expression of an ankyrin with a value of 3.4 highlights the importance of the cytoskeletal reorganisation during SE. Ankyrins are a family of spectrin binding proteins that are associated with several different membrane proteins and have been proposed to mediate interactions such as the regulation of transcription and the cell cycle. In addition the continued expression of Mis5/MCM6 confirms the involvement of these proteins in the initiation of genome replication of embryogenic cells. Also the expression of another lipid transfer protein during this timepoint at a value of 2.5 confirms the potential role of these genes in the initiation and development of SE.

One of the highest expressed genes at this timepoint is a beta expansin (B3). As previously discussed (1.5.2.), expansins function by weakening non-covalent binding between wall polysaccharides allowing turgor driven polymer-creep (McQueen-Mason and Cosgrove, 1994) and have been demonstrated in a number of growth related processes including auxin induced root formation and the emergence of primordia in the meristem where expansin expression precedes primordia outgrowth. In contrast to *Arabidopsis* beta expansins in maize are more numerous and more highly expressed than alpha expansins and are often expressed in specific tissues and cell types (Cosgrove *et al.*, 1997). It is thought that beta expansins have evolved special functions in the grasses and may have an important role in the regulation of SE in wheat. In addition, the expression of a second beta expansin at a value of 2.8 confirms the expression of this gene type in developing embryogenic calli. The expression of a wheat endochitinase gene at a value of 2.4 is consistent with the finding that the secreted proteins are required for the induction of SE as discussed (1.5.2) and also help confirm the differential expression of genes that are expected to be induced during SE.

The expression of indole-3-acetic acid (IAA) amino acid hydrolase suggests that there may be an endogenous control of auxin levels during SE. IAR3 was identified in *Arabidopsis* as a mutant that displayed insensitivity to IAA-Ala (Davies *et al.*, 1999). Conjugates of IAA are putative storage or inactivation forms of the growth hormone auxin which can be hydrolysed to supply the plants with the active hormone. There are also a number of genes expressed during this timepoint that are regulated by phytohormones other than auxin. For example the expression of alpha-amy2 at a value of 4.9 implicates the involvement of the gibberellin pathway during SE. The alpha-amy2 genes of wheat are a multigene family which have shown to be expressed in the aleurone cells of germinating grain under the control of gibberellin (Huttly *et al.*, 1988). Furthermore, the down-regulation of gamma thionin HTH3 also implies the regulation of the gibberellin pathway. Gibberellins are known to induce expression of genes such as alpha amylase in the aleurone layer of cereals, however a number of mRNA species are specifically repressed by gibberellin, such as HTH3, which is repressed during treatment with gibberellic acid (GA3) (Heck and HO, 1996). In addition, the expression of a systemic acquired resistance (SAR) DNA binding protein, involved in triggering cell death at the site of pathogenesis infection, is regulated by the ethylene pathway (Yu *et al.*, 2001). Protein phosphatase 2C (PP2C) is also expressed at a value of 2.2 and corresponds to an *Arabidopsis* PP2C gene that contains sequence similarity to the abscisic acid insensitive (ABI) genes ABI1 and ABI2 and was shown to be up-regulated by ABA (Rodriguez *et al.*, 1998). The presence of these genes confirms that the applications of auxin and/or cytokinin are not the only phytohormones responsible for regulating SE. Although auxin may be required for the initial impetus it appears that a more sophisticated interaction involving additional phytohormone pathways is required for the regulation of an embryogenic response.

The expression of a protein kinase at 2.4 is also of interest as it is a homologue of the *Arabidopsis* serine/threonine kinase 1 (ASK1) which has shown to be activated by osmotic stress in tobacco cells (Mikolajczyk *et al.*, 2000) and therefore may play a role in the cell signalling of hormone stress responses in plants. Two genes of particular interest involved in kinase signalling are a leucine rich receptor kinase and a somatic embryogenesis receptor like kinase (SERK). The amino acid sequence of SERK shows that its gene encodes a leucine-rich repeat containing receptor-like

kinase protein (Schmidt *et al.*, 1997). The expression of these genes confirms the importance of this gene family in the development of SEs in plants as previously discussed in 1.5.1.2. This particular gene shows homology to a SERK that was identified from maize (*Zea mays*) demonstrating the existence of this mechanism among the grasses as well as carrot and *Arabidopsis* cultures. Also identified at an expression value of 2 is a calcium dependent protein kinase (CDPK). CDPKs are essential sensor-transducers of calcium signalling in plants and have been shown in mung bean to be inducible when treated with IAA (Botella *et al.*, 1996). The role of calcium increase due to the application of auxin has been well documented (Gehring *et al.*, 1990) and has been shown as playing a major role as a second messenger in the transduction of auxin signalling (Yang and Poovaiah, 2000). The long lived calcium response may be mediated by the binding of calcium to calmodulin complexes which acts as a protein kinase phosphorylating proteins involved in a number of cellular processes such as differentiation.

Expressed at a value of 2.1 is the transcription factor IIA (TFIIA) identified as a component of the transcription initiation complex of RNA polymerase II and demonstrates the increased transcription activity in cells undergoing SE. This is consistent with the finding that the transition from unorganised cell proliferation to SE requires active RNA synthesis as previously discussed (1.5.). In this context there are several transcription factor genes expressed during this timepoint. Of particular interest is the expression of *CONSTANS* which encodes for a transcription factor that is essential for flowering and the activation of genes required for floral initiation (Simon *et al.*, 1996). It is also interesting to note that there is the expression of another MADS box gene (MADS9). HvMADS box9 belongs to the Agamous-like (AGL) sub group and was shown to be localised by *in situ* hybridisation experiments to specific regions of the floret meristem before the emergence of organ primordia (Schmitz *et al.*, 2000). The expression of a second *SCARECROW* like gene confirms the potential importance of this gene family in the meristematic activity in developing embryogenic calli.

Also expressed at a value of 2.1 is a SET (*Su(var)3-9,enhancer of zest, trithorax*) domain protein, a class of proteins implicated in regulating gene expression through histone methylation which is associated with chromatin accessibility (Springer *et al.*,

2003). A gene of particular interest with regards to methylation is the expression of cytosine 5-DNA methyltransferase at a value of 3.8. The expression of this gene demonstrates that there is an increase in DNA methylation activity during somatic embryo development. This is consistent with the finding that there was significant hyper-methylation following the application of auxin in Napier grass calli (Morrish and Vasil, 1989). In higher eukaryotes DNA methylation of cytosine residues appears to participate in the control of gene expression and developmental regulation as well as genomic imprinting and chromatin silencing. This raises interesting questions into the role of methylation in the transcriptional regulation of genes important for SE.

The highly down regulated myo-inositol phosphate synthase at a value of -16.4 indicates the reduction in the production of inositol phosphate. Interestingly, one of the earliest signal transduction events involved in the triggering of development by fertilisation in *Fucus* embryos is the breakdown of inositol phosphate followed by the subsequent increase in calcium and pH (Turner *et al.*, 1984). The role of inositol phosphate in the activation of egg development raises interesting questions about whether this signalling pathway is important for the initiation of SE.

Also down-regulated at a value of 3.9 is a WPK4 protein kinase identified in wheat as an SNF 1 related protein kinase (snRK) that contains a domain recognised by cytoskeletal elements and signal transduction proteins. WPK4 is also shown to be differentially regulated by cytokinin and sucrose (Ikeda *et al.*, 1999). It is possible that this protein plays a role in signalling and cytoskeletal reorganisation during SE. The continued expression of a SIP1 reiterates the need for the repression of normal germinative development during embryogenic callus formation. The down regulation of an ABA responsive protein again highlights the regulation of the ABA pathway during SE. Finally, there is the down regulation of a basic leucine zipper protein (bZip), a transcription factor that is differentially regulated by changes in temperature and is therefore probably involved in the regulation of a stress response to the exposure of high levels of auxin (2,4-D).

4.4.4. BW Tp0 Vs Tp4 minus SAV Tp0 Vs Tp4 (7Ds 2,4-D, 1D MS)

A large number of genes identified during this timepoint that have also been identified as being differentially expressed during the earlier timepoints. These include

HvMADS9, CRK1, beta expansin, WRKY2 DNA binding protein, HMGC2, ankyrin, PCNA, DNAJ-like, ATP dependent RNA helicase, Mis5/MCM6, Zinc finger protein, ABA responsive protein, SIP1, and Os bZip protein. Although it is possible that this may be a result of continued mRNA stability, it seems to suggest the need for the continued expression of these genes even after removal of 2,4-D from the tissue culture medium. There are, however, some genes that have been identified as being unique to this timepoint which are of particular interest as they may be specifically involved in the regulation of an early differentiation response of SEs that have been removed from high auxin media. The genes with highest DEVs are a number of hypothetical and unknown proteins or ESTs. Although there is no annotation for these genes to provide any clues as to there possible involvement in the early differentiation of SEs, the expression of these unknown genes during SE allows one to loosely suggest functional roles as being SE related. In addition, the fact that these genes are unknown opens the possibility of identifying novel genes involved in SE and the identification of novel gene families.

There are also a number of genes that are uniquely expressed during this timepoint which are annotated. These include a *Zea mays* response regulator 6 (ZmRR6) which encodes for a cytokinin responsive protein that localises in the nucleus and has been shown to be differentially expressed in vascular tissue (Nakazono *et al.*, 2003) indicating the differentiation of embryogenic cells into adult tissues. Expressed at a value of 2.8 is a polygalacturonase which is a cell wall degrading enzyme that hydrolyses pectin and other galacturonans. The expression of this gene may have some significance in the role of SE in the same way that arabinogalactan protein (AGPs) are hydrolysed by an endochitinase to promote SE as previously discussed in 1.5.2. Another gene of interest is the expression of an *Os*Ran protein expressed at a value of 2.2. Ran is a conserved Ras-like GTP binding protein implicated nucleocytoplasmic transport, cell cycle progression, spindle assembly, nuclear organisation, nuclear envelope assembly and functions to control progression through the cell cycle by regulating the transport of nucleic acids or proteins across the nuclear membrane. HvGPX12 encodes a glutathione peroxidase from barley (*Hordeum vulgare*) which was shown to be differentially expressed in a negative fashion in response to stress (Churin *et al.*, 1999) and is therefore probably involved as a stress response.

The down regulation of VIP2 protein is of particular interest as VIP2 is a viviparous 1 (VP1) interacting protein (VIP) which was also shown to interact with the *Arabidopsis ABSCISIC ACID INSENSITIVE 3* (*ABI3*) homologue. *VIP2* and *VP1* are highly expressed in dormant embryos and have a suggested role in the *VP1* mediated regulation of dormancy to germination transition in *A.fatua* (Jones *et al.*, 2000). *VP1* from maize has been shown to encode a seed specific transcription factor and mutants in *VP1* germinate precociously. The down regulation of such genes would be expected to result in the differentiation and development of SEs into adult structures which is consistent with the response observed of SEs that have been removed from a high auxin media. It is possible that the VIP2 gene is only expressed in the embryo axis and that the apparent down-regulation of VIP2 is as a result of a dilution effect on the mRNA population by the callus tissue. This highlights the need for further characterisation of the expression patterns of such transcripts by ISH which can determine their localised expression. As previously discussed inositol-based compounds may have an important functional role in the development and maintenance of SEs. Ssh1p is a membrane associated phosphatidyl inositol transfer protein which becomes rapidly phosphorylated and dissociates from the membrane under high osmotic stress conditions. Polyphosphoinositides (PPI) have also been shown to inhibit gelosin, an actin filament severing protein that has profound effects on actin filament organisation (Yu *et al.*, 1992) suggesting their importance as mediators of signalling events including the rearrangement of the cytoskeleton.

4.4.5. BW Tp3 (7Ds 2,4-D) Vs BW Tp4 (7Ds 2,4-D, 1D MS)

Overall the expression of genes during this hybridisation, which was designed to show differences in the transcription of genes as a result of the removal of 2,4-D from the tissue culture medium, shows only a few additional genes of interest. The genes with the highest DEVs are a putative protein, a hypothetical protein and an Incyte EST and therefore provide little indication as to their possible function during SE. Because starch is the main form of carbohydrate in plants, the up-regulation of a starch synthase indicates an increase in the metabolism of embryogenic calli when withdrawn from 2,4-D media. Similarly, the expression of an NADPH ubiquinone oxidoreductase and a ubiquinone biosynthesis protein at expression values of 3.6 and 3.3 respectively also suggest an increase in cell metabolism on the removal of 2,4-D.

The expression of such genes is not surprising as the removal of embryogenic calli from 2,4-D media results in the differentiation of SEs into shoot structures. Nevertheless they may play an important role in the regeneration of adult structures from the embryogenic calli. In addition, the expression of DNAJ-like genes at 2.6 and 2.5 again highlight the importance of these chaperone proteins throughout each of the different timepoints during SE. Additional genes of interest that are expressed during this hybridisation are the TAR (transactivation response) RNA loop binding protein expressed at a value of 2.1 and a nucleic acid binding protein at a value of 2. TAR is an RNA regulatory element which binds RNA polymerase and stimulates transcription elongation, suggesting there is an increase in transcriptional activity due to the removal of 2,4-D from the tissue culture medium.

4.4.6. SAV Tp4 Vs BW Tp4 (7Ds 2,4-D, 1D MS)

The genes identified from this hybridisation provide a more interesting perspective than the difference in expression between Bobwhite timepoint 3 versus Bobwhite timepoint 4, as it is the only direct comparison between the embryogenic and non embryogenic cultivars from all of the hybridisations. Topping the list is a peptidylprolyl isomerase at a DEV of 11.7. This cDNA belongs to a heat stress induced FK506 binding protein (FKBP) family that possess three FKBP-like domains and a calmodulin binding domain (Kurek *et al.*, 1999). The transcription of this cDNA was highest in imbibed mature embryos and two day old green shoots, which is consistent with the expected expression of genes involved in early shoot formation and emerging adult structures from differentiating embryos. When cells are exposed to elevating temperatures or other environmental stresses, heat shock proteins (HSPs) belonging to several different families are induced including an ATP dependent CLPB and a stress induced protein (STI1) which are also HSPs. Many of these HSPs function as chaperones and play important roles in maintaining correct protein folding during normal growth as well as under stress conditions. In particular, HSP101/ClpB gene expression can be induced by dehydration and ABA treatment indicating potential involvement of HSP100 in plant osmotic stress responses (Campbell *et al.*, 2001). It can be seen from the large number of HSPs that are expressed at this timepoint that the differential expression of these genes during embryo development suggests a specific functional role for HSPs in plant cells undergoing SE.

In addition to the highly differentially expressed HSPs there is a phosphoinositide-specific phospholipase C (PI-PLC) that is highly differentially expressed at a value of 9.1. The breakdown of inositol phosphatase as a primary signal in egg activation has already been discussed and the expression of this gene further confirms the involvement of this important signalling mechanism in the development of embryogenic cells. Furthermore, PI-PLC contains a C-terminal domain that shows homology to the C2 domain of protein kinase C. A targeted knockout of PI-PLC gene from the moss *Physcomitrella patens* was generated via homologous recombination and showed a loss of sensitivity to cytokinin and a reduced ability for protonemal filaments to grow negatively gravitropically in the dark (Repp *et al.*, 2004) implying a significant role for this gene in cytokinin signalling and gravitropism.

Another gene of interest with a high DEV is the secretory carrier membrane protein (SCAMP) expressed at a value of 7. SCAMPs are integral membrane proteins found in the secretory and endocytotic carriers implicated to function in membrane trafficking and are highly conserved across the plant and animal kingdoms (Hubbard *et al.*, 2000). The importance of secretory molecules in embryogenic cultivars has been discussed (1.5.2). It is possible that SCAMP mediates the secretion of molecules such as AGPs and endochitinases that have been shown to increase embryogenicity in SEs

AT hook DNA binding domains are protein domains that were first described in HMG proteins, proteins that bind preferentially to the minor grove of AT rich regions in dsDNA (Reeves and Nissen, 1990). Since their discovery, this motif has been observed in other DNA-binding proteins from a wide range of organisms and it appears that the AT-hook motif is an auxiliary protein motif that cooperates with other DNA-binding activities and facilitates changes in the structure of the DNA (Aravind and Landsman, 1998). HMG proteins are involved in the transcriptional regulation of genes containing AT rich regions and have been identified as being differentially expressed in several of the previous timepoints during SE. It is difficult to say whether there is any connection between the expression of this AT hook DNA binding gene and the expression of the previously identified HMG proteins. However, the relatively high DEV of 5.3 suggests it may be an important transcriptional regulator of genes specifically involved in the development of embryogenic calli.

Table 12: Shortlist of genes for further characterisation by *in situ* hybridisation.

Incyte Clone ID	Annotated Gene Name	Differential Expression Value						Accession Number	E-value
		0v1	0v2	0v3	0v4	3v4	4v4		
702008465	hypothetical protein; 53156-50996	5.9	-	-	-	-	-	AAG50962	9 00E-13
701989150	Incyte EST	5.9	-	-	-	-	-		No Hit
701996695	Incyte EST	4.1	-	-	-	-	-		No Hit
702039485	lipid transfer protein	4.3	-	-	-	-	-	CAA69949	4.00E-28
702006345	contains similarity to DsPTP1 protein~gene_id.MQM1.1	4	2.7	-	2.6	-	-	BAB10045	3.00E-57
701960043	putative NADPH-dependent oxidoreductase	3.4	3.9	2.2	-	-	-	AAF13742	2.00E-54
702013541	Incyte EST	3.3	4.3	-	-	-	-		No Hit
702001685	Incyte EST	2.9	4.9	-	-	-	-		No Hit
702015235	HMGc2	2.4	4.6	3.7	2.8	-	-	CAA69606	7.00E-08
701988495	DnaJ homologue	2.1	3.5	2.9	2.5	-	-	BAB11149	3.00E-23
702015038	Rice (O. sativa) gene for proliferating cell nuclear antigen (PCNA).	2.1	3.2	3.1	2.4	-	-	X54046	1.00E-81
701967005	protein kinase MK6	2	-	-	-	-	-	CAB82852	3.00E-12
701958460	Nt-gh3 deduced protein	2	-	-	-	-	-	AAD32141	1.00E-11
701958329	contains similarity to a DNAJ-like domain	2	-	2	2.5	2.5	-	AAB61072	5.00E-06
701989078	Incyte EST	-	18.4	-	-	-	-		No Hit
702039563	Incyte EST	-	12.8	-	-	-	-		No Hit
701700561	Incyte EST	-	11.2	-	-	-	-		No Hit
701965472	Incyte EST	-	10.2	-	-	-	-		No Hit
701996792	putative peptide transporter	-	6.3	-	-	-	-	AAG46154	9.00E-40
701960557	SCARECROW transcriptional regulator-like	-	4	-	-	-	-	BAB10182	2.00E-32
701957921	protein phosphatase 2C	-	4	2.2	-	-	-	CAA05875	1.00E-06
702006188	Oryza sativa embryo-specific (Ose731) mRNA, complete cds.	-	3.7	-	-	-	-	AF049892	0.00E+00
702014751	CRK1 protein	-	3.5	-	4	-	-	CAB89665	2.00E-36
701985820	CXC domain containing TSO1-like protein 1	-	3.4	-	-	-	-	AAF69125	7.00E-14
701987309	putative cell cycle protein kinase	-	3.1	2.4	-	-	-	AAG60199	4.00E-64
702007940	putative MADS domain transcription factor GGM13	-	2.6	-	-	2.6	-	CAB44459	3.00E-53
701993859	abscisic acid responsive elements-binding factor	-	2.6	2	-	-	2.1	AAF27182	3.00E-49
701969217	mis5 protein	-	2.5	2.4	2.1	-	-	CAB75412	3.00E-89
701991445	GYMNOS/PICKLE	-	2.5	-	-	-	-	AAF07084	2.00E-73
701995510	Dactylis glomerata somatic embryogenesis related protein mRNA	-	2.2	-	-	-	-	AY011122	0.00E+00
702009335	Ser-Thr protein kinase-like protein	-	2	-	-	-	-	BAB02869	4.00E-40
702046792	Rice lip19 mRNA for basic/leucine zipper protein.	-	-2.1	-2.3	-2.5	-	-	X57325	4 00E-33
701997623	ethylene-responsive transcriptional coactivator-like protein	-	-2	-	-	-	-	BAB01997	No Hit
701961663	Incyte EST	-	-	12	-	-	-		No Hit
702036677	Incyte EST	-	-	7.7	-	-	-		No Hit
701961028	Incyte EST	-	-	5	-	-	-		6.00E-19
702011085	SAR DNA binding protein	-	-	2.4	-	-	-	BAA31260	2.00E-15
701966508	Hordeum vulgare mRNA for MADS-box protein 9 (m9 gene).	-	-	2.2	4.6	-	-	AJ249147	3.00E-07
702006429	flowering protein CONSTANS, putative; 7571-5495	-	-	2.2	2	-	-	AAG52620	4.00E-05
701959805	leucine-rich receptor-like protein kinase, putative; 84911-81624	-	-	2.2	-	-	-	AAG51803	9.00E-18
701995104	transcription factor IIA small subunit	-	-	2.1	-	-	-	CAA11524	5.00E-04
703425203	somatic embryogenesis receptor-like kinase 3	-	-	2.1	-	-	-	CAC37642	2.00E-03
701988922	Incyte EST	-	-	-	11.6	-	-		No Hit
701968396	Incyte EST	-	-	-	11	-	-		No Hit
701998612	Nicotiana tabacum DNA-binding protein 2 (WRKY2) mRNA.	-	-	-	2.9	-	-	AF096299	5.00E-14
702017431	Hordeum vulgare gene for ids2, complete cds.	-	-	-	2.7	-	-	D15051	7.00E-17
701995007	Incyte EST	-	-	-	-	8.5	-		No Hit
701965813	Incyte EST	-	-	-	-	6.3	-		No Hit
701965494	putative protein	-	-	-	-	5.7	-	CAB96849	1.00E-71
702874380	TAR RNA loop binding protein	-	-	-	-	2.1	-	AAC50379	9.00E-25
701968565	nucleic acid binding protein	-	-	-	-	2	-	AAD31844	4.00E-15
701995191	Incyte EST	-	-	-	-	-	14.2		No Hit
702035890	T.aestivum mRNA for peptidylprolyl isomerase.	-	-2.5	-	-	-	11.7	Y07636	0.00E+00
701991275	Oryza sativa secretory carrier membrane protein (SC) mRNA.	-	-	-	-	-	7	AF225922	5.00E-19
701968407	AT-hook DNA-binding protein (AHP1)	-	-	-	-	-	5.3	AAB80677	4.00E-07
701966045	cell division related protein-like	-	-	-	-	-	2.1	BAA98200	3.00E-42
702042046	similar to mouse Glt3 or D. malanogaster transcription factor IIB	-	-	-	-	-	2	AAH05152	1 00E-05

The expression of an RNA binding protein at a DEV of 5 is also of interest, as dsRNA binding proteins are involved in the regulation of miRNAs and therefore the presence of this gene raises interesting questions about the possible role of this regulatory mechanism in the development of SE.

An interesting protein implicated in cell cycle control is a cullin 3-like protein that is expressed at a value of 4. Cullin 3 was isolated in a 2 hybrid screen to identify proteins that could bind cyclin E (Singer *et al.,* 1999). Overexpression of cullin 3 specifically increased ubiquitination of cyclin E whereas the deletion of the cullin 3 gene resulted in the accumulation of cyclin E and had specific effects on S-phase regulation (Singer *et al.,* 1999). Cyclin E is an evolutionary conserved protein whose essential function is to promote the cell cycle transition from G1 to S-phase by binding to and activating cdk 2. On entry into S phase, cyclin E is abruptly destroyed by the proteasome to which it is targeted by ubiquitination. Therefore the expression of cullin suggests that there are a greater number of actively dividing cells in the embryogenic response that are in the S-phase of the cell cycle.

4.5. Gene Shortlist

Following the validation, screening, annotation and evaluation of the genes identified as being differentially expressed in embryogenic wheat calli, a subset of these genes were selected for further analysis based on their DEVs, expression profiles and functional significance. The genes that comprise this shortlist have all been discussed with regards to their potential role in SE (4.4.1 – 4.4.6). A significant number of the bacterial clones selected were either, contaminated and unable to be used, or did not generate a PCR product, leaving the shortlist of genes for *in situ* hybridisation analysis as seen in Table 12 opposite. Each Incyte clone is listed with its annotated gene name derived from the best sequence alignment in genbank along with the accompanying accession number and E-value, and is presented with the differential expression value during each of the GEM hybridisations. The corresponding rice orthologues can be seen in Table 13 below. The wheat EST sequences were used to perform Blast searches of the rice database and the presented rice hits represent the highest scoring alignment for each sequence. The E-values for each alignment are given along with the full rice mRNA sequence (see appendix CD for spreadsheet containing sequences).

Table 13: Sequence alignment searches of wheat EST sequences against the rice genome.

Clone ID	Incyte Annotated Gene Name (Best Hit)	E-value	Rice hit	E -value
702008465	hypothetical protein, 53156-50996	9 00E-13	>Paxneb protein-like	3.00E-64
701989150	Incyte EST	No Hit	>unknown protein	2 00E-15
701996695	Incyte EST	No Hit	No Hit	/
702039485	lipid transfer protein	4 00E-28	phospholipid transfer protein homolog - rice	1.00E-21
702006345	contains similarity to DsPTP1 protein-gene_id MQM1.1	3 00E-57	putative PROPYZAMIDE-HTPERSENSITIVE 1	9 00E-66
701960043	putative NADPH-dependent oxidoreductase	2.00E-54	OSJNBa0064H22.5	6 00E-84
702013541	Incyte EST	No Hit	No Hit	/
702001685	Incyte EST	No Hit	putative CGI-94 protein	2 00E-08
702015235	HMGc2	7.00E-08	putative HMG type nucleosome/chromatin assembly factor D	2 00E-42
701988495	DnaJ homologue	3 00E-23	putative DnaJ homolog, subfamily C, member 9	4 00E-81
702015038	Rice (O. sativa) gene for proliferating cell nuclear antigen (PCNA).	1.00E-81	SPATULA-like	8 00E-33
701967005	protein kinase MK6	3.00E-12	No Hit	/
701958460	Nt-gh3 deduced protein	1 00E-11	putative Nt-gh3 deduced protein	2 00E-12
701958329	contains similarity to a DNAJ-like domain	5 00E-06	putative DNAJ heat shock N-terminal domain-containing protein	1 00E-57
701989078	Incyte EST	No Hit	No Hit	/
702039563	Incyte EST	No Hit	No Hit	/
701700561	Incyte EST	No Hit	nuclear transport factor 2 (NTF-2)	9 00E-09
701965472	Incyte EST	No Hit	putative transposase	3 00E-06
701996792	putative peptide transporter	9 00E-40	putative peptide transporter	2.00E-70
701960557	SCARECROW transcriptional regulator-like	2 00E-32	OSJNBa0064M23 7	1.00E-27
701957921	protein phosphatase 2C	1.00E-06	OSJNBa0052K01	7 00E-13
702006188	Oryza sativa embryo-specific (Ose731) mRNA, complete cds.	0.00E+00	embryo-specific protein	e-112
702014751	CRK1 protein	2 00E-36	putative CRK1 protein	4 00E-80
701985820	CXC domain containing TSO1-like protein 1	7 00E-14	'hypothetical protein, contains Tesmin/TSO1-like CXC domain	2 00E-04
701987309	putative cell cycle protein kinase	4 00E-64	putative cell cycle protein kinase	5 00E-76
702007940	putative MADS domain transcription factor GGM13	3 00E-53	putative MADS box protein ZMM17	8.00E-76
701993859	abscisic acid responsive elements-binding factor	3.00E-49	putative bZIP transcription factor ABI5	5 00E-22
701991445	GYMNOS/PICKLE	3 00E-89	chromatin-remodeling factor CHD3	e-110
701995510	Dactylis glomerata somatic embryogenesis related protein mRNA	2.00E-73	No Hit	/
702009335	Ser-Thr protein kinase-like protein	0 00E+00	putative serine/threonine-protein kinase Mak	0 00E+00
702046792	Rice lip19 mRNA for basic/leucine zipper protein.	4 00E-40	basic/leucine zipper protein	4 00E-40
701997623	ethylene-responsive transcriptional coactivator-like protein	4 00E-33	putative ethylene-responsive transcriptional coactivator	2 00E-59
701961663	Incyte EST	No Hit	No Hit	/
702036677	Incyte EST	No Hit	OSJNBb0052B05 14	6 00E-49
701961028	Incyte EST	No Hit	No Hit	/
701969217	mis5 protein	6 00E-19	putative minichromosome maintenance family protein	3 00E-52
702011085	SAR DNA binding protein	2.00E-15	SAR DNA binding protein	2 00E-17
AJ249147	Hordeum vulgare mRNA for MADS-box protein 9 (m9 gene).	3 00E-07		/
702006429	flowering protein CONSTANS, putative, 7571-5495	4 00E-05	putative zinc-finger protein	1 00E-09
701959805	leucine-rich receptor-like protein kinase, putative, 84911-81624	9 00E-18	putative leucine-rich repeat receptor-like kinase	6 00E-32
701995104	transcription factor IIA small subunit	5 00E-04	No Hit	/
703425203	somatic embryogenesis receptor-like kinase 3	2 00E-03	No Hit	/
701988922	Incyte EST	No Hit	No Hit	/
701968396	Incyte EST	No Hit	No Hit	/
701998612	Nicotiana tabacum DNA-binding protein 2 (WRKY2) mRNA.	5 00E-14	WRKY9	1 00E-39
702017431	Hordeum vulgare gene for ids2, complete cds.	7.00E-17	No Hit	/
701995007	Incyte EST	No Hit	unknown protein	7.00E-16
701965813	Incyte EST	No Hit	No Hit	/
701965494	putative protein	1 00E-71	putative WD-40 repeat protein	2 00E-11
702874380	TAR RNA loop binding protein	9 00E-25	No Hit	/
701968565	nucleic acid binding protein	4 00E-15	putative nucleic acid binding protein	9 00E-23
701995191	Incyte EST	No Hit	No Hit	/
702035890	T aestivum mRNA for peptidylprolyl isomerase.	0 00E+00	OSJNBa0091C07.4	2.00E-99
701991275	Oryza sativa secretory carrier membrane protein (SC) mRNA.	5 00E-19	secretory carrier membrane protein	5.00E-19
701968407	AT-hook DNA-binding protein (AHP1)	4 00E-07	putative AT-hook DNA-binding protein	2 00E-14
701966045	cell division related protein-like	3 00E-42	OSJNBa0083D01 10	7 00E-61
702042046	similar to mouse Glt3 or D. malanogaster transcription factor IIB	1 00E-05	No Hit	/

In Table 13, the vast majority of wheat sequences have corresponding orthologues with similar or equivalent function. However, several of the rice orthologues possess annotation that provides additional information about the function of the corresponding wheat gene sequences where there is currently no inferred function. These include, the hypothetical protein sequence which aligns with a PAXNEB protein, an RNA polymerase II elongator protein subunit. Several of the Incyte ESTs (702001685, 701700561 and 701965472) which align with a putative CGI-94 protein, a nuclear transport factor 2 (NTF-2) protein and a putative transposase respectively, and a putative protein that aligns with a WD-40 repeat protein. The additional

information provided by these rice searches highlights the need for further characterisation of the shortlisted sequences. To address this need, the software programs Clustalx and Treeview were used to perform multiple sequence alignments to generate neighbour-joining trees to determine whether particular functions assigned in other organisms may be relevant.

4.6. Phylogenetic Analysis of Gene Shortlist Sequences

The following phylogenetic trees were produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Wheat sequences are labelled by species and gene name, Incyte ID, and previously annotated gene function. For Example, TaEST|702039485|LTP where TaEST = *Triticum aestivum* EST; 702039485 = Incyte clone ID; and LTP = Lipid Transfer Protein. Aligned sequences are labelled by species and gene name, database source, and accession number. For example, OsLTP2|embl|U31766 where OsLTP2 = *Oryza sativa* Lipid Transfer Protein 2; embl = European molecular biology database; U31766 = Accession number. Scale bars indicate the number of nucleotide replacements per site. In general, the trees fall into three main categories: 1) Sequences with confirmed or equivalently inferred function; 2) Sequences with newly inferred function and; 3) Sequences with no additionally inferred function. The majority of sequences fall into the first category where their function is confirmed, and a number of examples are presented in figures 16-22, including sequences that have revealed additional insights into their function and evolutionary conservation. There are only a handful of gene sequences that have generated newly inferred function and are presented in figures 23-27. Finally, a number of sequences fall into the third category either as a result of there being an insufficient number of sequence alignments, or due to the aligned sequences having no annotation from which to infer any functional significance. Examples of these can be seen in figure 28 and figure 29. Because there are too many trees to present all of them in this chapter, a complete list is presented in the appendix and appendix CD. The CD includes folders containing the original .txt files obtained from the BLAST searches, and the corresponding .aln and .phb files created by the multiple sequence alignments and phylogenetic neighbour-joining trees in Clustalx (the software files for Clustalx and Treeview are also supplied in order to view these files on a PC).

4.6.1. Gene Sequences with Confirmed and Equivalent Function

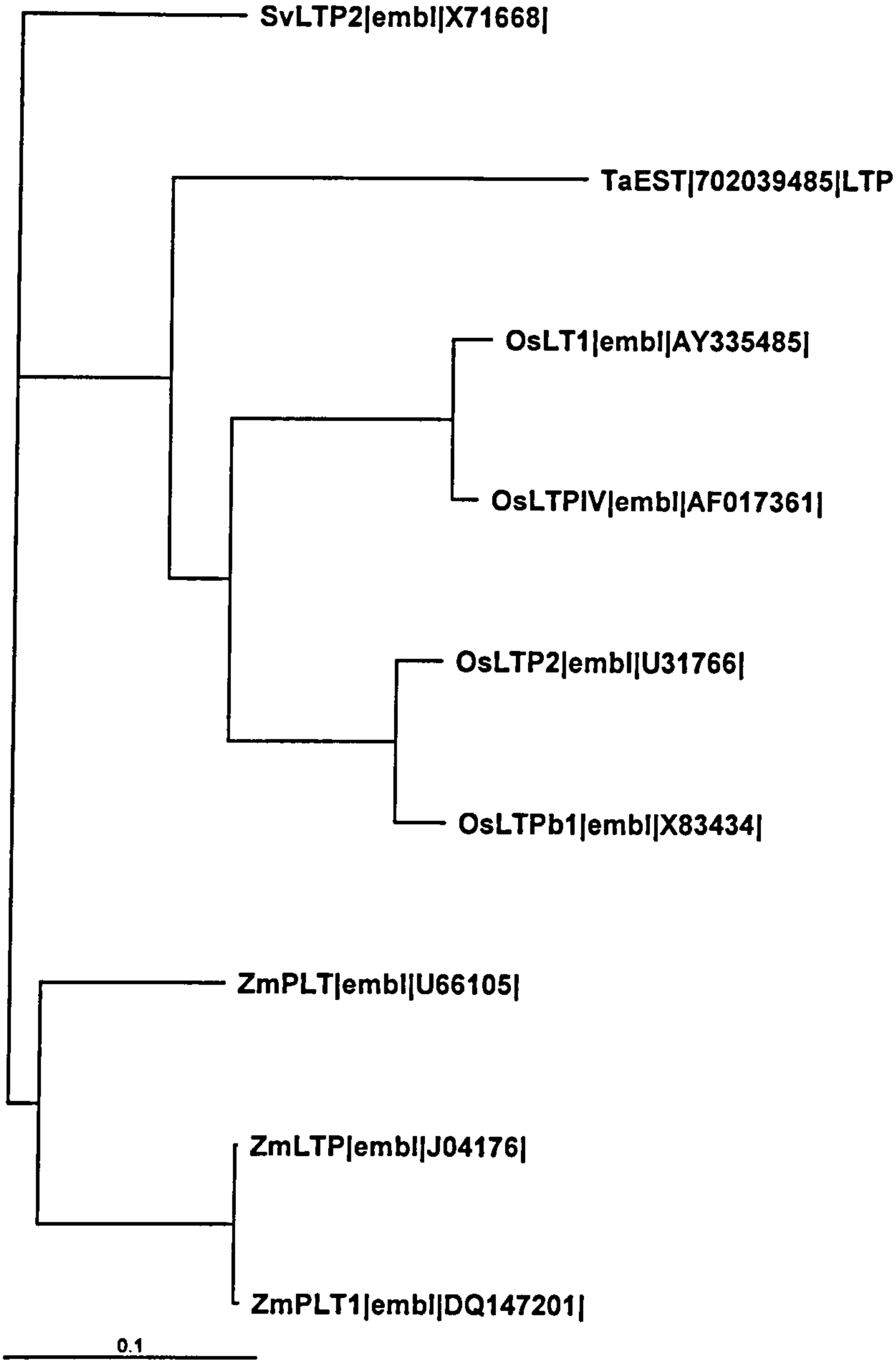


Figure 16: Phylogenetic tree for TaEST|702039485|LTP produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Sv=*Sorghum vulgare;* Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays*; LTP=Lipid Transfer Protein; PLT=Phospholipid Transfer Protein; EST=Expressed sequence tag. Scale bar indicates the number of nucleotide replacements per site.

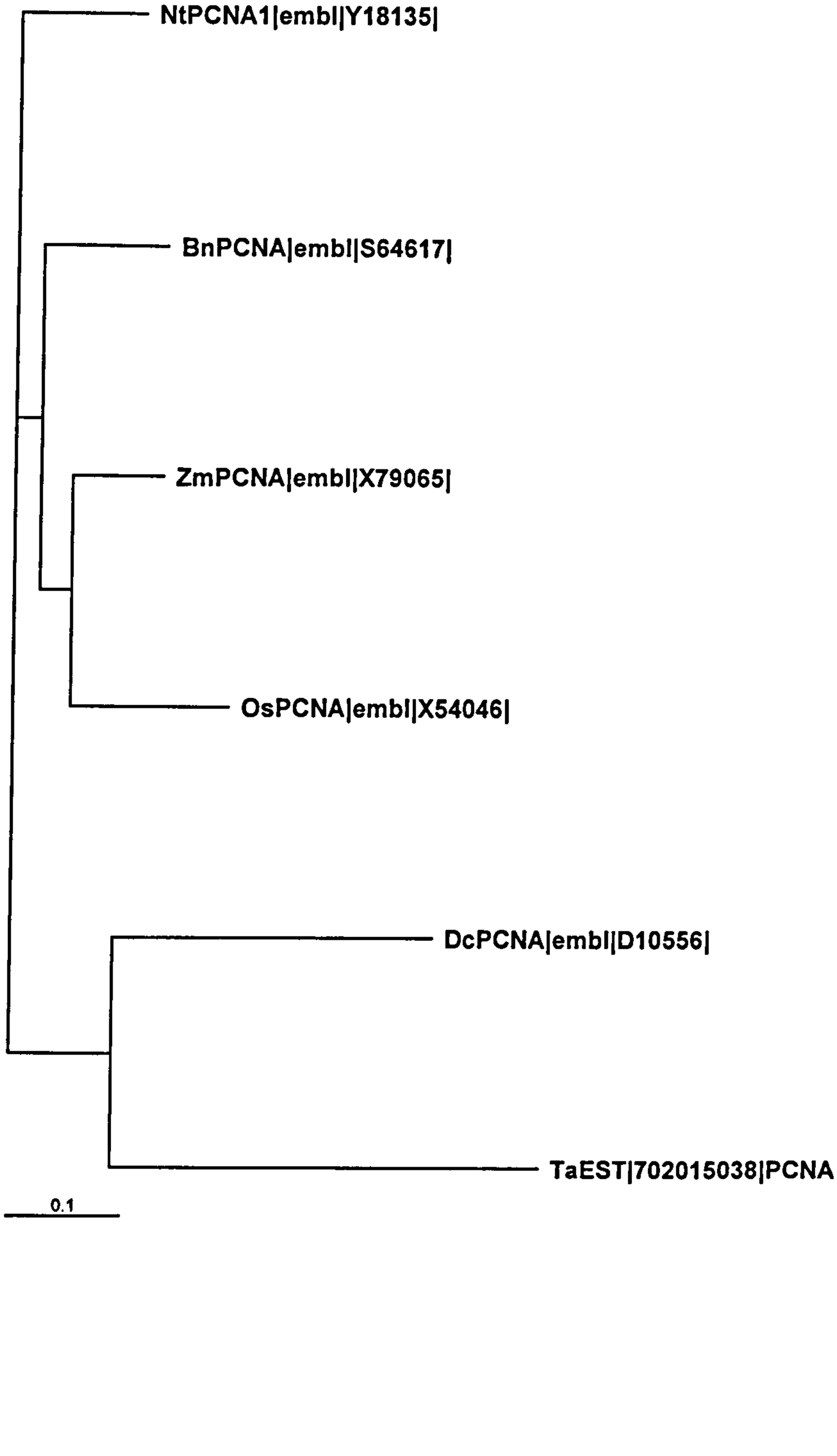


Figure 17: Phylogenetic tree for TaEST|702015038|PCNA produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* Dc=*Daucus carota;* Bn=*Brassica napus;* Nt=*Nicotiana tabacum;* PCNA= Proliferating cell nuclear antigen; EST=Expressed sequence tag. Scale bar indicates the number of nucleotide replacements per site.

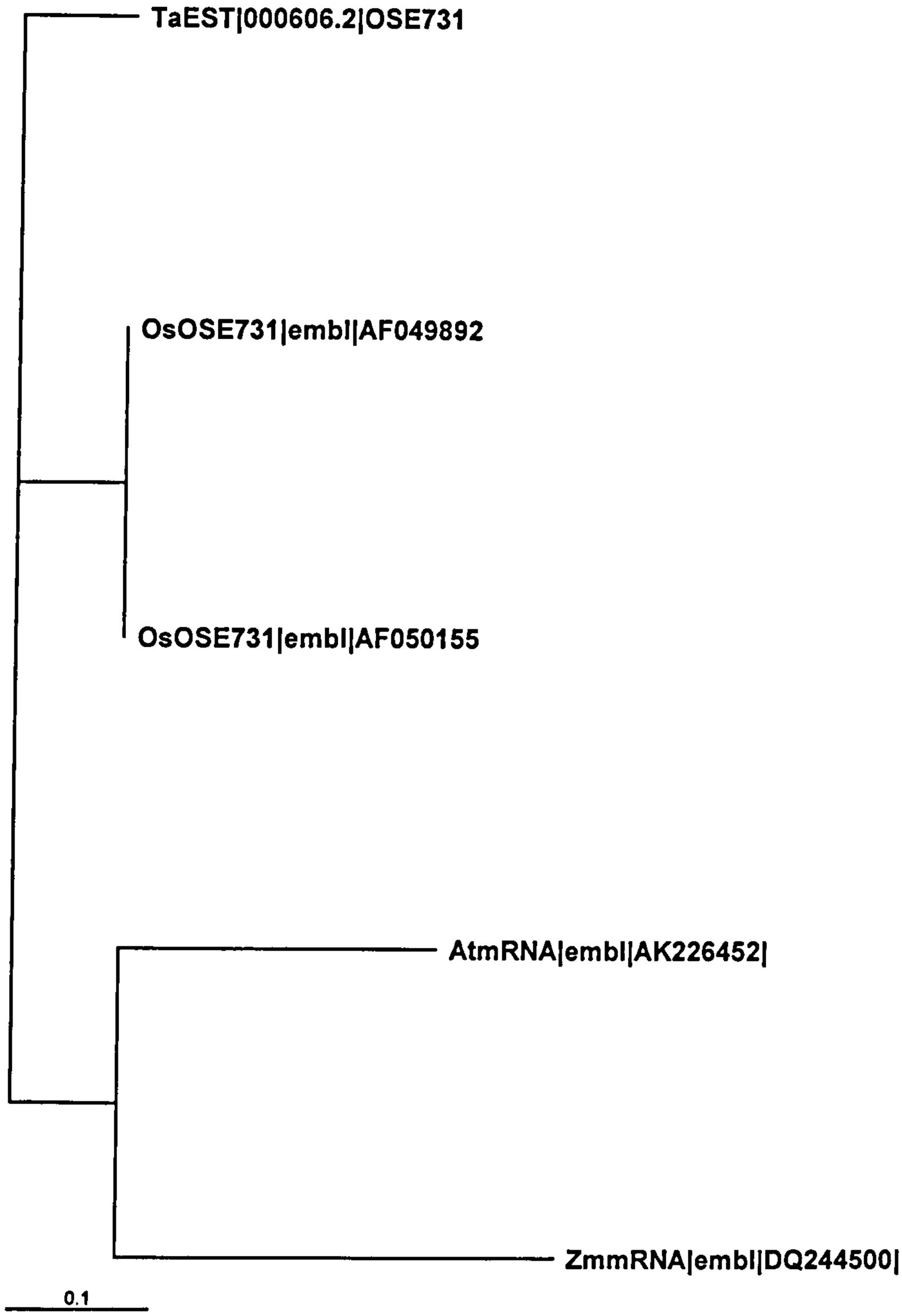


Figure 18: Phylogenetic tree for TaEST|000606.2|OSE731 produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* At=*Arabidopsis thaliana;* OSE731=*Oryza sativa* embryo specific 731 protein; EST=Expressed sequence tag; mRNA=mRNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

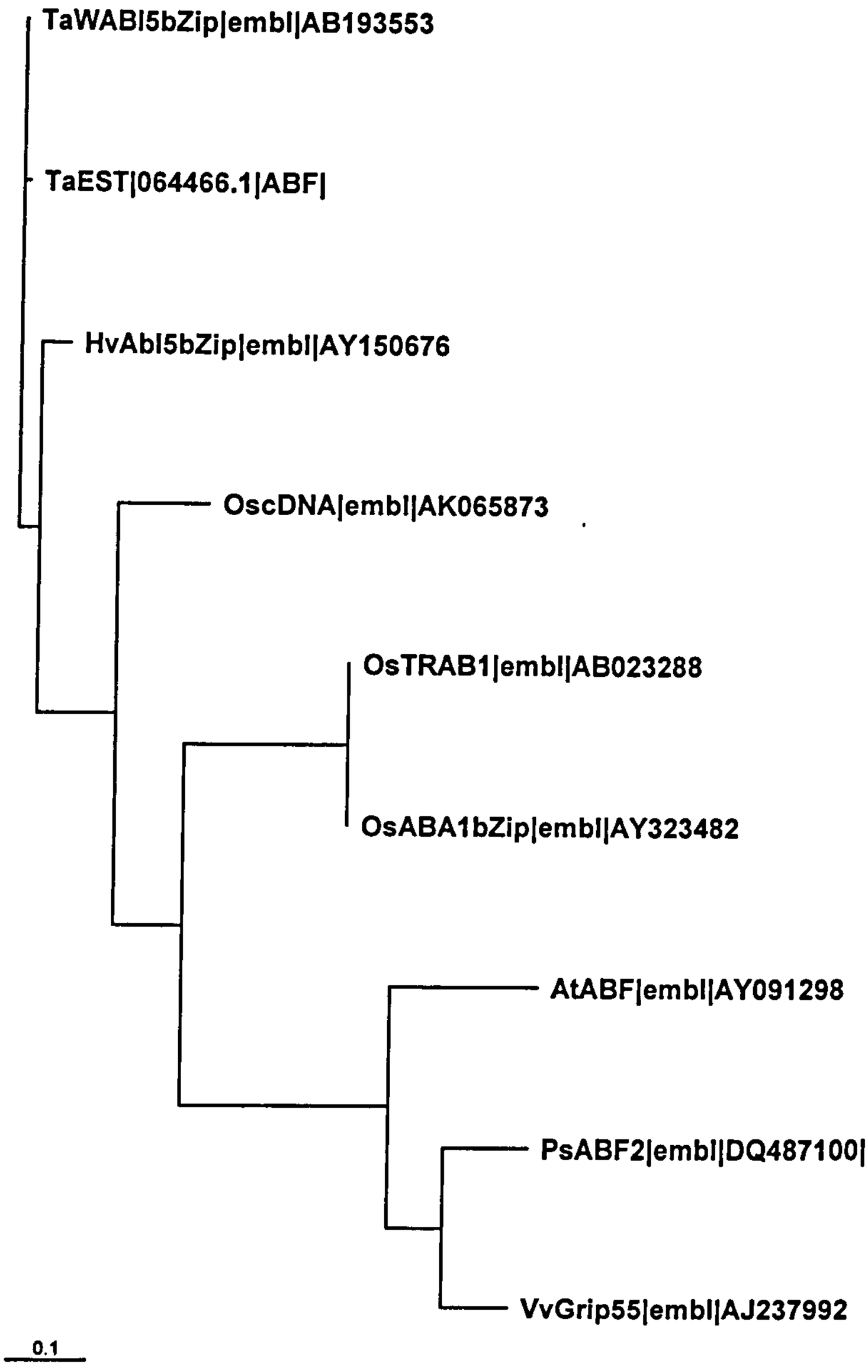


Figure 19: Phylogenetic tree for TaEST|064466.1|ABF produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Hv=*Hordeum vulgare;* At=*Arabidopsis thaliana;* Ps=*Populus suaveolens*; Vv= *Vitis vinifera*; ABI=ABSCISIC ACID-INSENSITIVE; TRAB=ABA-regulated transcription factor; ABA=ABA deficient; ABF=abscisic acid responsive elements-binding factor; Grip=ripening-related bZIP protein; EST=Expressed sequence tag; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

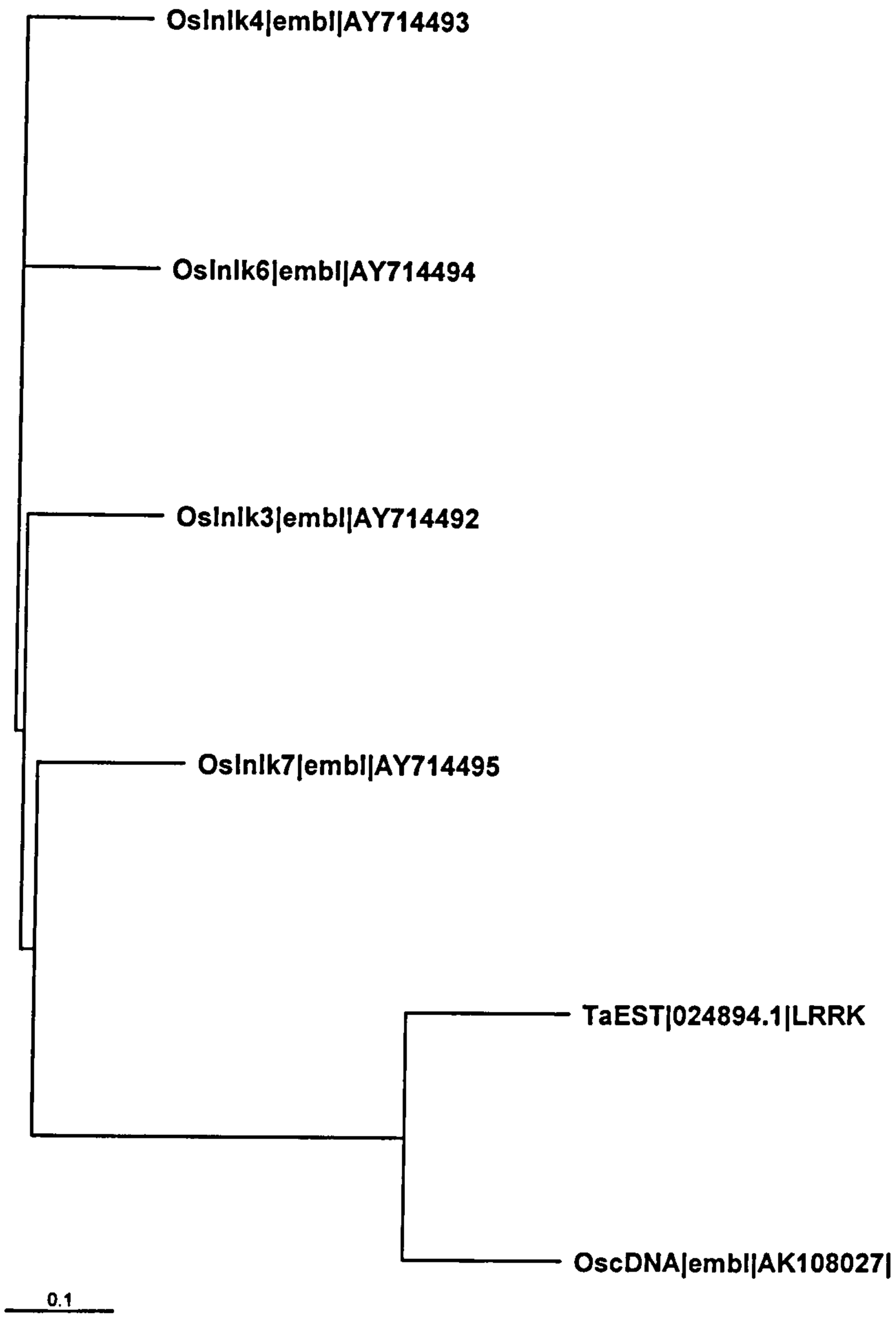


Figure 20: Phylogenetic tree for TaEST|024894.1|LRRK produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* LRRK=Leucine Rich Receptor Kinase; Inlk=putative leucine-rich repeat receptor-like kinase; EST=Expressed sequence tag; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

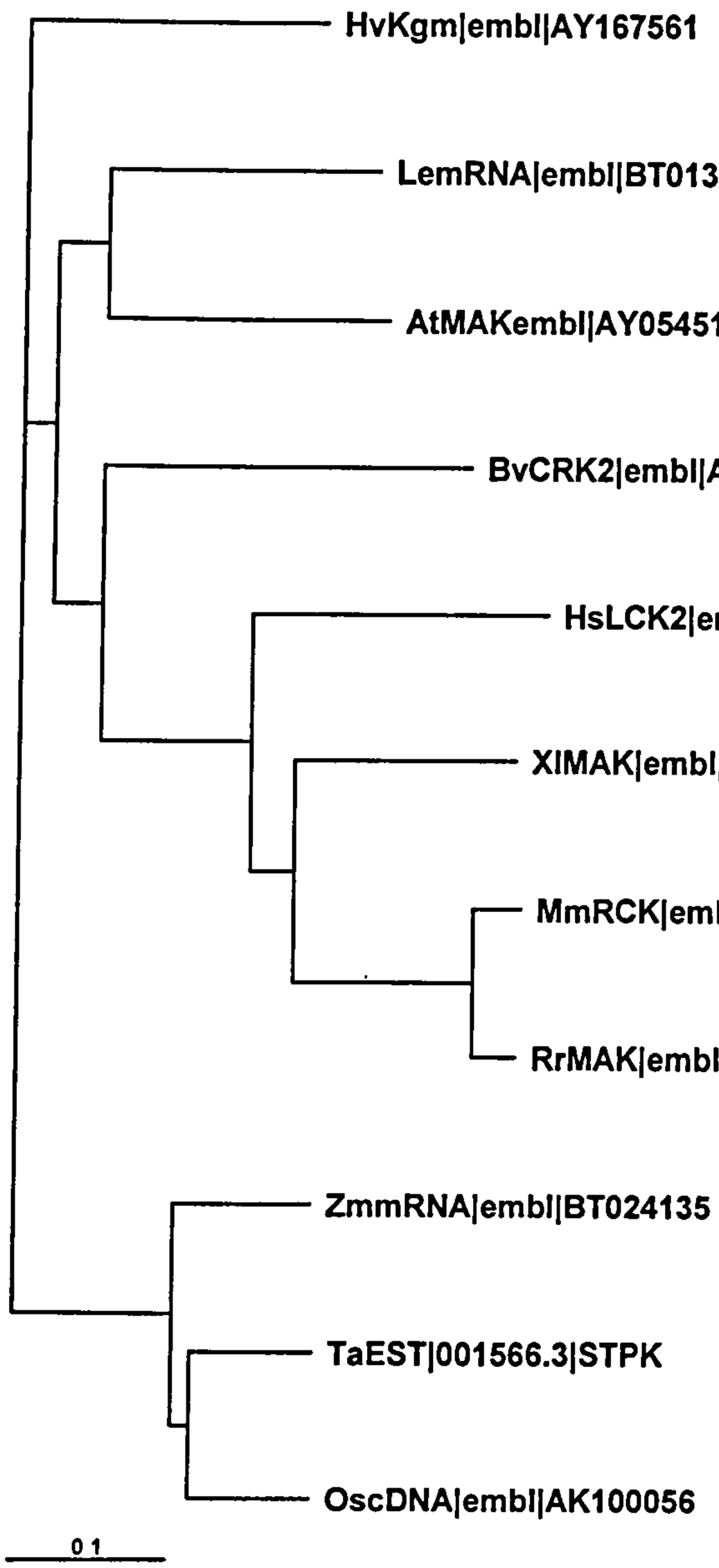


Figure 21: Phylogenetic tree for TaEST|001566.3|STPK produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* Hv=*Hordeum vulgare;* At=*Arabidopsis thaliana;* Bv=*Beta vulgaris;* Le=*Lycopersicon esculentum;* Mm=*Mus musculus;* Xl=*Xenopus laevis;* Hs=*Homo Sapiens;* Dm=*Drosophila melanogaster;* Rr=*Rattus rattus*; GM=GAMYB-binding protein; MAK=male germ cell-associated kinase; CRK= Cdc2-related serine/threonine protein kinases; LCK=laryngeal cancer kinase; RCK=male germ cell protein kinase; EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

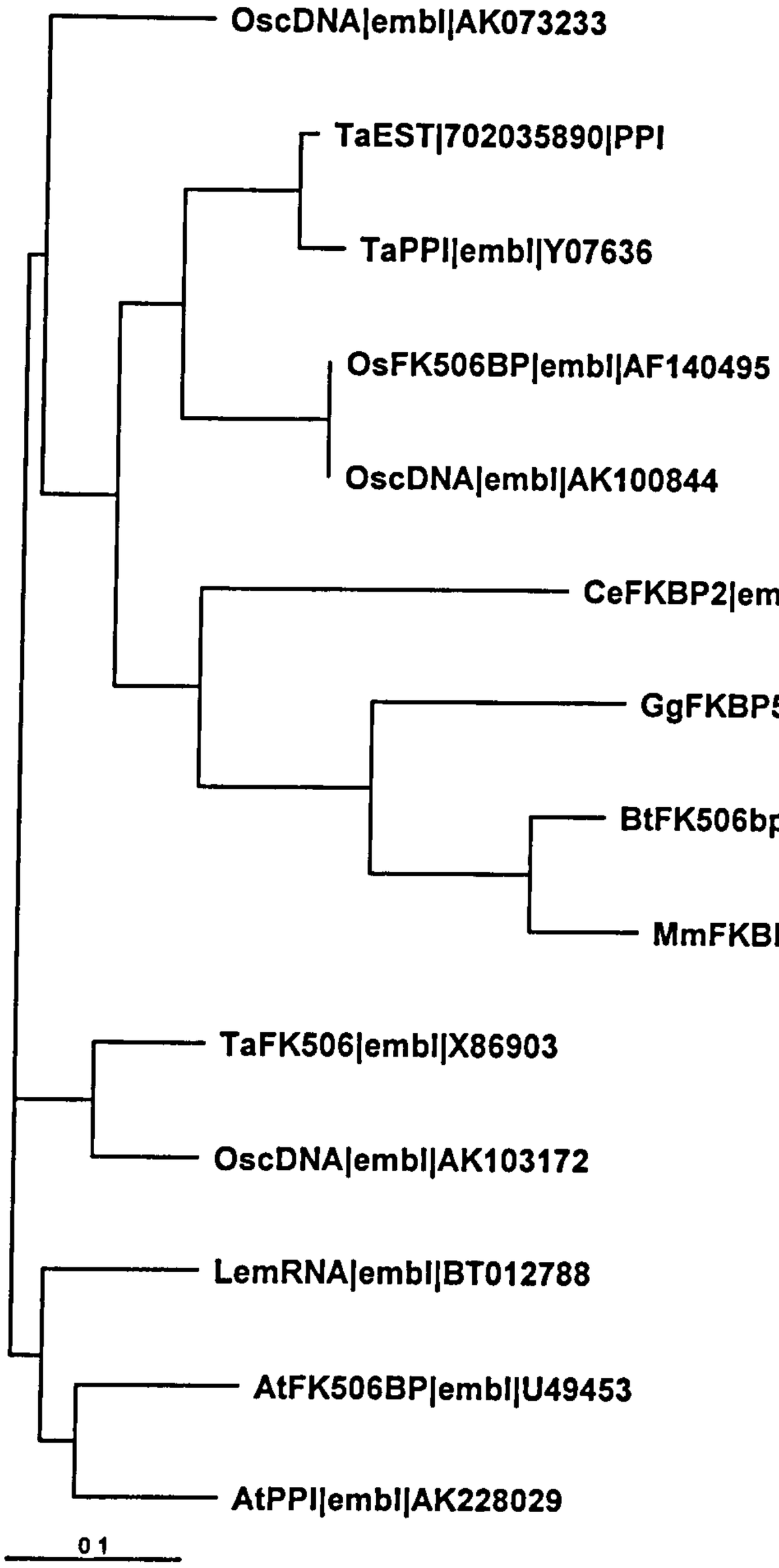


Figure 22: Phylogenetic tree for TaEST|702035890|PPI produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* At=*Arabidopsis thaliana;* Le=*Lycopersicon esculentum;* Mm=*Mus musculus;* Gg=*Gallus gallus;* Ce=*Caenorhabditis elegans;* Bt=*Bos Taurus;* PPI=Peptidyl-Prolyl Isomerase; FK506BP=FK506 Binding Protein; EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

The trees presented above are examples the ISH shortlisted gene sequences whose previously annotated function has been confirmed by the alignment with orthologues that possess the same or equivalent function. The first example presented in figure 16 shows the tree generated from the alignments of the wheat sequence previously identified as a lipid transfer protein (LTP). The wheat sequence groups closely with a number of rice LTPs in a defined clade, clearly demonstrating that this wheat sequence is indeed a LTP. Additional lipid transfer proteins in *S.vulgare* and *Z.mays* suggest that these genes are derived from a common ancestor. Figure 17 shows the tree generated from the alignments of the wheat sequence previously identified as a proliferating cell nuclear antigen (PCNA). The wheat sequence closely aligns with a *D.Carota PCNA* gene and also shares common ancestry with *PCNA* genes from several other plant species. This also clearly demonstrates that this wheat sequence is in fact a *PCNA* gene. Figure 18 shows the tree generated from the alignments of the wheat sequence previously identified as a rice embryo specific protein (OSE731). The wheat sequence groups closest to 2 rice *OSE731* genes, implicating that the wheat sequence is an OSE731 orthologue. Interestingly, the rice and wheat genes share common ancestry with *A.thaliana* and *Z.mays* genes of unknown function suggesting that these unknown sequences may be embryo specific 731 orthologues in their respective organisms. Figure 19 shows the tree generated from the alignments of the wheat sequence previously identified as an ABA responsive elements-binding factor (ABF). This sequence groups closely with the *T.aestivum* and *H.vulgare ABI5* gene and shares common ancestry with a number of ABA-regulated transcription factors from several different plant species confirming that this gene is involved in ABA-regulated transcription and is possibly an orthologue of the *ABA-INSENSITIVE 5* gene. Figure 20 shows the tree generated from the alignments of the wheat sequence previously identified as a leucine rich receptor protein kinase (LRRK) -like gene. The wheat sequence most closely aligns with an *O.sativa* cDNA of unknown function. However, the sequence also shares common ancestry with several genes from the *O.sativa* Inlk family, a putative leucine-rich repeat receptor-like kinase family of proteins. This suggests that the wheat sequence is related to a putative leucine-rich repeat receptor-like kinase but is lacking enough annotation to be conclusive. Figure 21 shows the tree generated from the alignments of the wheat sequence previously identified as a Serine-Threonine protein kinase (STPK)-like protein. The wheat sequence most closely groups with *Z.mays* and *O.sativa* gene sequences of unassigned

function. However, it also shares common ancestry with a number of male germ cell-associated kinase (MAK) gene sequences in both the plant and animal kingdoms. The Mak-type Cdc2-like protein kinases are a relatively uncharacterized group of proteins. However, members of the Mak subgroup are clearly Cdc2-related serine/threonine protein kinases and the expression of the *MAK* gene in sugar beet has been demonstrated to be correlated with cell division activity or cell differentiation (Fowler *et al.*, 2000). Furthermore, this gene also shares common ancestry with a MAK protein kinase which is highly expressed in mammalian testicular germ cells during meiosis (Matsushime *et al.*, 1990). This suggests that the wheat sequence is a STPK like protein that is related to the Mak-type Cdc2-like protein kinases. Figure 22 shows the tree generated from the alignments of the wheat sequence previously identified as a peptidyl-prolyl isomerase (PPI). The wheat sequence groups closely with a wheat PPI and also groups closely with a number of FK 506 binding proteins in a defined clade that includes organisms from both the plant and animal kingdoms. FK506 is a macrolide antibiotic that reduces peptidyl-prolyl isomerase activity by binding to the immunophilin FKBP-12 (FK506 binding protein) and acts as an immunosuppressant by inhibiting T-cell proliferation. This grouping clearly demonstrates that this wheat sequence is a PPI and suggests that it is an evolutionary conserved sequence that plays a role in the regulation of cell proliferation.

4.6.2. Gene Sequences with Newly Inferred Function

Figure 23: Phylogenetic tree for TaEST|702008465|HypPro produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* Le=*Lycopersicon esculentum;* At=*Arabidopsis thaliana;* EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function; elo1=*ELONGATA1* gene. Scale bar indicates the number of nucleotide replacements per site.

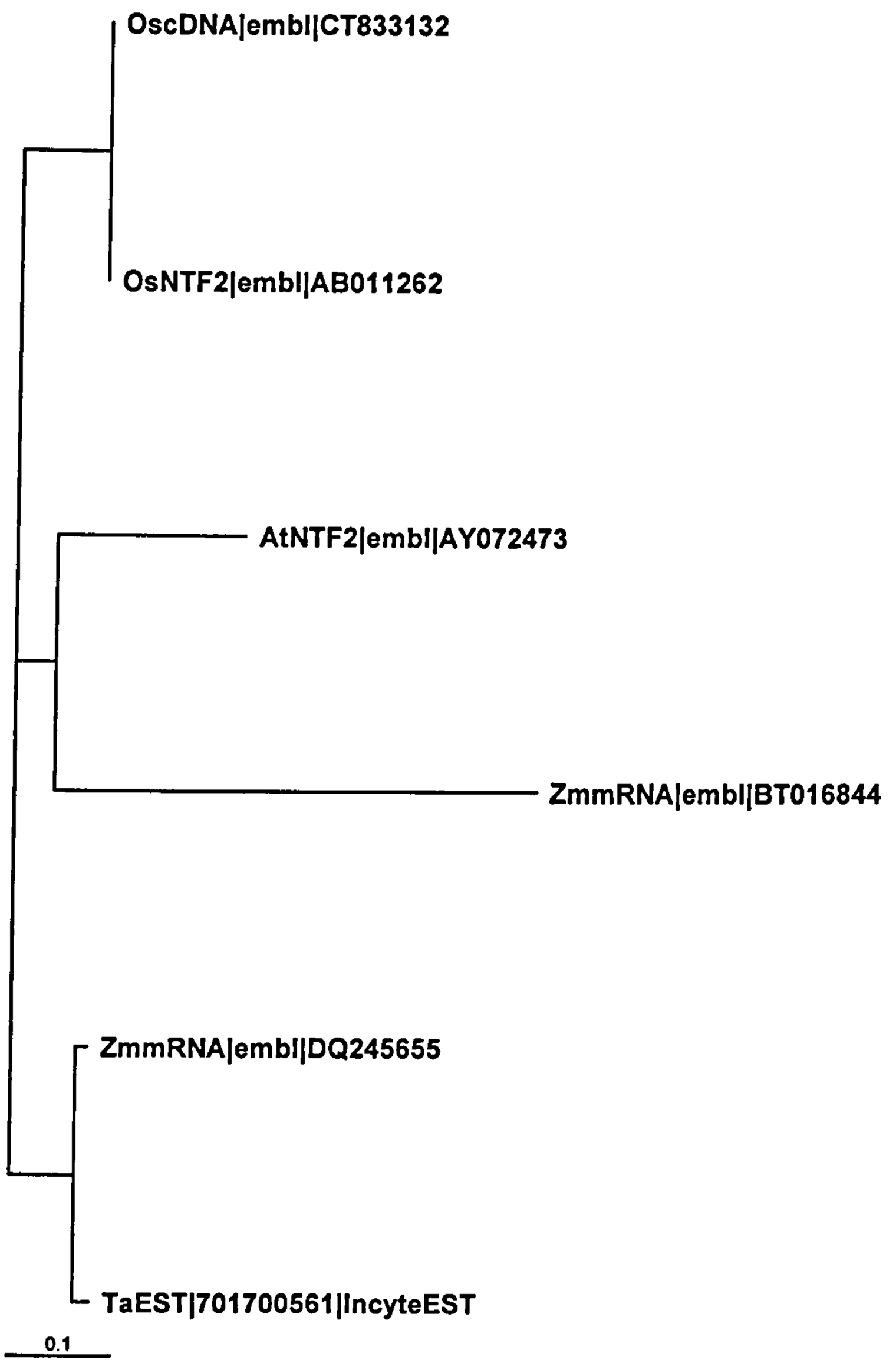


Figure 24: Phylogenetic tree for TaEST|701700561|IncyteEST produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* At=*Arabidopsis thaliana;* NTF2=Nuclear Transport Factor 2; EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

Figure 25: Phylogenetic tree for TaEST|701961663|IncyteEST produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Mm=*Mus musculus;* Hs=*Homo sapiens;* RFP=ring finger protein; EST=Expressed sequence tag; mRNA=mRNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

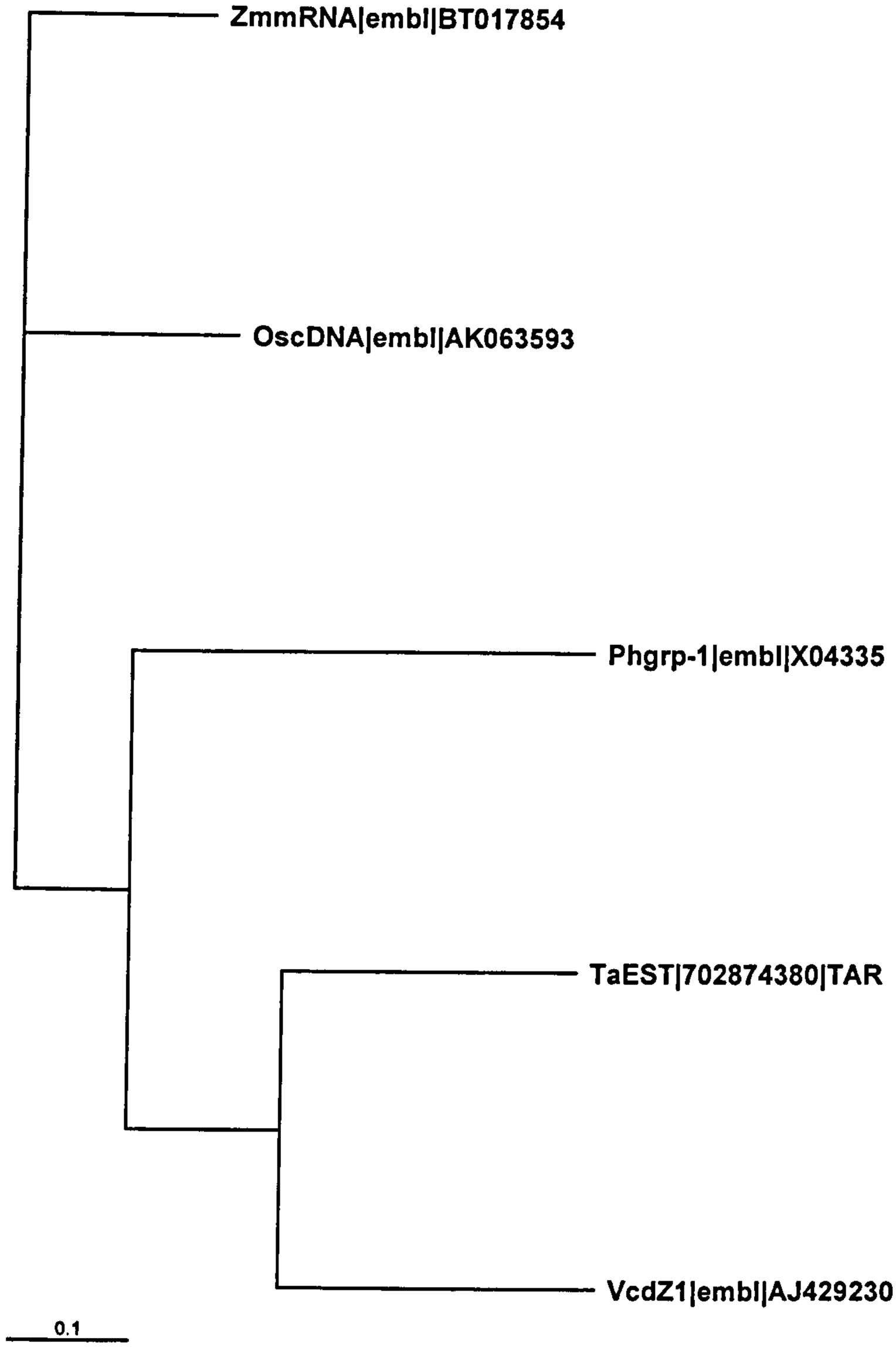


Figure 26: Phylogenetic tree for TaEST|702874380|TAR produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* Ph=*Petunia hybrida;* Vc=*Volvox carteri;* TAR=TAR RNA loop binding protein; dZ1=pherophorin-dz1 protein; grp-1= glycine-rich protein 1; EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

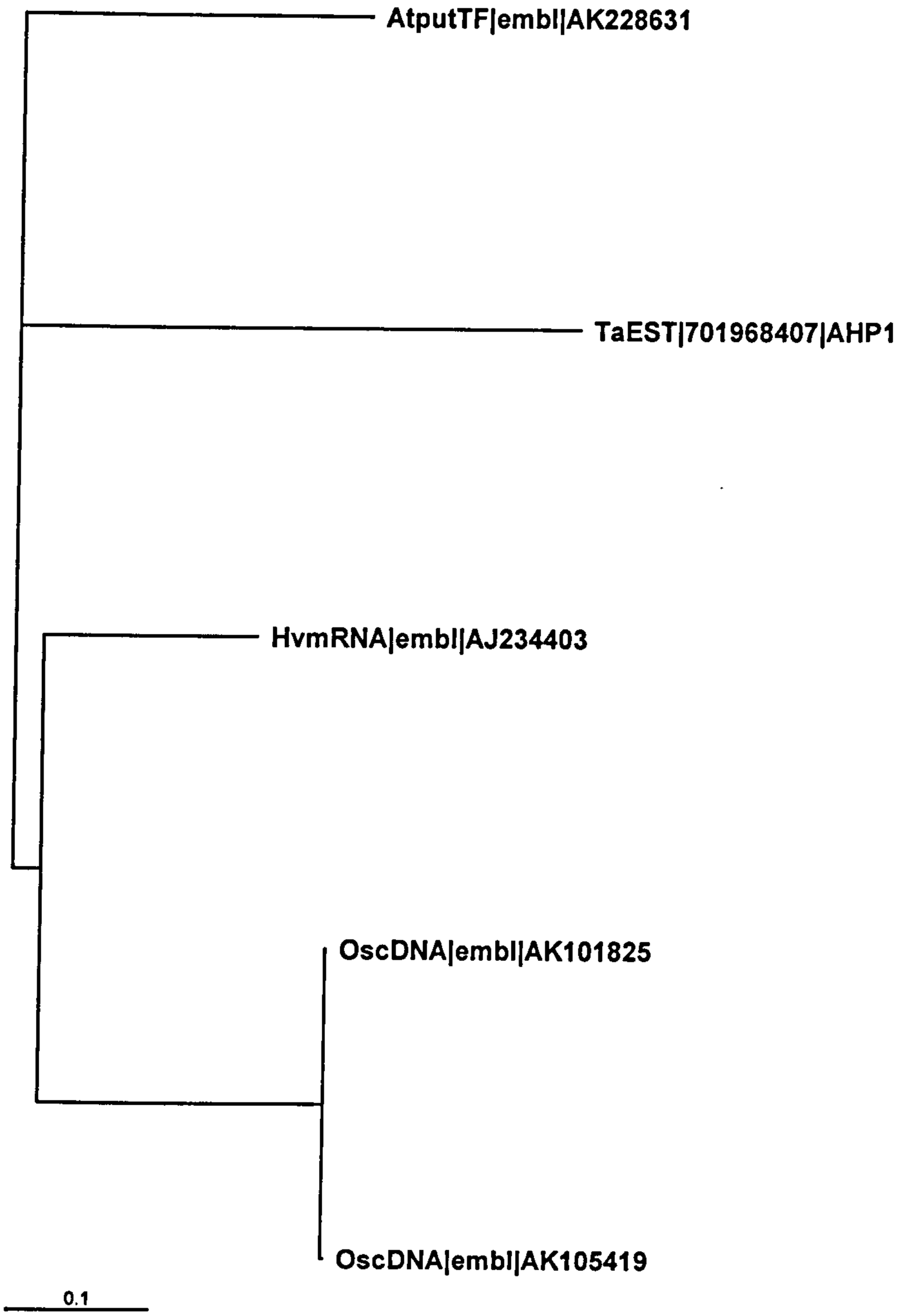


Figure 27: Phylogenetic tree for TaEST|701968407|AHP1 produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Hv=*Hordeum vulgare;* At=*Arabidopsis thaliana;* AHP1=Alkyl HydroPeroxide reductase; PutTF=Putative Transcription Factor; EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

The trees presented above are ISH shortlisted gene sequences whose alignments reveal new functional annotation. The first example presented in figure 23 shows the tree generated from the alignments of the wheat sequence previously identified as a hypothetical protein. The wheat sequence most closely groups with *O.sativa* and *Z.mays* sequences with unassigned function. However, it also shares a common ancestry with an *A.thaliana ELONGATA1* gene which encodes for a subunit of RNA polymerase II. Mutants of this gene have reduced root growth as a result of decreased cell division implicating a role for this gene in cell proliferation during organ growth (Nelissen *et al.*, 2005). Although there is a lack of additional orthologues supporting the annotation for this sequence, it can be inferred as being related to the *ELONGATA1* gene and may play a role in cell proliferation. Figure 24 shows the tree generated from the alignments of the wheat sequence previously identified as a unique Incyte EST. The wheat sequence most closely groups with a *Z.mays* sequence with unassigned function. However, it also shares a common ancestry with an *A.thaliana* and *O.sativa* nuclear transport factor2 (NTF2) sequence implicating this gene as being a wheat orthologue for NTF2. Figure 25 shows the tree generated from the alignments of the wheat sequence also previously identified as a unique Incyte EST. This wheat sequence most closely groups with a *H.sapiens* ring finger protein (RFP) and also shares a common ancestry with a *M.musculus* RFP. Although there are only a small number of aligned sequences with this Incyte EST it is strongly implicated as being an RFP which are known to be mediators of ubiquitin ligase activity. Figure 26 shows the tree generated from the alignments of the wheat sequence previously identified as a TAR RNA loop binding protein (TAR). This wheat sequence most closely groups with a *P.hybrida* glycine rich structural protein and a *V.carteri* pherophorin-dz1 protein, a hydroxyproline-rich glycoprotein involved in extracellular matrix assembly. This strongly suggests that this sequence is actually related to a glycoprotein that may be involved in extracellular matrix assembly and not as a TAR protein as was previously identified. Figure 27 shows the tree generated from the alignments of the wheat sequence previously identified as an alkyl hydroperoxide reductase (AHP1). This wheat sequence most closely groups with an *A.thaliana* putative transcription factor. Although there is a lack of supporting alignments it is possible that this gene is a putative transcription factor rather than an AHP1 as was previously identified.

4.6.3. Gene Sequences with No Additionally Inferred Function

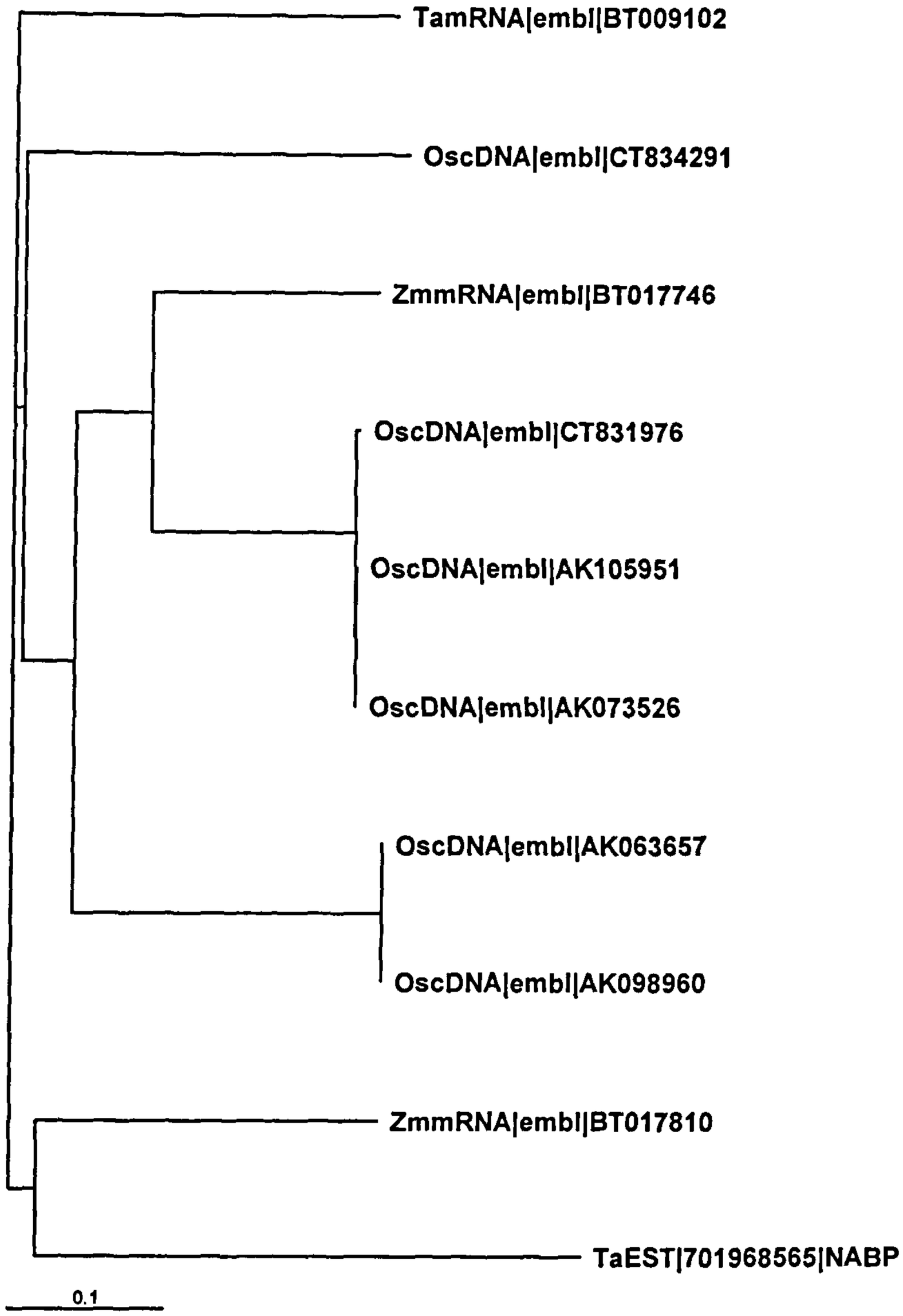


Figure 28: Phylogenetic tree for TaEST|701968565|NABP produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* NABP=Nucleic Acid Binding Protein; EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

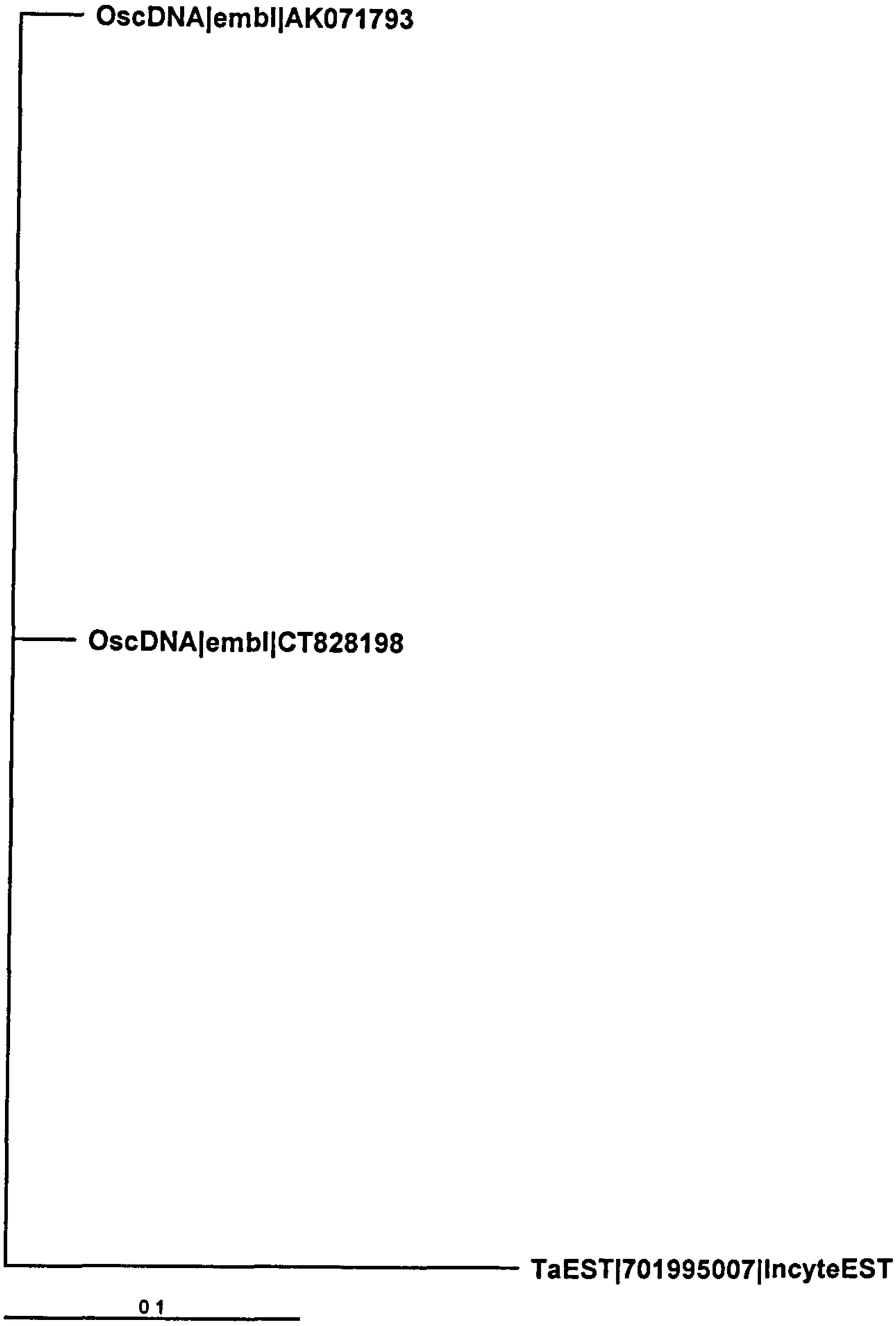


Figure 29: Phylogenetic tree for TaEST|701995007|IncyteEST produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* EST=Expressed sequence tag; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

The trees presented above are examples of ISH shortlisted gene sequences whose alignments do not infer any functional annotation. These are either as a result of there being an insufficient number of sequence alignments, or due to the aligned sequences having no annotation from which to infer any functional significance. The first example presented in figure 28 shows the tree generated from the alignments of the wheat sequence previously identified as a nucleic acid binding protein (NABP). Although there is a good tree produced from a number of different gene sequences there is no corresponding annotation. As a result no functional annotation can be inferred from this tree or from any other sequences that result in trees with similarly poor annotation. Figure 29 shows the tree generated from the alignments of the wheat sequence previously identified as an Incyte EST. As there is a lack of alignments for this sequence it is not possible to perform a bootstrapped neighbour-joining tree. It is therefore not possible to infer any function this sequence or any other sequences that retrieve similarly poor alignments.

Chapter Five

In situ Hybridisation

5 Chapter Five: *In situ* Hybridisation

5.1 Introduction

In situ hybridisation (ISH) is a cytochemical technique originally developed by Gall and Pardue over 30 years ago for localising specific nucleic acid sequences in biological specimens (Gall and Pardue, 1969). Since then the ISH technique has developed into an invaluable research tool in the field of cell and developmental biology as it provides direct visualisation of the spatial expression patterns of specific RNA or DNA sequences at the cellular level which is crucial for the elucidation of gene function. Because ISH takes advantage of the specific annealing of the complementary nucleic acids, in theory any sequence can be studied by constructing an 'anti-sense' labelled probe to the sequence of interest, enabling rapid analysis of the spatial and temporal expression of gene sequences in cells and tissues. Although cDNA microarrays are a powerful tool for identifying gene expression they cannot assign the expression to particular cell types. ISH sufficiently maintains cellular architecture and tissue anatomy so that gene expression profiles can be associated with specific cell types. The rationale behind using ISH for this part of the study is for two reasons: firstly, to confirm the differential expression of the selected set of genes identified from the microarray analysis and secondly, to provide further information concerning the spatial and temporal expression of the identified genes. The ISH approach was chosen for its suitability to serve both purposes as it has a distinctive advantage over other expression detecting methods such as northern blot hybridisation or RT-PCR, *i.e.* it can both confirm gene expression and reveal spatial and temporal information. However, the relatively high number of selected genes makes it difficult to carry out such analysis by conventional ISH methods. To overcome this difficulty, I took advantage of the high-throughput ISH method developed by (Drea *et al.*, unpublished) in Dr John Doonan's laboratory at the John Innes Centre (JIC) in Norwich. This newly developed method allows analysis of a large number of genes simultaneously through semi-automated procedures. Genes with confirmed expression profiles can be identified and analysed to assess the extent to which they are localised within tissues associated with somatic embryo formation and organ differentiation. Those genes whose expression profiles are associated with such tissues can then be selected as candidates for further investigation.

5.2. Materials and Methods

5.2.1. Fixation

Fixation is one of the most critical steps in ISH as poorly fixed material will not generate any signal. Once properly fixed however, tissues can be stored in 70% EtoH at 4^0C for several months without any noticeable reduction in signal intensity. Material for the different developmental timepoints was cultured and collected as described in Chapter 2.2. and placed into 25 ml screw top glass vials containing 4% paraformaldehyde fixative in phosphate buffered saline (PBS). This was prepared fresh each time using 100ml of 1x PBS pH 6.5-7 (Sigma, P4417) adjusted to pH 11 using 10M NaOH solution (Sigma, 8045). In the fume cupboard 4g paraformaldehyde (Sigma, P6148) was weighed out, added to the pre-warmed (60^0C) PBS solution and shaken vigorously until dissolved. If necessary a few drops of 10M NaOH were added to clear the solution. The solution was then cooled on ice and the pH brought back down to 7 using 1M H_2SO_4 (Aldrich, 33974-1). Finally, 100µl (0.1%) Tween®20 (Sigma, P9416) and 100µl (0.1%) Triton®-X-100 (Sigma, T8787) were added to help break down lipids that form non-permeable barriers in plant tissue and cell walls, thereby aiding the penetration of fixative. The vials were placed in a vacuum desiccator and a vacuum applied using a bench top oil pump for approximately 10 minutes to ensure complete penetration of the fixative. The paraformaldehyde solution was replaced with fresh solution after vacuum treatment and left overnight at 4^0C. The following day the paraformaldehyde solution was poured off and the tissue rinsed with ice-cold PBS to remove any residual paraformaldehyde solution. The fixed material was then successively vacuum infiltrated in 50% EtOH and 70% EtOH for 1 ½ hrs each, on ice and stored at -20^0C.

5.2.2. Embedding

For embedding, fixed samples (5.2.1.) were transferred to the Sakura, Tissue Tek Vacuum infiltration processor (VIP) for an automated sequence of dehydration and infiltration steps as outlined in Table 14:

Table 14: Embedding steps in Tissue Tek vacuum infiltration processor (VIP)

Solution	Time	Temperature
80% Ethanol/H20	1.5hrs	35^{0}C
90% Ethanol/H20	2hrs	35^{0}C
100% Ethanol	1hr	35^{0}C
100% Ethanol	1.5hrs	35^{0}C
100% Ethanol	2hrs	35^{0}C
Xylene (Sigma, 29588-4)	0.5hr	35^{0}C
Xylene	1hr	35^{0}C
Xylene	1.5hrs	35^{0}C
100% paraffin wax (Sigma, P3558)	1hr	60^{0}C
100% paraffin wax	1hr	60^{0}C
100% paraffin wax	2hrs	60^{0}C

The tissue was then transferred to the Tissue Tek Embedding Console for embedding. The tissues were orientated using pre-warmed metal moulds and tweezers for longitudinal cross sections on a heated metal plate and then embedded in wax in plastic mounting blocks on ice for approximately ten minutes. The wax embedded tissues were then labelled and stored in a sterile, sealed plastic container at 4^{0}C.

5.2.3. Tissue Sectioning

Silicone isolator chambers (9mm diameter) (Grace Biolabs, JTR8R) were placed onto pre-treated Polylysine slides (BDH) ensuring a water tight seal. Just before sectioning, 2-3 drops of DEPC treated water was applied to each of the 8 chambers on the slide. The wax blocks were cut to a trapezoid shape, so that the longer of the two parallel faces lay at the bottom and was first to strike the blade. Sections of 8μm thick were cut using a Leica Microtome (RM2125RT) and 2-4 sections were floated onto the DEPC treated water in each chamber. For each timepoint enough sections were cut to analyse each probe sequence and placed into the hybridisation oven overnight at 42^{0}C to dry out.

5.2.4. Tissue Pre-Treatment

Slides were placed into the VP2000 slide processor (Vysis) on the pre-treatment programme shown below in Table 15. Dehydrated Slides were removed from the VP2000 slide processor and stored in slide boxes (BDH) at 4^{0}C until required for the hybridisation steps.

Table 15: Tissue pre-treatment steps in VP2000 slide processor.

Solution	Time	Temperature
Xylene	20 min	RT
Xylene	20 min	RT
100% ETOH	10 min	RT
95% ETOH**	2 min	RT
85% ETOH	2 min	RT
50% ETOH	2 min	RT
30% ETOH	2 min	RT
1xPBS	4 min	RT
Proteinase K (Sigma, P2308) (2mg/ml in 100mM Tris 10mM EDTA pH 7.5)	30 min	37^{0}C
Glycine (0.2% in PBS)	2 min	RT
1xPBS	4 min	RT
Acetic anhydride (Sigma, A6404)		
(0.5% in 0.1M triethanolamine pH8) (Sigma, T-9534)	10 min	RT
1xPBS	4 min	RT
30% ETOH	2 min	RT
50% ETOH	2 min	RT
85% ETOH	2 min	RT
95% ETOH**	2 min	RT
100% ETOH	10 min	RT

5.2.5. Preparation and Labelling of RNA Probes

5.2.5.1. DNA Template Preparation

DNA templates for probe preparation were provided as cDNA inserts in modified pSPORT1 vectors (Invitrogen™) with an SP6 RNA polymerase promoter at the 3' end and a T7 RNA polymerase promoter at the 5' end. Due to the inconsistency of transcripts made with SP6, templates were generated by PCR using a forward primer; T7.2 5'GAATTGTAATACGACTCACTATAGGGCCAGTGAATTGAATTTAGG 3' that incorporated a T7 RNA polymerase binding site (underlined) at the 3' end of the insert and a reverse primer; R7.2 5' AGGGAAAGCTGGTACGCCTGC 3' that lay inside the native T7 RNA polymerase binding site. Bacterial cultures containing the EST inserted plasmids were grown overnight at 37^{0}C in a 96 well plate. Ten microlitres of the overnight cultures were added to 90ul of water in a fresh 96 well plate and heated on a thermocycler at 95^{0}C for 2 minutes to lyse the cells and cooled on ice. The denatured DNA templates and T7 primers were added with PCR reaction buffer (1X), dNTPs (0.25mM), taq DNA polymerase (2.5 Units) and DMSO (5%v/v) in 50ul PCR reactions in 96 well plates and run on the following cycle: 94^{0}C for 3 minutes, followed by 30 cycles of 94^{0}C for 45 secs, 63^{0}C for 45 secs 72^{0}C for 90 secs, with a final extension of 72^{0}C for 6 minutes. PCR products were cleaned up using a

'montage 96 well' clean up kit (Millipore, LSKP09601) and checked on an E-gel 96® well agarose gel (Invitrogen, G7008-02).

5.2.5.2. Digoxigenin (DIG)-Labelling via *In Vitro* Transcription

In vitro transcription was carried out in 10µl reactions at 37^{0}C for 2 hrs in 96 well plates using 2 µl DNA template, 1x T7 RNA polymerase buffer, RNasin (20 units) DTT (10mM), 1X NTP labelling mix, T7 RNA polymerase(20 units). After 1 hr, 1µl was removed from each sample and run on a gel to check for transcription.

5.2.5.3. Carbonate Hydrolysis

Following transcription, each reaction was mixed with 10 µl carbonate buffer (200mM, pH10) and incubated at 60^{0}C for 30 minutes to reduce the length of the probes by hydrolysis. The hydrolysis reactions were terminated by the addition of 10µl of 7.5M ammonium acetate (Sigma, A1542) that had been filter sterilised with a Nalgene 0.45 micron filter and the labelled probes precipitated with three volumes (90µl) of ice-cold ethanol. The plates were spun at 4000rpm for 35 minutes to pellet the probes, then inverted and spun at 500rpm for 30 secs followed by a few minutes air-drying to remove the last traces of ethanol before finally dissolving each probe in 30µl TE (100mM Tris, 10mM EDTA).

5.2.5.4. Probe Evaluation

A crude, qualitative evaluation of the probes was carried out by dot blots. One microlitre of each probe (1:100 dilution in DEPC H_2O) was spotted on to a nitrocellulose filter (Hybond, RPN203B) alongside 1:10 and 1:100 dilutions of the supplied standard DIG-RNA control and other selected control experimental probes. The probes were cross-linked to the filter by a UV-transilluminator and the filter soaked briefly in 1x Tris-buffered Saline (TBS) (10mM Tris, 250mM NaCl). The filter was then placed in 1% blocking solution [1g Blocking reagent (Roche, 1096176) in 100ml of 1x TBS] for 30 minutes, 1:5000 dilution of anti-DIG-Alkaline phosphatase (AP) (Roche, 1093274) in 1xTBS for 30-minutes, 1xTBS for 5 minutes, AP buffer (100mM Tris, 100mM NaCl pH9.5, 50mM $MgCl_2$) for 5 minutes and Nitro-Blue Tetrazolium Chloride (NBT) / 5-Bromo-4-Chloro-3'-Indolyphosphate p-Toluidine Salt (BCIP) (2µl/1.5µl per ml of AP buffer) solution for up to 10 minutes to develop.

5.2.6. Hybridisation

Hybridisation solution was made as a master mix (Table16) and aliquoted into the wells of the microtitre plate. Probes were diluted 1:100 in hybridisation solution and mixed well with a multi-channel pipette.

Table 16: Hybridisation solution master mix

Quantity	Ingredient	Final concentration
100µl	10x Salts [3M NaCl, 0.1M Tris-HCL (pH8), 0.1M NaPO4 buffer (pH6.8), 50mM EDTA]	1x Salts
500µl	De-ionised Formamide [50ml of formamide (Sigma F-9037) added to 5g of mixed bed ion-exchange resin (Fluka, 06441) and filtered twice, through Whatman #1 filter paper	50% Formamide
100µl	50% Dextran sulphate [5g Dextran sulphate (Sigma, D-8906) in 5ml H20]	5% Dextran Sulphate
5µl	100 mg/ml tRNA [type XXI tRNA (Sigma R-4251) in H20]	0.5 mg/ml
10µl	100x Dernhardts [5ml H20 added to lypholised powder (Sigma, D9905) containing 50mg of BSA, Ficoll and PVP]	1x Dernhardts
10µl	10 mg/ml salmon testis DNA (Fluka, 31149) in H20	0.1 mg/ml
275µl	H_20	
1000µl		

Hybridisation chambers (Grace Biolabs) were applied to the pre-treated slides, ensuring a tightly formed seal. The probes were denatured by heating to 85^0C for 2 minutes and placed on ice. Forty microlitres of the diluted probes were individually added to each chamber and cover slips placed over the chambers to prevent evaporation. The slides were then placed into the hybridisation oven on metal trays and hybridised overnight at 50^0C.

5.2.7. Washing

Hybridisation chambers were removed by briefly immersing in 2x sodium saline citrate (SSC) (300mM NaCl, 30mM sodium citrate)/50% formamide wash solution and the slides placed in the VP2000 processor and put through the washing programme seen in Table 17.

Table 17: VP2000 processor washing steps

Solution	Time	Temperature
2xSSC/50% formamide (constant agitation)	15 min	40^{0}C
2xSSC/50% formamide (constant agitation)	20 min	50^{0}C
1xSSC/50% formamide (constant agitation)	20 min	50^{0}C
1xSSC (150mM NaCl, 15mM Na3 citrate)	5 min	RT
1xTBS (10mM Tris, 250mM NaCl)	5 min	RT

5.2.8. Antibody Staining and Detection

The slides were removed from the VP2000 and placed face up in plastic slide boxes, 8 slides to a box on a horizontal shaker, and the following solutions were added for staining:

Table 18: VP2000 antibody staining and detection steps

Solution	Time	Temperature
1xTBS containing 1% blocking reagent	1Hr	RT
1xTBS containing 1:3000 dilution of Anti-Dig AP and 0.05% Tween-20	1Hr	RT
1xTBS	10 min	RT
1xTBS	10 min	RT
1xTBS	10 min	RT
1xTBS	10 min	RT
AP-buffer	5 min	RT

The slides were developed in AP buffer containing NBT (0.1mg/ml) and BCIP (0.075mg/ml) in slide boxes that had been sealed with a lid and parafilm and were placed in the dark for approximately 16 hrs. To stop the reaction, slides were washed several times in water and briefly rinsed in 70% EtOH and 100% EtOH to remove any excess staining. Slides were then air dried and mounted in 2-3 drops of Entellan (Merck, 362134A). The slides were air dried overnight in a fume cupboard ready for viewing under a microscope.

5.2.9. Microscopy and Visualisation

Mounted slides were examined under a Nikon E800 microscope and brightfield digital images taken with a Nikon Coolpix digital camera. For easy comparison and consistency, sections of different time points were examined and the images taken under the same microscope and camera settings.

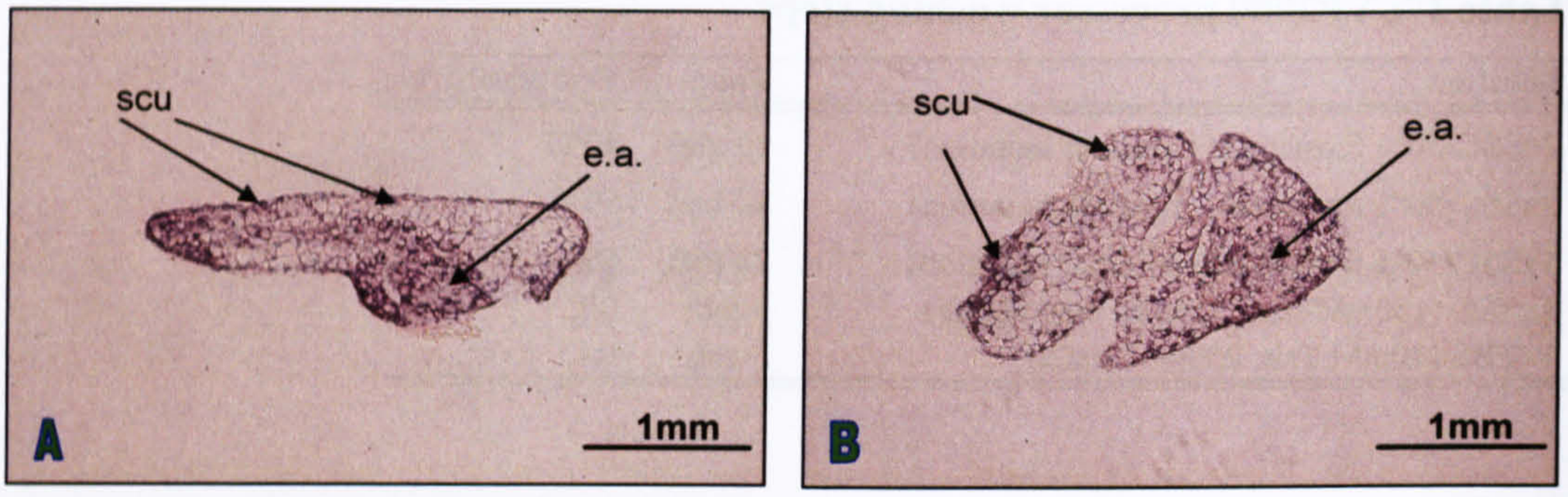


Figure 30 (A-D). Embryogenic BW calli at various stages (Tp1-4) showing the bound positive control probe, Histone 4. Immature embryos were fixed after the following treatments A) TP1= 7 Hrs 2,4-D Treatment, B) TP2 = 4 Days 2,4-D treatment, C) TP3 = 7 Days 2,4-D treatment, D) TP4 = 7 Days 2,4-D treatment + 1 Day MS. Tissue sections were cut 8um thick. The strong probe signal in individual cells demonstrates a typical cell cycle-like pattern of expression and confirms the integrity of the fixed material. *e.a.* = embryonic axis, *scu* = scutellum, *emb* = embryoid body, *u.c.* = unorganised cells, *c.d.c.* = cytoplasmically dense cells. Bar = 1mm

5.3. Results

The isolation of wheat embryos from their caryopses and the exposure to high auxin (2,4-D) causes a disruption to their normal developmental programme which is characterised by several aspects. Firstly, there is the suppression of germinative growth from the embryonic axis. Secondly, there is the loss of cellular specification. Thirdly, there is the rapid onset of cellular division. This can lead to the reorganisation of proliferating cells into ordered embryonic structures or the development of unorganised cell masses. This disruption to immature wheat embryo development can distort embryonic anatomy after 3-4 days making it difficult to describe anatomical structures. However, the ISH images have been annotated to highlight the early anatomical details of immature embryos as well as cytoplasmically dense proliferating cells, embryonic structures and unorganised cell regions that develop in embryogenic calli. The use of several positive controls and a negative control allow for the determination of a true signal from the experimental gene probes and ensures confidence in the expression of these genes in embryogenic cells and tissues. The ISH experiments were conducted using sections of BW calli only as there was no readily available source of SAV material and insufficient time to produce the material for the ISH experiments. Consequently, there is a lack of a comparable non-embryogenic control which would provide conclusive confirmation that a gene expressed in embryogenic tissue is not expressed in non-embryogenic tissues and therefore specifically required for SE. Nevertheless, the ISH results do provide confirmation of the expression of the cDNA microarray identified gene sequences and reveal further information about their localised expression patterns within the tissues they were identified and an insight into their potential functional roles in SE.

5.3.1. Histone 4 Positive Control

Histone 4 is one of four histones that form the eukaryotic nucleosome core which are complexed to DNA in chromatin and chromosomes and are synthesised during DNA replication in cells undergoing division. The expression of histone 4 is therefore indicative of cells that are active and dividing and provides a useful positive control that not only demonstrates the integrity of the fixed tissue but also provides information about the nature of cell proliferation within the developing embryogenic

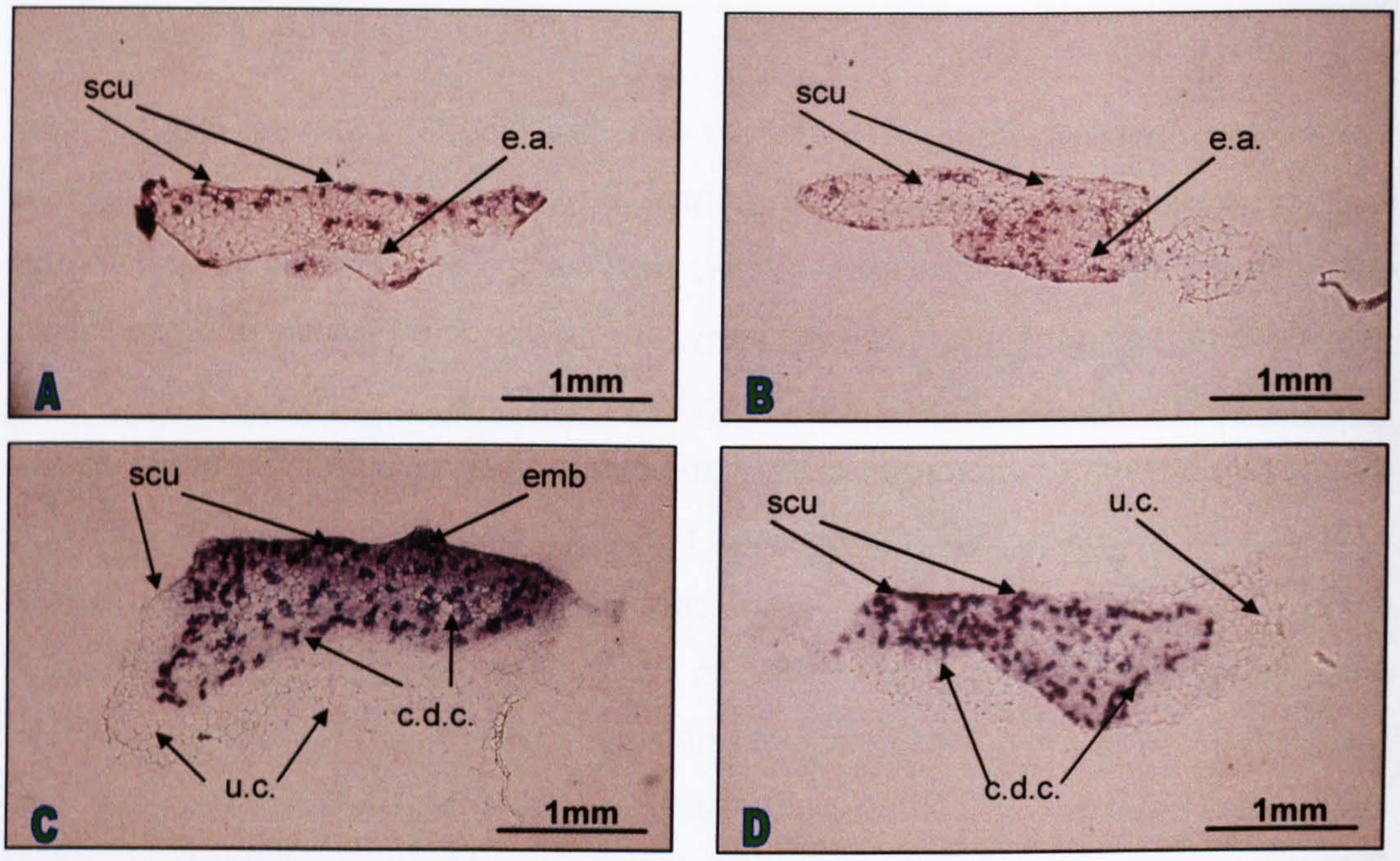


Figure 31 (A-D). Embryogenic BW calli at various stages (Tp1-4) showing the bound positive control probe, Cyclin-D. Immature embryos were fixed after the following treatments A) TP1= 7 Hrs 2,4-D Treatment, B) TP2 = 4 Days 2,4-D treatment, C) TP3 = 7 Days 2,4-D treatment, D) TP4 = 7 Days 2,4-D treatment + 1 Day MS. Tissue sections were cut 8um thick. The strong probe signal in individual cells demonstrates a typical cell cycle-like pattern of expression and confirms the integrity of the fixed material. *e.a.* = embryonic axis, *scu* = scutellum, *emb* = embryoid body, *u.c.* = unorganised cells, *c.d.c.* = cytoplasmically dense cells. Bar = 1mm

callus tissues. Figure 30 shows a uniformed distribution of the H4 signal in timepoints 1 + 2 (A + B) demonstrating that the tissue is actively undergoing cell division and proliferation. In timepoints 3 + 4 (C + D) the signal is more localised to regions of the callus in which there are cytoplasmically dense, actively dividing cells. It can also be seen that the histone 4 ISH signal is localised to individual cells which is typical of cell cycle regulated genes.

5.3.2. Cyclin-D Positive Control

Cyclin-D is also used as a positive control as it is expressed in all proliferating cell types. It is involved in the progression of cells through the cell cycle and is essential to cell division and is therefore also indicative of cells that are active and dividing. Figure 31 demonstrates that there is an ISH signal in each of the four timepoints which is localised to individual cells in a typical cell cycle like manner. The expression of cyclin-D is particularly strong in timepoints 3 + 4 (C + D) with timepoint 3 providing a particularly good example of a developing embryogenic calli. The ISH signal indicates the regions within the calli that are undergoing rapid cell proliferation with particularly strong expression in the upper layer of the scutellar tissue. Embryoid bodies (emb) can be seen to develop from the scutellum and appear to have a particularly dense signal. A clear contrast can be seen between the proliferating and non proliferating regions of the callus with a complete lack of signal in the regions of calli that are immediately in contact with the media. These cells appear more vacuolated than their cytoplasmically dense counterparts and are typically responsible for the generation of unorganised callus tissue.

5.3.3. Poly-T Positive Control

The Poly-T probe is also used as a positive control to show a general ISH signal in each cell as the majority of mRNA molecules contain mRNA tails. This control in particular allows one to check that the pre-treatment and hybridisation steps have made all of the cells penetrable for probe hybridisation and provides a reference for the potential localised expression for mRNA species in the treated tissues at each timepoint. The presence of an ISH signal at each timepoint therefore demonstrates that the tissue has been fixed and treated properly and that the mRNA species have maintained sufficient integrity as to allow for the hybridisation and signalling of the complementary cDNA probes. Figure 32 (A+B) illustrates that the expression of

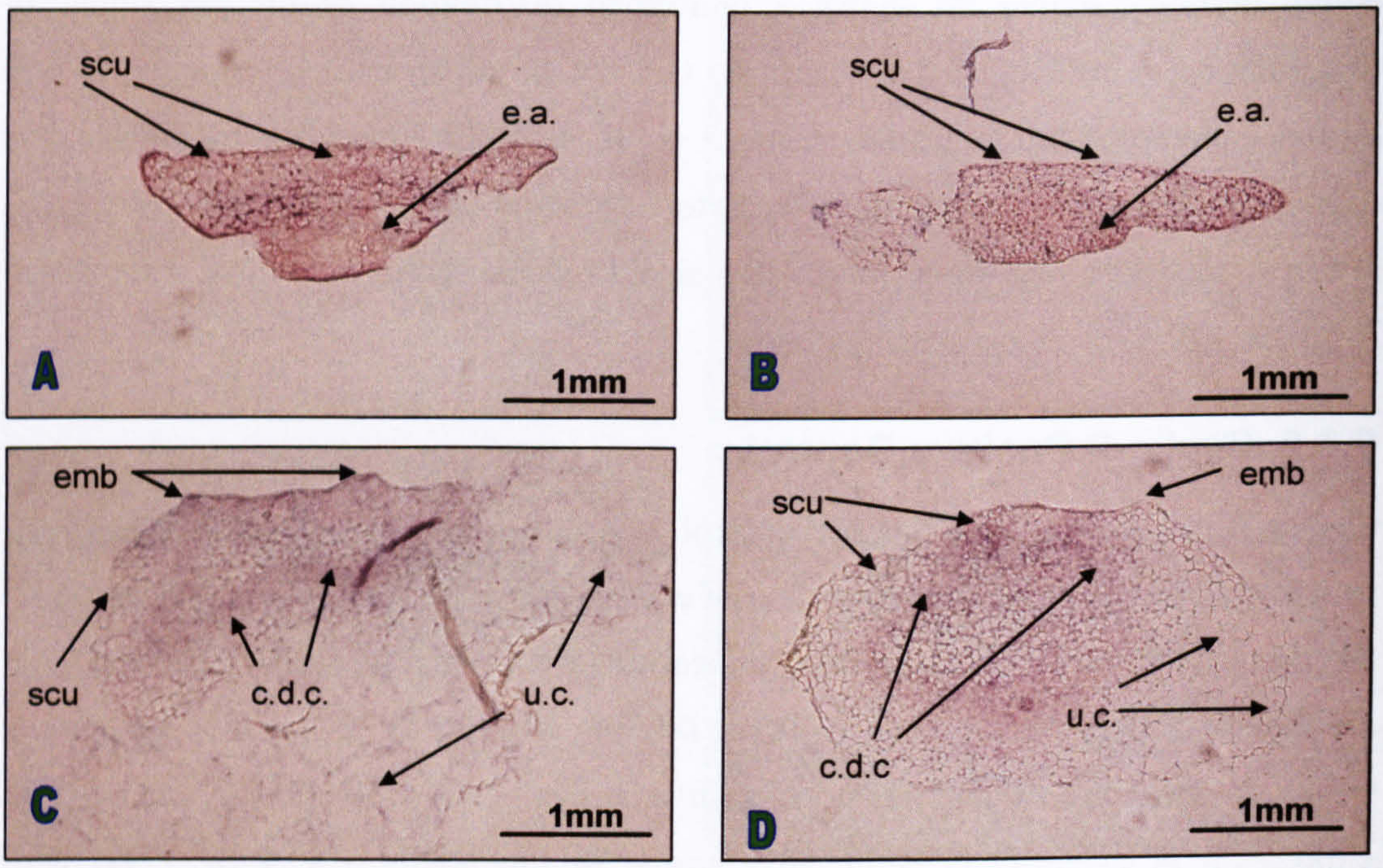


Figure 32 (A-D). Embryogenic BW calli at various stages (Tp1-4) showing the bound positive control probe, Poly-T. Immature embryos were fixed after the following treatments A) TP1= 7 Hrs 2,4-D Treatment, B) TP2 = 4 Days 2,4-D treatment, C) TP3 = 7 Days 2,4-D treatment, D) TP4 = 7 Days 2,4-D treatment + 1 Day MS. Tissue sections were cut 8um thick. The poly-T control demonstrates the pre-treatment and hybridisation steps have made all of the cells penetrable for probe hybridisation and provides a reference for the localised expression for mRNA species. *e.a.* = embryonic axis, *scu* = scutellum, *emb* = embryoid body, *u.c.* = unorganised cells, *c.d.c.* = cytoplasmically dense cells. Bar = 1mm

mRNA species in timepoints 1 and 2 is fairly ubiquitous and uniform, whereas the poly-T signal at timepoints 3 and 4 (Fig.18 C+D) appears to be restricted to the cytoplasmically dense, actively dividing regions of the calli highlighting the importance of gene expression within these regions during embryogenic callus formation and SE.

5.3.4. Alpha Thionin Negative Control

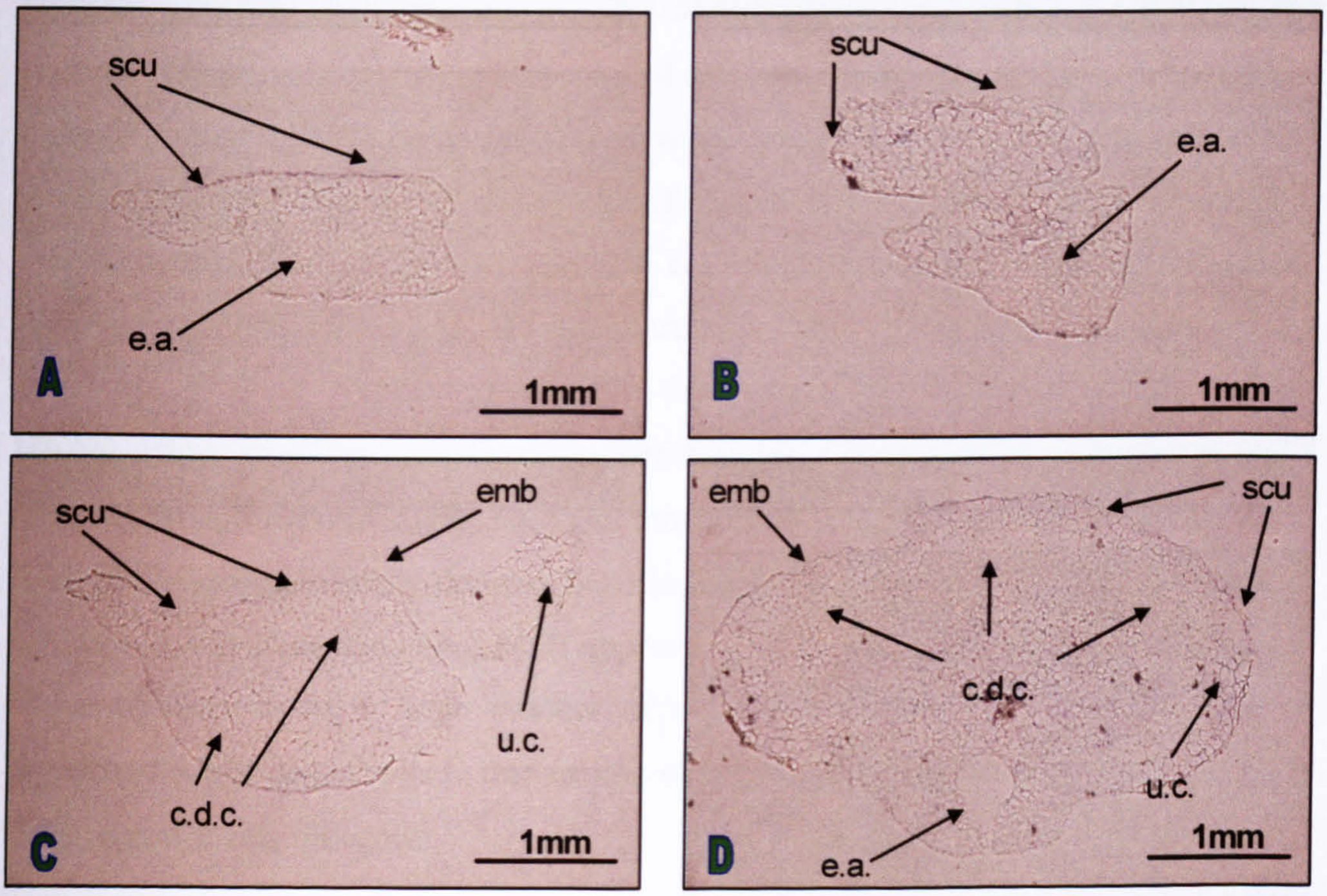


Figure 33 (a-d) Embryogenic BW calli at various stages (Tp1-4) showing the bound negative control probe, alpha-thionin. Immature embryos were fixed after the following treatments A) TP1= 7 Hrs 2,4-D Treatment, B) TP2 = 4 Days 2,4-D treatment, C) TP3 = 7 Days 2,4-D treatment, D) TP4 = 7 Days 2,4-D treatment + 1 Day MS. Tissue sections were cut 8um thick. The lack of ISH signal allows the assessment of background signalling from non-specific binding. *e.a.* = embryonic axis, *scu* = scutellum, *emb* = embryoid body, *u.c.* = unorganised cells, *c.d.c.* = cytoplasmically dense cells. Bar = 1mm

Alpha thionin is a wheat endosperm specific storage protein and provides a useful control for assessing the level of background signal that may be caused by random or non-specific probe binding. It allows one to differentiate between non specific background signalling and a genuine ISH signal of bound experimental probe. (Fig. 33 A-D) clearly demonstrates the uniformly distributed low level background signalling in each of the different timepoints and tissues. This allows for the interpretation of genuine gene expression with a high degree of confidence among candidate genes that demonstrate an ISH signal.

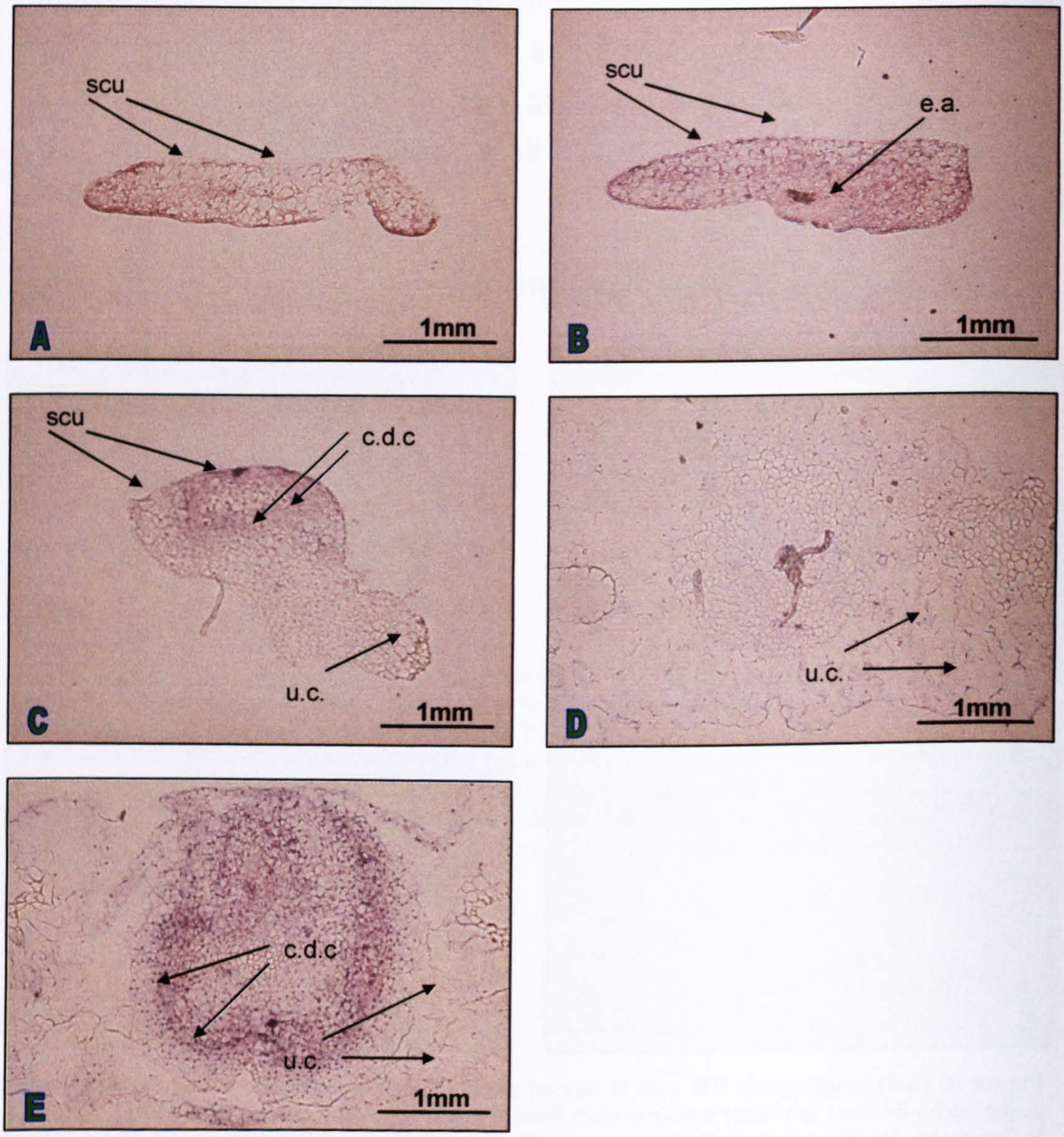


Figure 34 (A-E). Embryogenic BW calli at various stages (Tp1-4) showing the bound experimental probe, dual specificity protein phosphatase (DsPTP1)-like. Immature embryos were fixed after the following treatments A) TP0 = untreated, B) TP1 = 7 Hrs 2,4-D Treatment, C) TP2 = 4 Days 2,4-D treatment, D) TP3 = 7 Days 2,4-D treatment, E) TP4 = 7 Days 2,4-D treatment + 1 Day MS. Tissue sections were cut 8um thick. The ISH signals confirm the general differential expression patterns identified in the microarrays, that there is an increase in the *in situ* signal at time points 1, 2 and 4 (B,C+E) *.e.a.* = embryonic axis, *scu* = scutellum, *emb* = embryoid body, *u.c.* = unorganised cells, *c.d.c.* = cytoplasmically dense cells. Bar = 1mm

5.3.5. Contains Similarity to DsPTP1- like Protein

The first experimental probe presented in the results is the dual specific phosphatase (DsPTP1) that was identified in the microarray hybridisations as being differentially expressed at timepoints 1 (7Hrs 2,4-D), 2 (4 days 2,4-D) and 4 (7 days 2, 4-D, 1 day MS) at values of 4, 2.7, and 2.6 respectively. Figure 34 shows that the ISH signals confirm the general differential expression patterns identified in the microarrays. That is, that there is an increase in the *in situ* signal at time points 1, 2 and 4 (Fig. 34 B,C+E). It is important to note here that although ISH is a sensitive technique for detecting gene expression and can confirm the expression of a gene identified from cDNA microarray experiments, it is difficult to quantify the level of differential expression between samples and as such can only be used in a qualitative capacity. There is a clear ISH signal in the untreated sample timepoint 0 (Fig.34A) and timepoint 1 (Fig.34B) that appears to be fairly uniform and ubiquitous with a small noticeable increase in the ISH signal in timepoint 1 compared to timepoint 0. The ISH signal in timepoint 2 (Fig.34C) and timepoint 4 (Fig.34E) appears to be expressed in specific cytoplasmically dense cellular regions of the developing calli. Timepoint 3 (7 days 2,4-D treatment) (Fig.34D) appears to lack a genuine ISH signal, however, there appears to be a large amount of unorganised callus tissue and a lack of cytoplasmically dense cells in this section which may be responsible for the lack of an ISH signal at this timepoint.

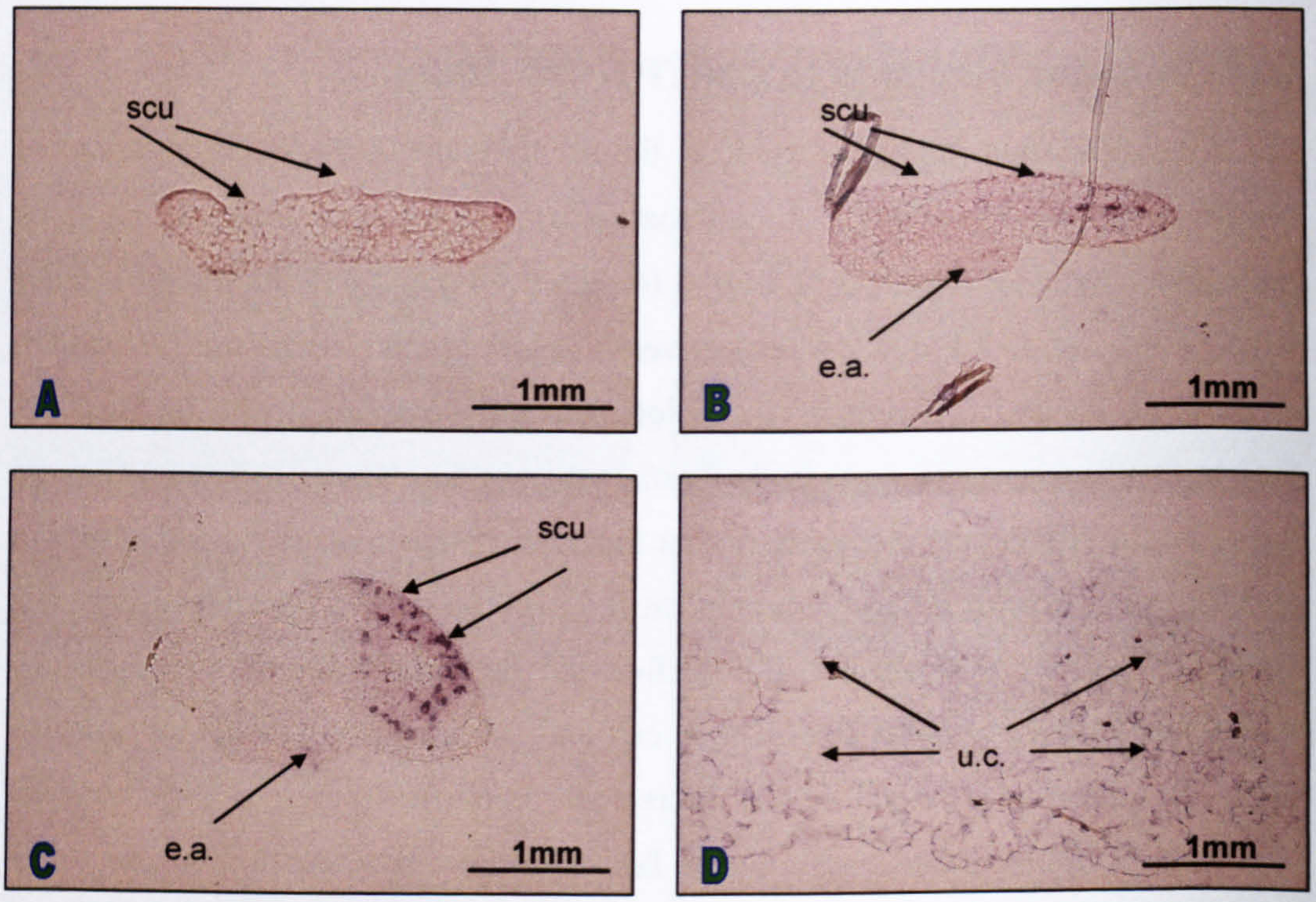


Figure 35 (A-E). Embryogenic BW calli at various stages (Tp1-4) showing the bound experimental probe, Abscisic Acid Responsive Elements-Binding Factor (ABF). Immature embryos were fixed after the following treatments A) TP0 = untreated, B) TP1 = 7 Hrs 2,4-D Treatment, C) TP2 = 4 Days 2,4-D treatment, D) TP3 = 7 Days 2,4-D treatment, E) TP4 = 7 Days 2,4-D treatment + 1 Day MS. Tissue sections were cut 8um thick. The spatial and temporal expression of this gene is particularly interesting as it is not expressed in untreated immature embryo samples (A) but appears to be expressed in a small number of individual cells during TP1 (B) and increases in expression during TP2 (C) and TP4 (E) *.e.a.* = embryonic axis, *scu* = scutellum, *u.c.* = unorganised callus, *c.d.c.* = cytoplasmically dense cells. Bar = 1mm

5.3.6. Abscisic Acid Responsive Elements-Binding Factor

The spatial and temporal expression of the second experimental probe, an ABA responsive elements binding factor (ABF), is particularly interesting as it does not appear to be expressed in untreated immature embryo samples in timepoint 0 (Fig. 35 A), but appears to be switched on and expressed in a small number of individual cells after only seven hours 2,4-D treatment during timepoint 1 (Fig. 35 B). The expression of this gene appears to be localised to a specific layer of cells within the scutellar tissue. The number of cells expressing this gene increases after four days 2,4-D treatment in timepoint 2 (Fig. 35 C) which continue to be localised to the scutellar regions of the developing callus, in particular at the surface layers of the scutellum. A large number of cells continue to express this gene after seven days 2,4-D treatment plus one day MS in timepoint 4 (Fig. 35 E). However, this gene does not appear to be expressed at all after seven days 2,4-D treatment in timepoint 3. This may be due to the quality of the particular section in question which appears to be largely comprised of unorganised cells and lacks cytoplasmically dense regions in which typical gene expression occurs. This gene was identified from the microarray hybridisations as being only differentially expressed after four days 2,4-D treatment during timepoint 2, however, the results from the *in situ* experiments clearly reveal the expression of this gene at timepoints 1 and 4 (Fig. 35 B+E) The overall spatial expression of this gene is indicative of a cell cycle like regulated pattern of gene expression and is similar to those observed in the histone 4 and cyclin D controls. However, what makes this gene a particularly interesting candidate for further analysis is that it is not observed in untreated cells and appears to be specifically induced by the application of 2,4-D after only a few hours.

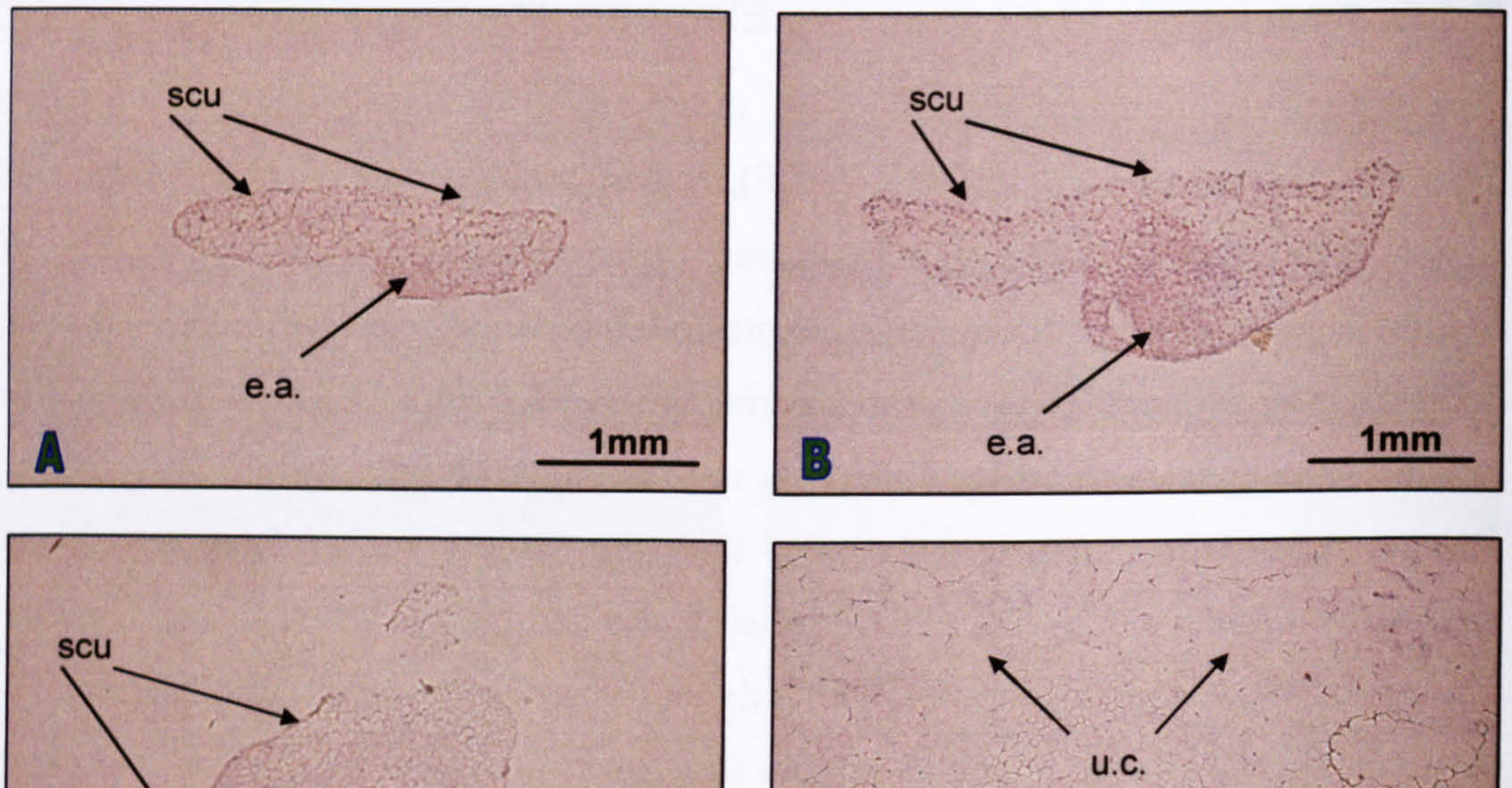


Figure 36 (A-E). Embryogenic BW calli at various stages (Tp1-4) showing the bound experimental probe, *Oryza sativa* embryo specific (Ose 731) mRNA. Immature embryos were fixed after the following treatments A) TP0 = untreated, B) TP1 = 7 Hrs 2,4-D Treatment, C) TP2 = 4 Days 2,4-D treatment, D) TP3 = 7 Days 2,4-D treatment, E) TP4 = 7 Days 2,4-D treatment + 1 Day MS. Tissue sections were cut 8um thick. The spatial and temporal expression of this gene is interesting because it is significantly up-regulated in central regions of the calli after removal from the 2,4-D media (E) Furthermore the expression of this gene rapidly drops off after removal from the 2,4-D media (E).*e.a.* = embryonic axis, *scu* = scutellum, *u.c.* = unorganised cells, *c.d.c.* = cytoplasmically dense cells. Bar = 1mm

5.3.7. *O.sativa* Embryo Specific Protein (OSE731)

The next experimental probe *Oryza sativa* embryo specific (OSE731) mRNA was identified during the microarray hybridisations as being expressed after four days 2,4-D in timepoint 2 at a value of 3.7. Although there is some signal detected after seven hours 2,4-D treatment in timepoint 1 (Fig. 36 B) within the cells that comprise the embryonic axis, the ISH results fail to detect any significant signal at timepoint 2 (Fig. 36 C). The reason for this may simply be due to the particular section in question being derived from an incorrectly fixed tissue, although this is unlikely due to the presence of a signal at this timepoint in the positive controls. A more likely explanation would be that the gene is expressed in a specific, localised area of the callus that is not represented in the particular section in question which may result in the gene not appearing to be expressed. Interestingly, a significant signal is detected in calli that have been removed from the 2,4-D medium in timepoint 4 (Fig. 36 E). Taken alone, the ISH experiments suggest that this gene is expressed at a low level in the embryonic axis of immature embryos, which is consistent with its annotation, and is only significantly induced in calli on removal of 2,4-D from the tissue culture medium. Therefore this gene forms another interesting candidate for further analysis as it appears to be specifically induced by the removal of 2,4-D from the tissue culture medium and is likely to play a role in the early events that lead to the regeneration of adult structures from somatic embryos and embryogenic calli.

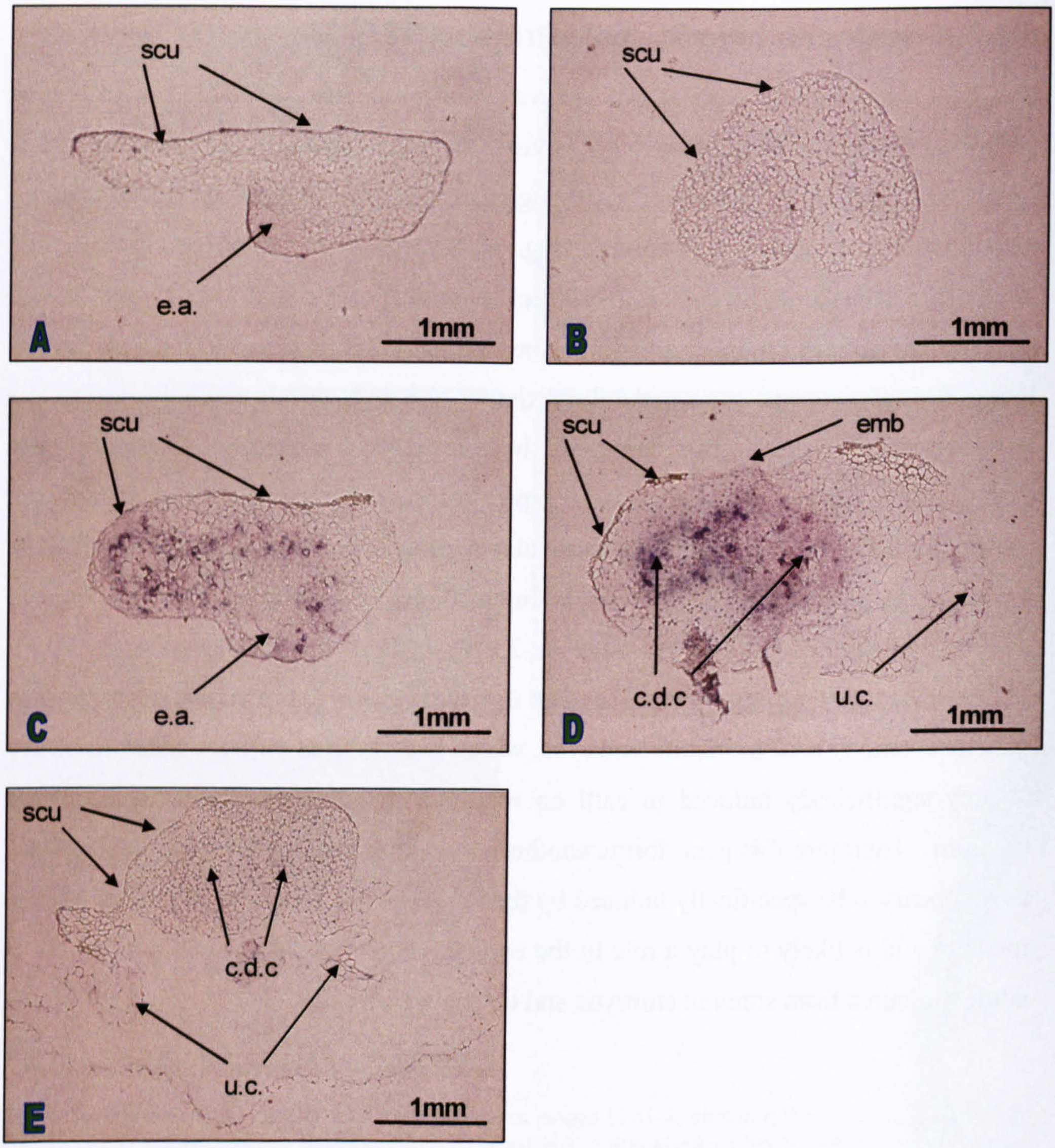


Figure 37 (a-e). Embryogenic BW calli at various stages (Tp1-4) showing the bound experimental probe, leucine rich receptor kinase (LRRK). Immature embryos were fixed after the following treatments A) TP0 = untreated, B) TP1 = 7 Hrs 2,4-D Treatment, C) TP2 = 4 Days 2,4-D treatment, D) TP3 = 7 Days 2,4-D treatment, E) TP4 = 7 Days 2,4-D treatment + 1 Day MS. Tissue sections were cut 8um thick. The spatial and temporal expression of this gene is interesting as it is not expressed in untreated immature embryo samples (A) but appears to be expressed in central regions of the calli after 4 days (C) and 7 days 2,4-D treatment (D) Furthermore, the expression of this gene rapidly drops off after removal from the 2,4-D media (E). *e.a.* = embryonic axis, *scu* = scutellum, *u.c.* = unorganised cells, *c.d.c.* = cytoplasmically dense cells, *emb* = embryoid body. Bar = 1mm

5.3.8. Leucine-Rich Receptor-Like Protein Kinase (LRRK)

The next experimental probe presented in Fig. 37 A-E is the leucine-rich receptor-like protein kinase (LRRK) which also possesses an interesting pattern of expression. The LRRK gene was identified in the cDNA microarray experiments as being differentially expressed after seven days 2,4-D treatment (TP3) which is confirmed by the ISH results (Fig. 37 D) as well as revealing the signal at timepoint 2 after four days 2,4-D treatment (Fig. 37 C). It can be seen that there is no detectable signal in the untreated samples at timepoint 0 (Fig. 37 A) or after seven hours 2,4-D treatment at timepoint 1 (Fig. 37 B). However, after four days 2,4-D treatment in timepoint 2 (Fig. 37 C) and seven days treatment in timepoint 3 (Fig. 37 D) there is significant up-regulation of this gene in the embryogenic calli. There is a similar pattern of expression at both of these timepoints for this gene which appears to be localised in a central region of the developing calli rather than at the surface of the scutellum where somatic embryo-like structures usually emerge. The fact that there is no expression of this gene in untreated samples suggests that this gene is specifically induced as a result of 2,4-D treatment. However, as there is also no expression after seven hours 2,4-D treatment it is likely to be regulated by earlier molecular events that occur upstream of the LRRK pathway. Interestingly, there is no signal detected after twenty-four hours removal of the 2,4-D media, suggesting that this gene is specifically and transiently expressed during the establishment of embryogenic calli and becomes quickly down regulated on the removal of 2,4-D from the tissue culture medium.

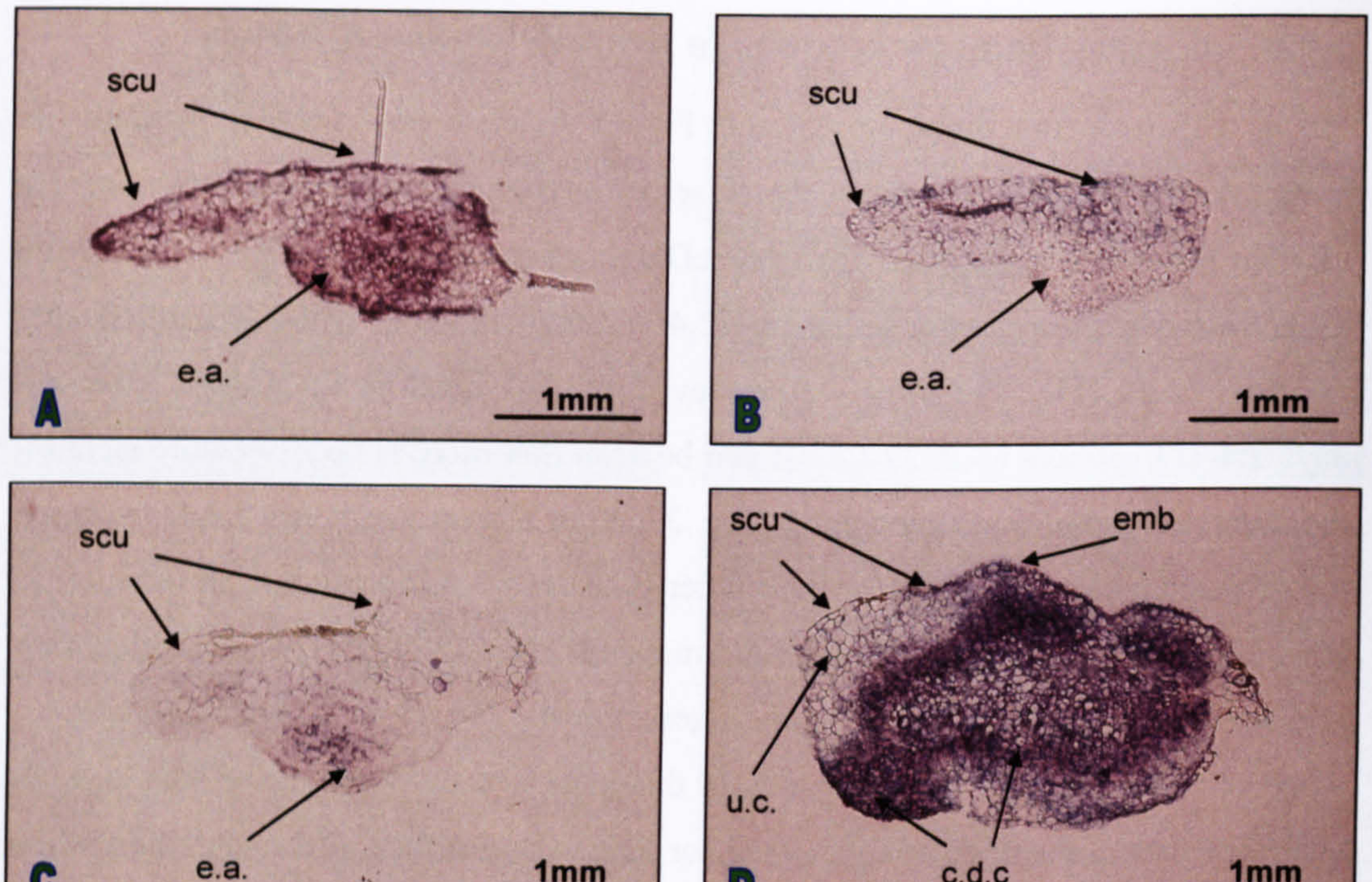


Figure 38 (a-e). Embryogenic BW calli at various stages (Tp1-4) showing the bound experimental probe, Incyte EST (ID = 701961028). Immature embryos were fixed after the following treatments A) TP0 = untreated, B) TP1 = 7 Hrs 2,4-D Treatment, C) TP2 = 4 Days 2,4-D treatment, D) TP3 = 7 Days 2,4-D treatment, E) TP4 = 7 Days 2,4-D treatment + 1 Day MS. Tissue sections were cut 8um thick. The Incyte EST sequence was identified in the cDNA microarray experiments as being differentially expressed after seven days 2,4-D treatment (TP3) which is confirmed by the ISH data *e.a.* = embryonic axis, *scu* = scutellum, *u.c.* = unorganised cells, *c.d.c.* = cytoplasmically dense cells, *emb* = embryoid body. Bar = 1mm

5.3.9. Incyte EST (Incyte Clone ID = 701961028)

Another experimental probe that was identified in the cDNA microarray experiments as being differentially expressed after seven days 2,4-D treatment (TP3) was the Incyte EST sequence (clone ID = 701961028) with a DEV of 5. This result is confirmed by the ISH data that demonstrates a significant increase in signal intensity of this gene during timepoint 3 (Fig. 38 D). It can also be seen that this gene sequence is strongly expressed in the untreated immature embryo sample at timepoint 0 (Fig. 38 A) and appears to be localised to the embryonic axis and the edge of the scutellum. There is a noticeable reduction in signal intensity in timepoint 1 after 7 hours (Fig. 38 B) and in timepoint 2 after 4 days 2,4-D treatment (Fig. 38 C) with the *in situ* signal at timepoint 2 appearing to be restricted to cells associated with the embryonic axis. It can be seen that in timepoint 3 after seven days 2,4-D treatment (Fig. 38 D) there is a significant increase in signal intensity which appears to be localised to the cytoplasmically dense cellular regions of the callus. On removal of 2,4-D from the medium (TP4) there again appears to be a significant reduction in the *in situ* signal. Taken together these *in situ* results suggest that this Incyte EST is expressed in untreated immature embryos and is down regulated when cultured on high auxin media until after seven days where it then appears to be specifically and transiently induced. Following a further 24 hours culture on media without any auxin, there is a subsequent reduction in signal intensity suggesting that this gene may be implicated in the establishment of embryogenic competence.

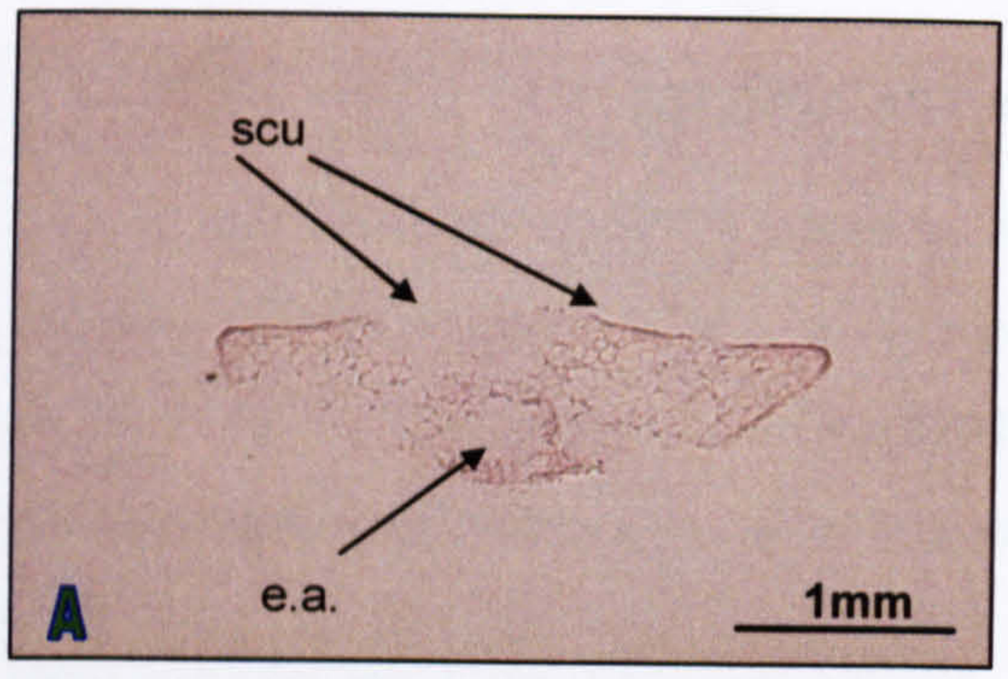


Figure 39 (a-e). Embryogenic BW calli at various stages (Tp1-4) showing the bound experimental probe, protein phosphatase 2C (PP2C) - like. Immature embryos were fixed after the following treatments A) TP0 = untreated, B) TP1 = 7 Hrs 2,4-D Treatment, C) TP2 = 4 Days 2,4-D treatment, D) TP3 = 7 Days 2,4-D treatment, E) TP4 = 7 Days 2,4-D treatment + 1 Day MS. Tissue sections were cut 8um thick. This gene forms another interesting candidate for further analysis as it appears to be induced by the removal of 2,4-D from the tissue culture medium and may play an important role in the differentiation of SEs and embryogenic calli into adult structures *.e.a.* = embryonic axis, *scu* = scutellum, *u.c.* = unorganised cells, *c.d.c.* = cytoplasmically dense cells, *emb* = embryoid body. Bar = 1mm

5.3.10. Protein Phosphatase 2C (PP2C)

This Protein phosphatase 2C (PP2C) gene has a very similar pattern of expression to the OSE731 gene. It was originally identified in the cDNA microarray experiments as being expressed after four days (TP2) and seven days treatment with 2,4-D (TP3) at values of 4 and 2.2 respectively. However, like the OSE731 gene the *in situ* hybridisation results failed to detect any significant signal at these timepoints which may be explained for the same reasons as discussed earlier. These include, the particular section in question being derived from incorrectly fixed tissue, or the more likely explanation that the gene is expressed in a specific, localised area of the callus that is not represented in the section resulting in the gene not appearing to be expressed. However, a significant signal is detected in calli that have been removed from the 2,4-D medium in timepoint 4 (Fig. 39 E) suggesting that this gene is expressed in calli on removal of 2,4-D from the tissue culture medium. Therefore, like the OSE 731 gene, this gene also forms another interesting candidate for further analysis as it appears to be specifically induced by the removal of 2,4-D from the tissue culture medium and may play an important role in the differentiation of SEs and embryogenic calli into adult structures.

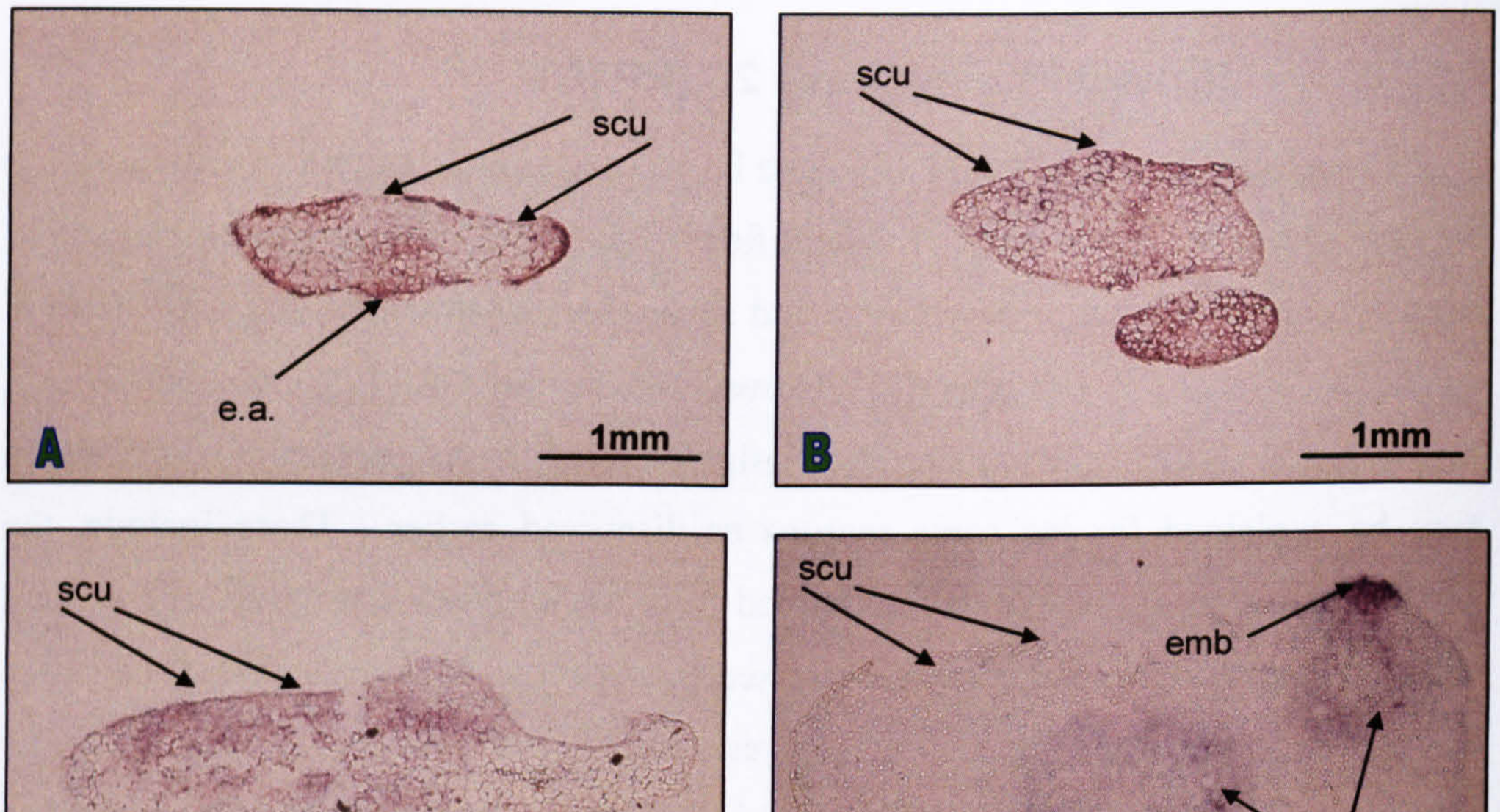


Figure 40 (a-e). Embryogenic BW calli at various stages (Tp1-4) showing the bound experimental probe, Ser-Thr protein kinase-like. Immature embryos were fixed after the following treatments A) TP0 = untreated, B) TP1 = 7 Hrs 2,4-D Treatment, C) TP2 = 4 Days 2,4-D treatment, D) TP3 = 7 Days 2,4-D treatment, E) TP4 = 7 Days 2,4-D treatment + 1 Day MS. Tissue sections were cut 8um thick. *e.a.* = embryonic axis, *scu* = scutellum, *u.c.* = unorganised cells, *c.d.c.* = cytoplasmically dense cells, *emb* = embryoid body. Bar = 1mm

5.3.11. SER-THR Protein Kinase-like Protein

Another experimental probe that was originally identified as being differentially expressed in the cDNA microarray experiments after four days 2,4-D treatment (TP2) at a value of 2 is the serine-threonine protein kinase-like protein. There appears to be an *in situ* hybridisation signal detected during timepoint 0 in untreated embryos (Fig. 26 A), after 7 hours 2,4-D treatment in timepoint 1 (Fig. 40 B), and after 7 days 2,4-D treatment in timepoint 3 (Fig. 40 D). Following removal for 24hrs on MS media in timepoint 4 there appears to be a complete reduction in the ISH signal (Fig. 40 E). The fact that this gene appears to be expressed in untreated IEs may account for the undetected differential expression at timepoints 1 and 3 in the microarray experiments. Furthermore, it suggests that this gene may also play a role in zygotic embryogenesis as well as in somatic embryogenesis. The patterns of expression of this genes signal at timepoint 0 and timepoint 1 (Fig. 40 A+B) appears to be fairly ubiquitous whereas the spatial expression at timepoint 2 appears to be localised to the scutellar tissue (Fig. 40 C). In addition, the spatial expression at timepoint 3 appears to be specifically localised to cytoplasmically dense cellular regions of the embryogenic calli and in particular to an emerging embryoid body region (Fig. 40 D). Therefore the expression of this gene implicates a functional role in emerging somatic embryo-like structures of embryogenic calli under the control of high auxin levels as well as playing a role in the establishment of zygotic embryogenesis.

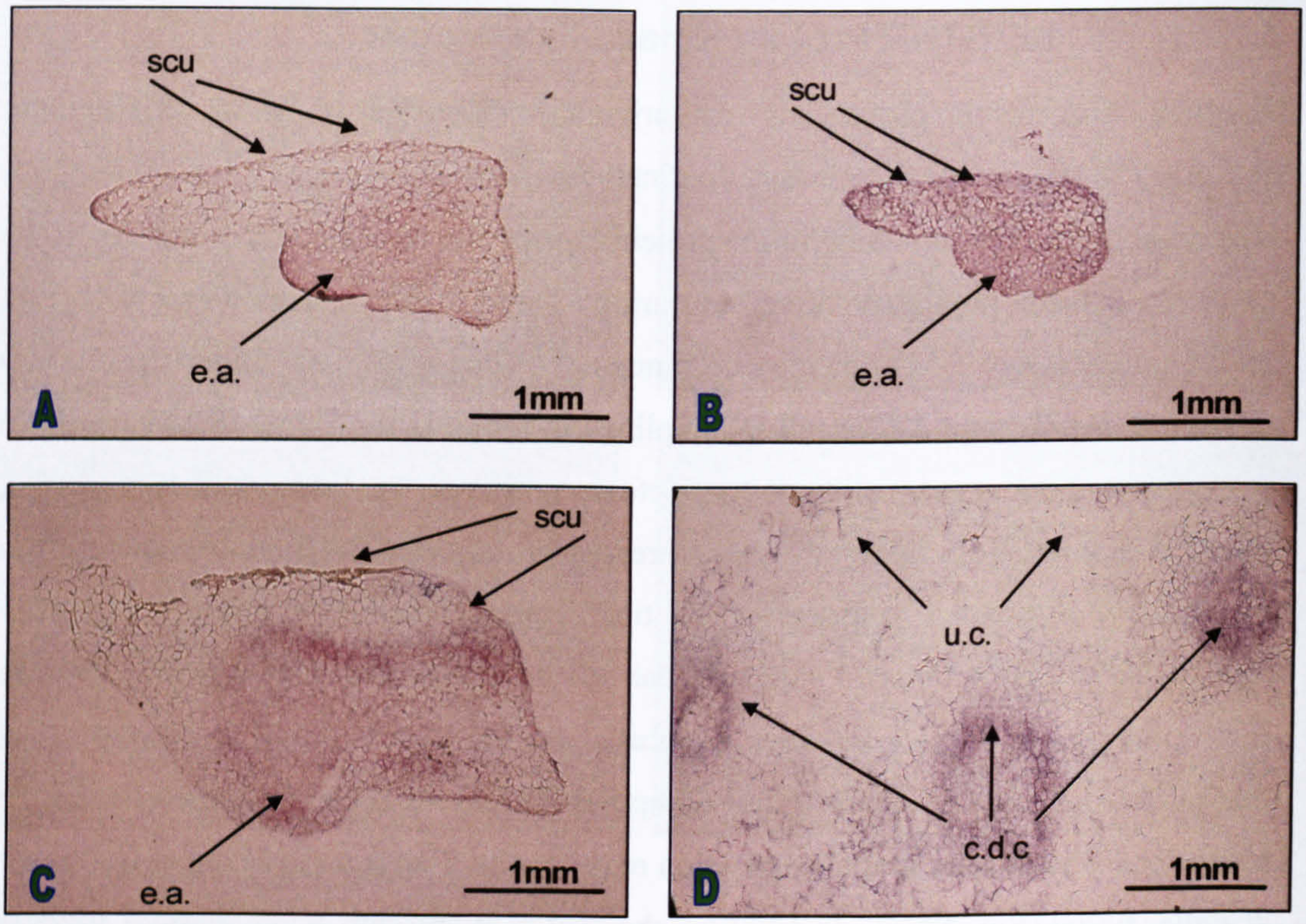


Figure 41 (a-e). Embryogenic BW calli at various stages (Tp1-4) showing the bound experimental probe, SCARECROW (SCR) transcriptional regulator-like. Immature embryos were fixed after the following treatments A) TP0 = untreated, B) TP1 = 7 Hrs 2,4-D Treatment, C) TP2 = 4 Days 2,4-D treatment, D) TP3 = 7 Days 2,4-D treatment, E) TP4 = 7 Days 2,4-D treatment + 1 Day MS. Tissue sections were cut 8um thick. *e.a.* = embryonic axis, *scu* = scutellum, *u.c.* = unorganised cells, *c.d.c.* = cytoplasmically dense cells. Bar = 1mm

5.3.12. SCARECROW (SCR) Transcriptional Regulator-like

The next experimental probe that is presented in the results is a SCARECROW (SCR) transcriptional regulator-like gene that was originally identified in the cDNA microarray experiments as being expressed at a value of 4 after four days 2,4-D treatment (TP2). The ISH experiments reveals the expression of this gene during timepoint 0 in untreated embryos (Fig. 41 A), and timepoint 1 after 7 hours 2,4-D treatment (Fig. 41 B), which appears to be in a uniform and ubiquitous pattern. This pattern of expression changes after 4 days 2,4-D during timepoint 2, (Fig. 41 C) and appears to maintain a signal in cells of the embryonic axis and a layer of cells below the surface of the scutellum. There are also signals detected during timepoints 3 and 4 (Fig. 41 D+E) that appear to be localised to cytoplasmically dense regions of the developing embryogenic calli. Therefore the expression pattern of this gene suggests it is zygotically expressed in the immature embryo and has an associated functional role in the development of SE which is specifically localised to cytoplasmically dense cellular regions of embryogenic calli.

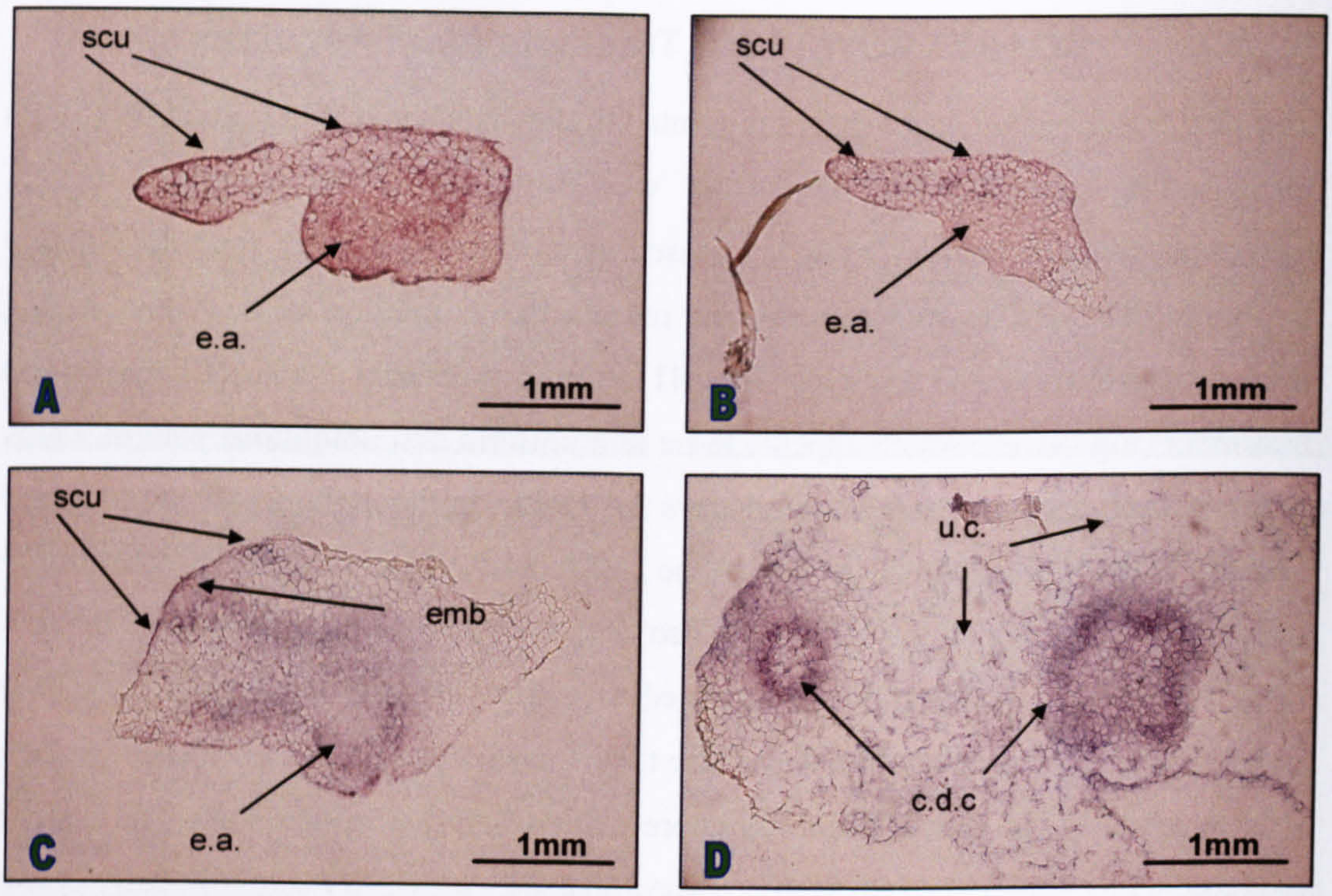
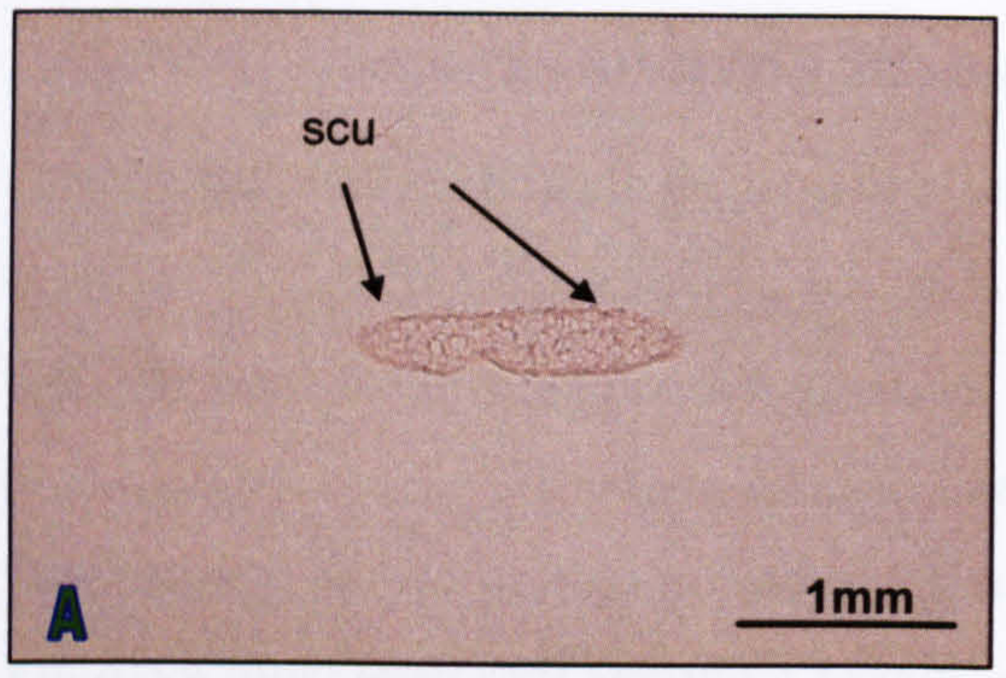
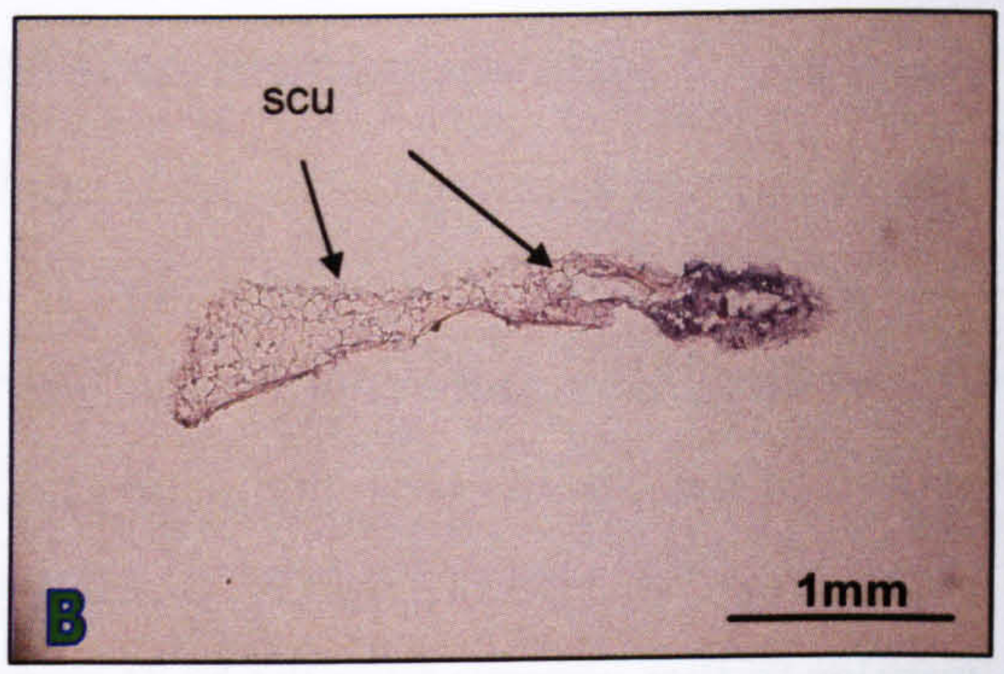


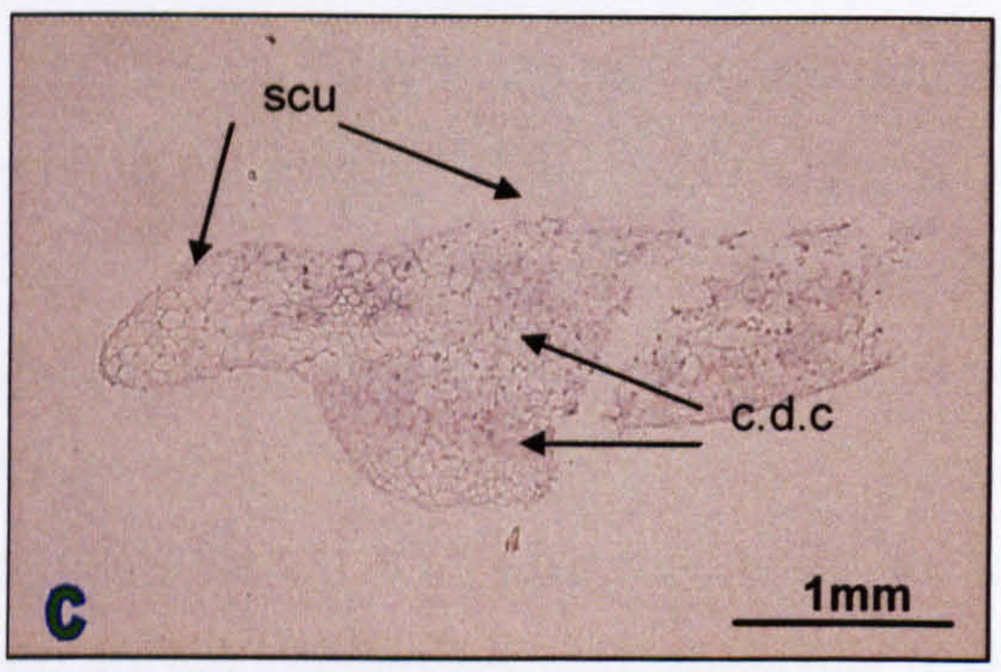
Figure 42 (a-e). Embryogenic BW calli at various stages (TP0-4) showing the bound experimental probe, somatic embryogenesis (SE) related protein. Immature embryos were fixed after the following treatments A) TP0 = untreated, B) TP1 = 7 Hrs 2,4-D Treatment, C) TP2 = 4 Days 2,4-D treatment, D) TP3 = 7 Days 2,4-D treatment, E) TP4 = 7 Days 2,4-D treatment + 1 Day MS. Tissue sections were cut 8um thick. *e.a.* = embryonic axis, *scu* = scutellum, *u.c.* = unorganised callus, *c.d.c.* = cytoplasmically dense cells. Bar = 1mm

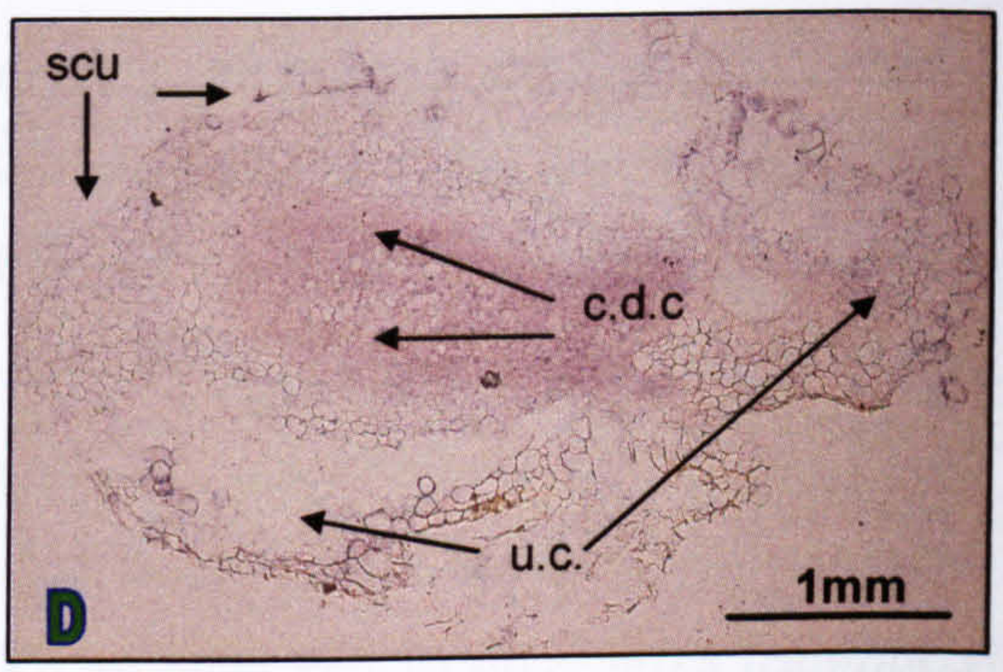
5.3.13. Somatic Embryogenesis (SE) Related Protein

The SE related protein was also originally identified in the cDNA microarray experiments as being expressed after four days 2,4-D treatment (TP2) at a value of 2.2. The ISH results confirm that there is a signal expressed in the embryo axis region of the calli as well as in a central band of cells above the embryo axis in calli during this timepoint (Fig. 42 C). The overall expression pattern of this gene is very similar to that observed for the SCR transcriptional regulator-like gene (Fig. 42 A-E). That is, that the expression of this gene during timepoint 0 in untreated embryos (Fig. 28 A), and in timepoint 1 after 7 hours 2,4-D treatment (Fig. 42 B), appears to be in a uniform and ubiquitous pattern. In addition, the pattern of expression also changes after 4 days 2,4-D during timepoint 2, to one of a more localised manner (Fig. 42 C). ISH signals are also detected at timepoints 3 and 4 which are localised to cytoplasmically dense embryogenic regions of the developing calli, indicating the continued expression of this gene in all tested stages of SE. As mentioned for the SCR transcriptional regulator-like gene, the expression pattern suggests that it is expressed in the zygotic embryo and may have a functional role in the development of cytoplasmically dense regions of embryogenic calli and therefore provides another interesting candidate for further analysis.









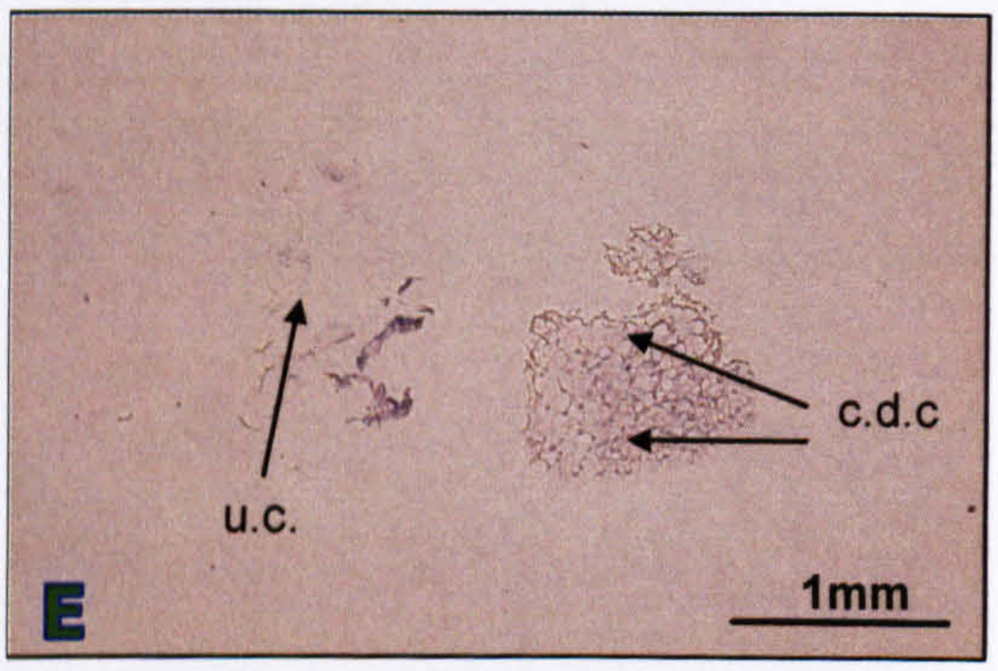


Figure 43 (a-e). Embryogenic BW calli at various stages (TP0-4) showing the bound experimental probe, Hypothetical Protein; 53156-50996. Immature embryos were fixed after the following treatments A) TP0 = untreated, B) TP1 = 7 Hrs 2,4-D Treatment, C) TP2 = 4 Days 2,4-D treatment, D) TP3 = 7 Days 2,4-D treatment, E) TP4 = 7 Days 2,4-D treatment + 1 Day MS. Tissue sections were cut 8um thick. *scu* = scutellum, *u.c.* = unorganised callus, *c.d.c.* = cytoplasmically dense cells. Bar = 1mm

5.3.14. Hypothetical Protein; 53156-50996

The final experimental probe presented here is the hypothetical protein that was originally identified in the cDNA microarray hybridisations as being differentially expressed after seven hours 2,4-D treatment (TP3) at a value of 5.9. The results from the ISH experiments confirm the increased expression of this gene after 7 Hrs 2,4-D treatment at timepoint one (Fig. 43 B) where a specific region of the scutellum shows a particularly strong signal, which is in contrast to the rest of the sample and the untreated sample in timepoint 0 (Fig. 43 A) which exhibit a low, uniform background signal. Low, background signals are also seen after 4 days 2,4-D treatment in timepoint 2 (Fig. 43 C) but there appears to be a slight increase in signal strength in timepoint 3 (Fig. 43 D), localised to cytoplasmically dense cells. As this appears to be at a low level it is possibly not sufficient enough to be detected as differentially expressed during the cDNA microarray hybridisations. Finally, a low background signal is also seen after 7 days 2,4-D treatment and 1 day MS in timepoint 4 (Fig. 43 E). These results are consistent with the microarray data that demonstrates a significant increase in the differential expression of this gene after 7 hrs 2,4-D treatment. This suggests that this gene is only transiently expressed for a short period following its initial induction thereby making it a potential candidate that is associated with the initiation of SE.

5.4. Discussion

The ISH experiments have demonstrated that this approach has served its required functions. Firstly, although there are some discrepancies, the results have generally confirmed positive differential expression signals for the genes that were identified as being differentially expressed during SE by the microarray experiments. Secondly, it has provided further information concerning the spatial and temporal expression of these identified gene sequences and additional insight into the possible function of these genes in SE. For example, the results of the *in situ* expression analysis for DsPTP1 provide conclusive confirmation of the identified microarray differential expression values. Furthermore, additional information about the localisation of this gene is illustrated by the change from a uniform and ubiquitous pattern in timepoint 1 to that of a specific expression in cytoplasmically dense cellular regions of the

developing calli in timepoint 2 (Fig. 34 C) and timepoint 4 (Fig. 34 E). The fact that this DsPTP1 gene has not been identified as being differentially expressed in any of the Savannah non-embryogenic hybridisations further implicates this gene as playing an important regulatory role in the initiation and development of SE. As previously discussed (4.4.1) this gene was shown to be regulated by calmodulin and dephosphorylates and inactivates MPK4, a mitogen activated protein kinase (MAPK) (Yoo *et al.*, 2004). MAPKs are a family of serine/threonine protein kinases involved in many cellular responses including cell proliferation and differentiation (Schaeffer and Weber, 1999). Therefore, it is possible that this gene is involved in the regulation of a MAPK signalling response that is specifically expressed during the initiation and development of SE.

An important pre-requisite for interpreting such results is the analysis and comparison of the control set of probes. For example, the Histone 4 and Cyclin D positive controls demonstrate the integrity of the fixed tissue and provide some background information about the nature of cell division and proliferation within the developing embryogenic calli. The expression of these genes is indicative of active and dividing cells and therefore demonstrates that although untreated embryos are active and dividing, the treatment with 2,4-D stimulates rapid cellular proliferation, particularly in the scutellar regions of mature embryogenic calli (TP3). In addition to the positive controls, the alpha-thionin negative control demonstrates the low level of random or non-specific binding of the experimental probe allowing for the validation of genuine probe hybridisation among candidate genes.

The results from the *in situ* experiments confirm that the ABF gene is differentially expressed during timepoint 2, however, they also clearly reveal the additional expression of this gene at timepoints 1 and 4 (Fig. 35 B+E). The pattern of expression of this gene is indicative of a cell cycle like regulated gene and is similar to those observed in the histone 4 and cyclin D controls i.e. strongly expressed in single cells in a random manner. However, what makes the expression pattern of this ABF gene particularly interesting is that it is not observed in untreated cells and appears to be specifically induced by the application of 2,4-D treatment. Furthermore, the annotation of this gene suggests that it is regulated by ABA. ABFs are induced by ABA and various stress treatments and have been shown to transactivate an ABRE-

containing reporter gene in yeast (Choi *et al.*, 2000). Therefore it can be seen that the ABA pathway is involved, probably as a response to high auxin stress, in mediating early signalling events in SE that takes place even before any changes in morphology are observed. The fact that this gene is not identified as being differentially expressed during any of the SAV hybridisations, suggests that this gene may play a significant role in the developmental switch of cells towards the initiation of the embryogenic stage and also continues to be involved in the subsequent development of embryogenic calli and the differentiation of SEs into adult tissues and structures.

The *Oryza sativa* embryo specific (OSE731) mRNA that was identified from the microarray hybridisations as being expressed during timepoint 2 is not confirmed by the ISH results which fail to detect any significant signal at this timepoint. This may be because the gene is actually expressed in a specific, localised area of the callus that is not represented in the particular section in question resulting in the failure to detect it by ISH. In addition there is some signal detected after seven hours 2,4-D treatment in timepoint 1 which is localised to the embryonic axis which is consistent with its annotation as being an embryo specific gene. However, a significant signal is detected in calli that have been removed from the 2,4-D medium in timepoint 4 (Fig. 36 E). This interesting expression pattern suggests that the OSE 731 gene is significantly induced following removal of 2,4-D and is therefore likely to play a role in the early events that lead to the regeneration of adult structures from somatic embryos and embryogenic calli. As previously discussed (4.4.2) a maize homologue of the rice embryo specific gene was shown to be involved in stress response during seed maturation and germination (Kollipara *et al.*, 2002) which would be consistent with the detection of this gene after removal of 2,4-D, i.e. a stage which is associated with the differentiation and regeneration of SEs into adult structures. A similar pattern of expression was observed for the PP2C gene. PP2C is a Mn+ or Mg+ dependent serine/threonine phosphatase and shows similarity to the ABA insensitive genes (ABI1 and ABI2). Furthermore, in *Arabidopsis* the kinase associated protein phosphatase (KAPP) is an enzyme that dephosphorylates the Ser/Thr receptor like kinase RLK5 which contains a C-terminal PP2C domain (Stone *et al.*, 1994). It is therefore possible that this gene is involved in the regulation of a signalling pathway that is important for the development of SE that may also be regulated by ABA.

The ISH results confirm the expression of the LRRK gene during timepoint 3 as well as revealing a similar signal at the earlier timepoint 2 indicating that this gene is transiently expressed during the establishment of embryogenic calli and becomes quickly down regulated on the removal of 2,4-D from the tissue culture medium. The transient expression pattern of the LRRK gene is of particular interest as (Schmidt *et al.*, 1997) identified a gene that was transiently and specifically expressed in single cells of carrot cultures that possessed the competence to form somatic embryos. This gene was identified as a leucine-rich repeat containing receptor-like kinase protein designated SERK. The signal in the wheat calli appears to be localised to cytoplasmically dense cells in a central region of the developing calli which may allow somatic cells the competence to form embryo structures that typically emerge from the edge of the scutellum. Furthermore, SERK expression ceased after the globular stage during SE which is analogous to the cessation of ISH signal between timepoint three and timepoint four of the developing wheat calli. It is quite possible therefore that this wheat sequence is a homologue of the SERK gene that provides cells with the competence to form somatic embryos. It would be interesting to further study the cells that express this gene by tracking their fate to determine whether or not they end up in somatic embryo structures.

The ISH data confirms that there is a significant increase in signal intensity of this Incyte EST gene during timepoint three (Fig. 38 D). The up-regulation of the Incyte EST gene at timepoint three suggests that this gene has a short developmental window during the development of SE. In addition, the apparent expression of this gene in immature embryos at timepoint zero suggests that this gene may also be associated with zygotic embryogenesis. The lack of annotation for this gene sequence makes it difficult to speculate how this gene may function. However, there is no detected differential expression of this gene during the non-embryogenic hybridisations and is therefore implicated as being associated with the establishment of SE. Further characterisation of this novel gene will be of particular interest for increasing our understanding of the molecular mechanisms involved in SE.

The role of the SCR gene was previously discussed (4.4.2) as being involved in the regulation of stem cell fate and asymmetric division (Di Laurenzio *et al.*, 1996) and is therefore likely to be involved in regulating the meristematic activity and asymmetric

division of embryogenic cells in SE. The expression pattern of this gene as observed by the ISH results suggests that not only is it zygotically expressed in the immature embryo (Fig. 41 A) but it can also be seen to play a functional role in the development of SE which is specifically localised to cytoplasmically dense cellular regions of embryogenic calli (Fig. 41 C-E). Again this gene was not shown to be differentially expressed in the non embryogenic hybridisations and therefore implicates it as playing an important role in the regulation of meristematic activity and asymmetric division of embryogenic cells in SE.

The ISH results for the SE related gene confirm the presence of a signal in calli during timepoint 2 (Fig. 42 B) that was identified in the cDNA microarray hybridisations. The overall expression pattern of the SE related gene (Fig. 42 A-E) is very similar to that observed for the SCR transcriptional regulator-like gene (Fig. 43 A-E), which indicates that it is also expressed in the zygotic embryo and may have a functional role in the development of cytoplasmically dense regions of embryogenic calli. The SE related gene does not appear to be expressed in any of the non embryogenic microarray hybridisations and therefore confirms the observation that this gene was differentially expressed in embryogenic but not non-embryogenic cultures (Alexandrova and Conger, 2002). As previously discussed, this gene shows strong similarity to the WRKY21 a DNA binding protein from *Arabidopsis* which encodes a nuclear localisation sequence and DNA binding domain (Alexandrova and Conger, 2002) implicating a functional role for this gene in the regulation of transcription during the initiation and development of SE.

The ISH results for the hypothetical protein confirm the increased expression of this gene at timepoint one (Fig. 43 B) where a specific region of the scutellum shows a particularly strong signal. As this gene is annotated as being a hypothetical protein it is difficult to ascertain any possible functional role during the initiation of SE. However, the fact that it is identified as being up-regulated at a value of 5.9 after only a few hours exposure to auxin and is confirmed by an *in situ* hybridisation signal suggests that this gene may play an important role in triggering the developmental switch of somatic cells towards an embryogenic state. Further characterisation of this gene will be of particular interest for increasing our understanding of the mechanisms responsible for the early de-differentiation and initiation responses involved in SE.

In Summary, the ISH experiments have been largely successful in both confirming the expression of these gene sequences and in revealing the spatial and temporal expression of the gene sequences that were identified in the cDNA microarray experiments. This has provided valuable information about the localised expression of these genes in their tissues, thereby improving their functional characterisation. However, to fully determine their functional role in SE, gene overexpression and knock out studies are required. Unfortunately, due to time constraints it was not possible to fully clone each gene in order to carry out overexpression studies. However, using the available clone DNA sequences, short regions were able to be used to carry out RNAi gene silencing experiments to obtain some functional data about the role of these shortlisted genes in SE.

Chapter Six

Gene Testing Experiments

6 Chapter Six: Gene Testing Experiments

6.1. Introduction

RNA silencing or RNA interference (RNAi) was first observed in plants as a mysterious immune response to viral pathogens as long ago as 1928 (Wingard, 1928) and was, retrospectively, also originally discovered in plants as post transcriptional gene silencing (PTGS) (Jorgensen *et al.,* 1996). However, it was not until a landmark paper that identified the trigger for gene silencing as double stranded RNA (Fire *et al.,* 1998) that the phenomenon of RNAi was described and the first real breakthrough in the mechanistic understanding of RNA silencing came about. Since then RNAi has been shown to occur in a wide variety of eukaryotic organisms as an evolutionarily conserved mechanism with several distinct pathways (Tijsterman *et al.,* 2002), the components of which have been identified at an impressive rate. It soon became clear that RNAi was more than just a response to exogenous genetic material as small RNAs, known as micro RNAs (miRNAs) were implicated in the regulation of gene expression and development (Reinhart *et al.,* 2000). RNAi has also unexpectedly been shown to function in heterochromatin silencing through the targeting of repeat sequence transcripts and transposable elements that are present in abundance throughout the genome (Lippman *et al.,* 2003). In addition to increasing our understanding of fundamental developmental and regulatory gene mechanisms, RNAi is being utilised as an experimental tool to revolutionise the exploration of gene function in a number of different organisms.

One of the most effective ways to determine the biological function of a protein is to examine the phenotype of an organism that contains some kind of mutation or disruption in its encoding gene (Hammond *et al.,* 2001). Typically such mutations have been generated in a directed manner via homologous recombination or by screening randomly mutagenised populations for lesions in a gene of interest. Alternative methods for silencing specific genes such as antisense methods using DNA or RNA have provided powerful approaches for probing gene function. However, this methodology has suffered from questionable specificity and incomplete efficacy (Hammond *et al.,* 2001). Recently (Fire *et al.,* 1998) from their investigations into these artefacts surprisingly found that double stranded RNA was

substantially more effective at producing interference than either strand individually. This discovery, termed RNA interference (RNAi), has revolutionised reverse genetic approaches and is currently being harnessed in large scale genomic efforts in plants. In particular (Smith *et al.,* 2000) showed that gene constructs encoding intron spliced RNA with a hairpin structure can induce RNA silencing in plants when directed against endogenous genes. The hairpin RNA (hpRNA) construct contains sense and anti-sense sequences flanking an intron spacer sequence which contributes to the stability of the reverted repeat sequence but is not required for specificity of RNA silencing. This intron is spliced out during pre mRNA processing to produce a loopless hpRNA which was shown to be 100% effective in silencing endogenous plant genes (Smith *et al.,* 2000).

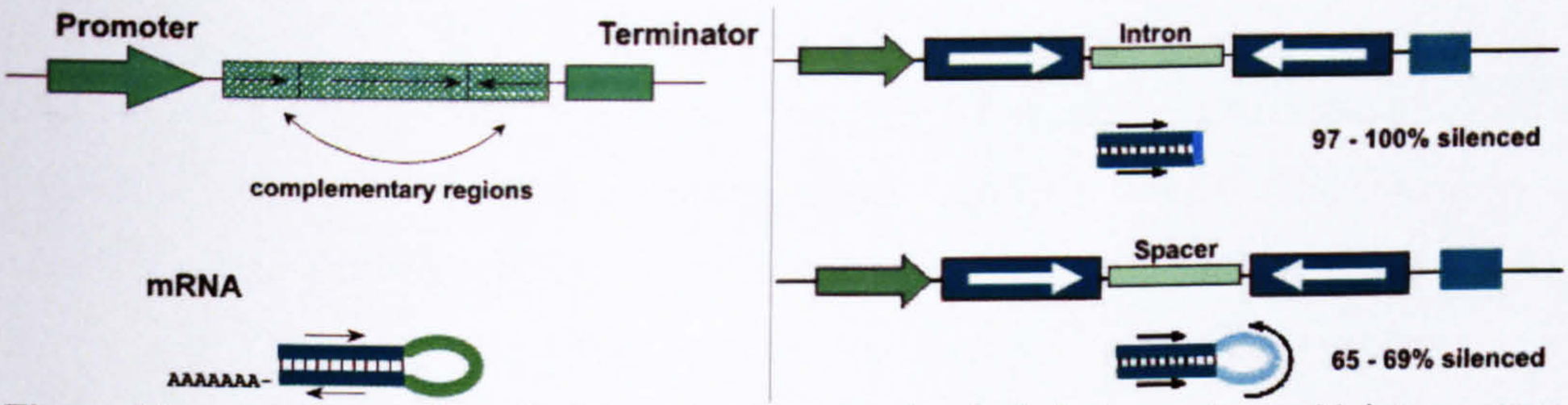


Figure 44. Diagram from (Smith *et al.*, 2000) showing the hairpin construct with intron spacer.

Based on this approach Gateway™ compatible transformation vectors, constructed by Syngenta, were used in this project to generate gene specific RNAi vectors. The Gateway™ technology utilises site-specific recombination to enable the cloning of DNA fragments into a destination vector without the need to use restriction enzymes and ligase and is based on the lambda phage site-specific recombination system (*att*B x *att*P > *att*L x *att*R), the genetic mechanism by which phage lambda integrates itself into the host genome of *E. coli*. The cloning is composed of two reactions; firstly, the cloning of a PCR product into the entry vector (BP reaction) and secondly, the site-specific recombination between the entry and destination vectors (LR reaction). The Gateway™ system also utilises a *ccd*B selection to ensure the highly efficient recovery of recombinant clones (>99%). Both the entry and destination vectors contain the *ccd*B gene as a selectable marker, which encodes a protein that disrupts the cell cycle and consequently inhibits the growth of standard *E. coli* strains. During recombination, the *ccd*B gene is replaced by the DNA fragment of interest, which enables the recombinant colonies to grow.

The ISH images presented in chapter five (5.3.5 to 5.3.14.) comprise a subset of the candidate genes presented in the shortlist in chapter four (Table 12). These genes were selected on the basis of their confirmed expression and localised signalling patterns for further genetic and functional characterisation. Due to the long turnaround time for generating wheat mutants and due to the time constraints faced on the project, it was decided that the best way to obtain functional data about these genes was to characterise them using RNAi technology. However, an important prerequisite for this approach was to establish the efficacy of this technique in wheat.

6.2. Materials and Methods

6.2.1. RNAi Transient Assays

The strategy to test whether RNAi works in wheat is to bombard the wheat IEs with the overexpression construct of a selectable/detectable marker gene either alone or together with a specific RNAi construct. Comparisons of marker gene expression in the IEs following the different bombardments will provide evidence whether the co-bombardment (of the overexpression and RNAi constructs of the marker gene) could lead to gene silencing.

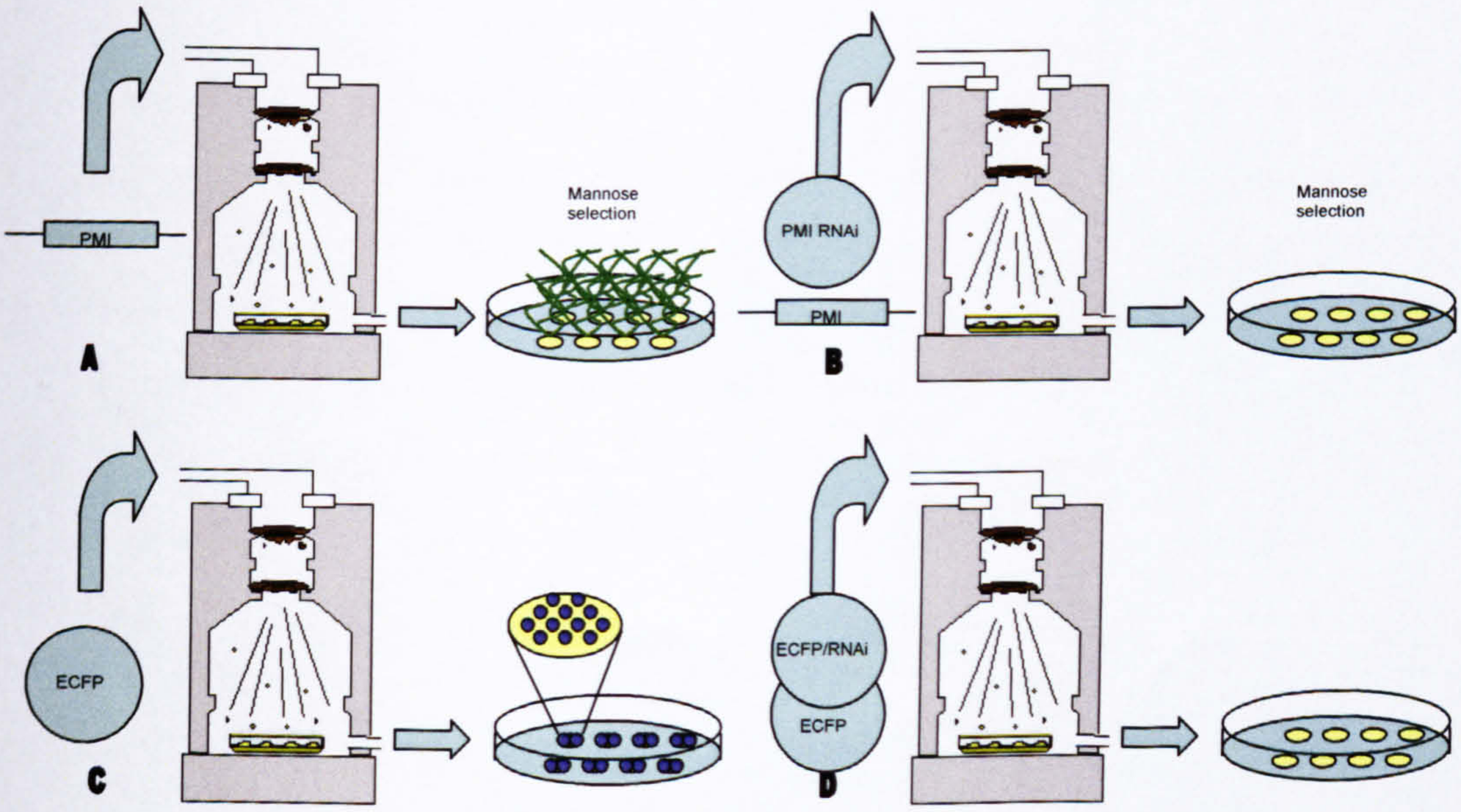


Figure 45. Diagram illustrating the strategy employed for testing RNAi gene silencing in wheat. A) IEs bombarded with the PMI construct would be expected to regenerate shoots on mannose selection. B) Successful silencing of the PMI construct by the co-transformed PMI RNAi construct would result in the inability of shoots to regenerate on mannose selection. C) IEs bombarded with the ECFP construct should result in the transient expression of the fluorescent protein. D) IEs co-bombarded with the ECFP and ECFP RNAi construct would result in silencing of the fluorescent protein.

6.2.1.1. Construction of the ECFP Visual Marker Vector

A linearised plasmid preparation of construct pCIB11251 containing an enhanced cyan fluorescent variant of the *Aequorea victoria*-derived green fluorescent protein gene *ecfp* (CLONTECH), driven by the strong constitutive *cestrum* yellow leaf curling virus (cmYLCV) promoter (Stavolone *et al.*, 2003) was used for transformation in the fluorescent protein transient assay. DNA was extracted using the QIAprep® spin miniprep kit (QIAGEN, 27104) according to the instruction manual and the plasmid DNA linearised by O/N digestion with HindIII at 37^{0}C. The linearised DNA was purified using the QIAquick® purification kit (QIAGEN, 28104) according to manufacturer's instructions and resuspended in TE buffer to a final concentration of 1 μg/μl ready for transformation.

6.2.1.2. Construction of the ECFP RNAi Vector

For the generation of ECFP RNAi entry clones a 250 bp PCR product was synthesised using the *attb* linked (underlined) gene specific primers (GSP) ECFP F; 5'GGGGAC AAGTTTGTACAAAAAAGCAGGCTCAAAGTTCATCGGCGACGACATGA 3' and ECFP R; 5'GGGGACCACTTTGTACAAGAAAGCTGGGTACCCTTCAGG CTGATCTCCCA 3' with the pCIB11251 construct as a template. BP reactions between the *AttB* linked PCR products and the pDONR™221 entry vector were carried out according to the Invitrogen™ Gateway™ instruction manual. Several of the resulting transformed colonies were grown and the DNA extracted using the QIAprep® spin miniprep kit. DNA from the Gateway™ compatible RNAi destination vector pCIB11079 (provided by Syngenta) containing an inverted *Att*R repeat separated by a rice intron and driven by the maize ubiquitin intron 1 promoter (Christensen *et al.*, 1992) was used for the generation of the ECFP RNAi transformation vectors. The LR recombination reaction between the ECFP pDONR™221 entry clones and the Gateway™ compatible RNAi destination vector pCIB11079 was carried out as described in the instruction manual. Primers were designed to sequence the inverted repeats and several colonies were sequenced to verify that the PCR products had integrated with the pCIB11079 RNAi destination vector. DNA was extracted from the sequence verified colonies using the QIAprep® spin miniprep kit according to manufacturer's instructions and resuspended in TE buffer to a final concentration of 1 μg/μl ready for transformation.

6.2.1.3. Donor Plant Material

Due to changes in the commercial operations at Syngenta, Bobwhite (BW) was no longer used as a cultivar to generate transgenics. Therefore, the spring variety Bolero was used as the wheat donor material for the RNAi assay and gene testing experiments as it has an identical response to tissue culture conditions as BW. The Bolero donor plants were grown as described for BW (section 2.2.1.) and IEs were aseptically isolated onto regeneration medium as described (section 2.2.2.). The IEs were cultured in the dark at 25°C/40% humidity for approximately one week then evaluated for embryogenic response as indicated by swelling of the scutellar tissue and the formation of glossy, smooth, pre-globular, structures. Only high quality, responding embryos were used for transformation.

6.2.1.4. Transformation

Approximately 100 wheat IEs with embryogenic responses, were arranged in a 2.5cm diameter circle with the scutellum facing upward. Wheat tissues were exposed to osmoticum-containing medium (0.2 M sorbitol and 0.2 M mannitol) for four hours prior to gene delivery. DNA for both the overexpression and RNAi constructs was precipitated at equal molar concentrations onto gold microcarriers according to the DuPont Biolistics™ manual. The DNA was delivered to the target tissue cells using the PDS-1000He Biolistics™ device with the distances between the rupture disk and the macrocarrier, the macrocarrier and the stopping screen, and the stopping screen and the target set at 8 mm, 10 mm and 7 cm respectively. To reduce tissue damage from the helium blast, a stainless steel mesh (McMaster-Carr), with 200 openings per linear inch horizontally and vertically, was placed between the stopping screen and the target tissue. After gene delivery, the wheat IEs were further subjected to high osmotic treatment in the dark with 25°C and 40% humidity for 16 hours and then moved to the callus induction (3MS3S) medium (2.2.3.2.) and incubated for one week while they were evaluated by fluorescence microscopy.

6.2.1.5. ECFP Fluorescence Microscopy

Images were taken of the calli at 24 hourly intervals post bombardment to monitor and record the transient expression levels of the ECFP protein. This allowed for the direct comparison of calli bombarded with the ECFP overexpression construct alone

alongside calli that had been co-bombarded with the ECFP overexpression construct and the ECFP RNAi construct. Fluorescent images were taken with an axiocam HR digital camera (Carl Zeiss) attached to an axioskop 2MAT microscope (Carl Zeiss) while using the filter set number 05 (excitation BP395-440 nm, emission LP470 nm). The acquired images were processed with the axiovision software program (Carl Zeiss) and loaded into Paintshop Pro for direct comparison of expression.

6.2.1.6. Construction of the PMI Selectable Marker Vector

A fragment preparation of construct pCIB9818 (provided by Syngenta) containing the *E. coli*-derived *manA* gene encoding phosphomannose isomerase (*pmi*) driven by the maize ubiquitin intron 1 promoter (Christensen *et al.*, 1992) was used for transformation in the selectable gene marker experiments. DNA was extracted using the QIAGEN® Plasmid Maxi Kit (QIAGEN, 12162) according to the instruction manual. Typically, 250 µg of plasmid DNA was digested in a final volume of 500µl and then run on 0.8% agarose gel in TBE buffer. The band containing the fragment DNA was cut from the gel and electroeluted for 3 hrs at 200 volts using Elutrap® (Schleicher and Schuell). The DNA was then precipitated and resuspended in TE buffer to a final concentration of 1 µg/µl.

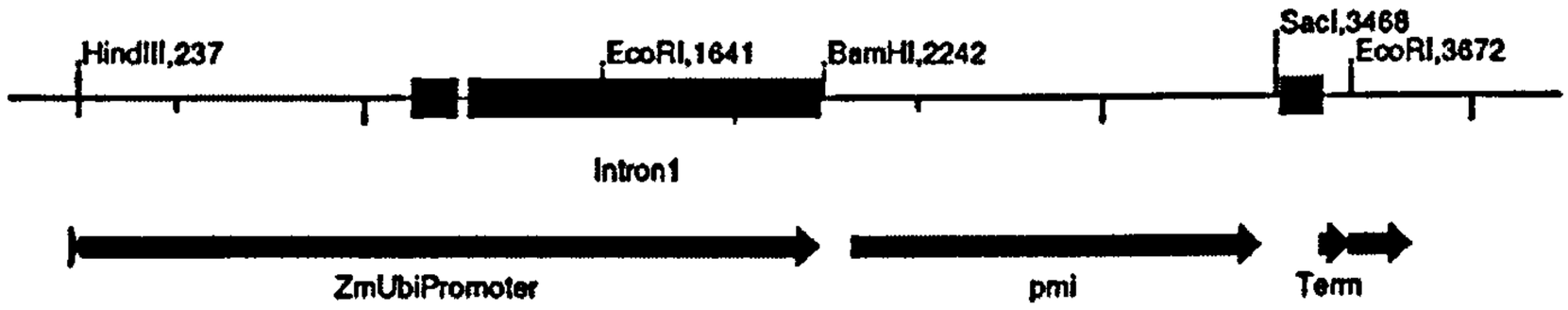


Figure 46. Diagram showing the fragment preparation of construct pCIB9818.

6.2.1.7. Construction of PMI RNAi Vectors

For the generation of PMI RNAi entry clones a 300 bp PCR product was synthesised using the *AttB* linked (underlined) gene specific primers (GSP) PMI F; 5'GGGGAC AAGTTTGTACAAAAAAGCAGGCTCAATGCAAAAACTCATTAACTCAGTGC 3' and PMI R; 5'GGGGACCACTTTGTACAAGAAAGCTGGGTATGGATGAAC CTGAATGGAGA with the pCIB9818construct as a template. BP reactions between the *AttB* linked PCR products and the pDONR™221 entry vector were carried out according to the Invitrogen™ Gateway™ instruction manual. Several of the resulting transformed colonies were grown and the DNA extracted using the QIAprep® spin miniprep kit. DNA from the Gateway™ compatible RNAi destination vector

pCIB11079 (provided by Syngenta) containing an inverted *Att*R repeat separated by a rice intron and driven by the maize ubiquitin intron 1 promoter (Christensen *et al.*, 1992) was used for the generation of the PMI RNAi transformation vectors. The LR recombination reaction between the PMI pDONR™221 entry clones and the Gateway™ compatible RNAi destination vector pCIB11079 was carried out as described in the instruction manual. Primers were designed to sequence the inverted repeats and several colonies were sequenced to verify that the PCR products had integrated with the pCIB11079 RNAi destination vector. DNA was extracted from the sequence verified colonies using the QIAprep® spin miniprep kit according to manufacturer's instructions and resuspended in TE buffer to a final concentration of 1 µg/µl ready for transformation.

6.2.1.8. Donor Plant Material

The Bolero donor plant material was prepared as described in 6.2.1.3.

6.2.1.9. Transformation

Transformation was carried out as described in section 6.2.1.4. Except after gene delivery and osmotic treatment the wheat IEs were then moved to the callus induction (3MS3S) medium (2.2.3.2.) and incubated for three weeks.

6.2.1.10. *Pmi* Selection and Regeneration

The embryogenic tissue that developed during the three-week initiation period was separated from the non-embryogenic tissue and placed on a regeneration/selection (NG) medium (2.2.3.3.). For *pmi* selection in wheat, the carbohydrate source in the NG medium was altered to contain 10 g/L mannose and 5 g/L sucrose (NG 1M 0.5S). Following growth in the light at 24°C for two weeks, emerging shoots were scored and transferred to MS medium with 15 g/L sucrose plus 10 g/L mannose, 300 mg/L glutamine and 150 mg/L asparagine (MS 1.5M 1S). Shoots which continued to grow were transferred to larger containers (Greiner Bio-one) with a medium consisting of ½ strength (2.15 g/L) MS salts (Sigma, M 5524), full strength MS vitamins and 15 g/L mannose (½ MS 1.5M 1A) and grown for a further 2 weeks. Shoots that continued to grow vigorously were considered to be transgenic and were sampled (6.2.2.5.) to calculate the transformation efficiency.

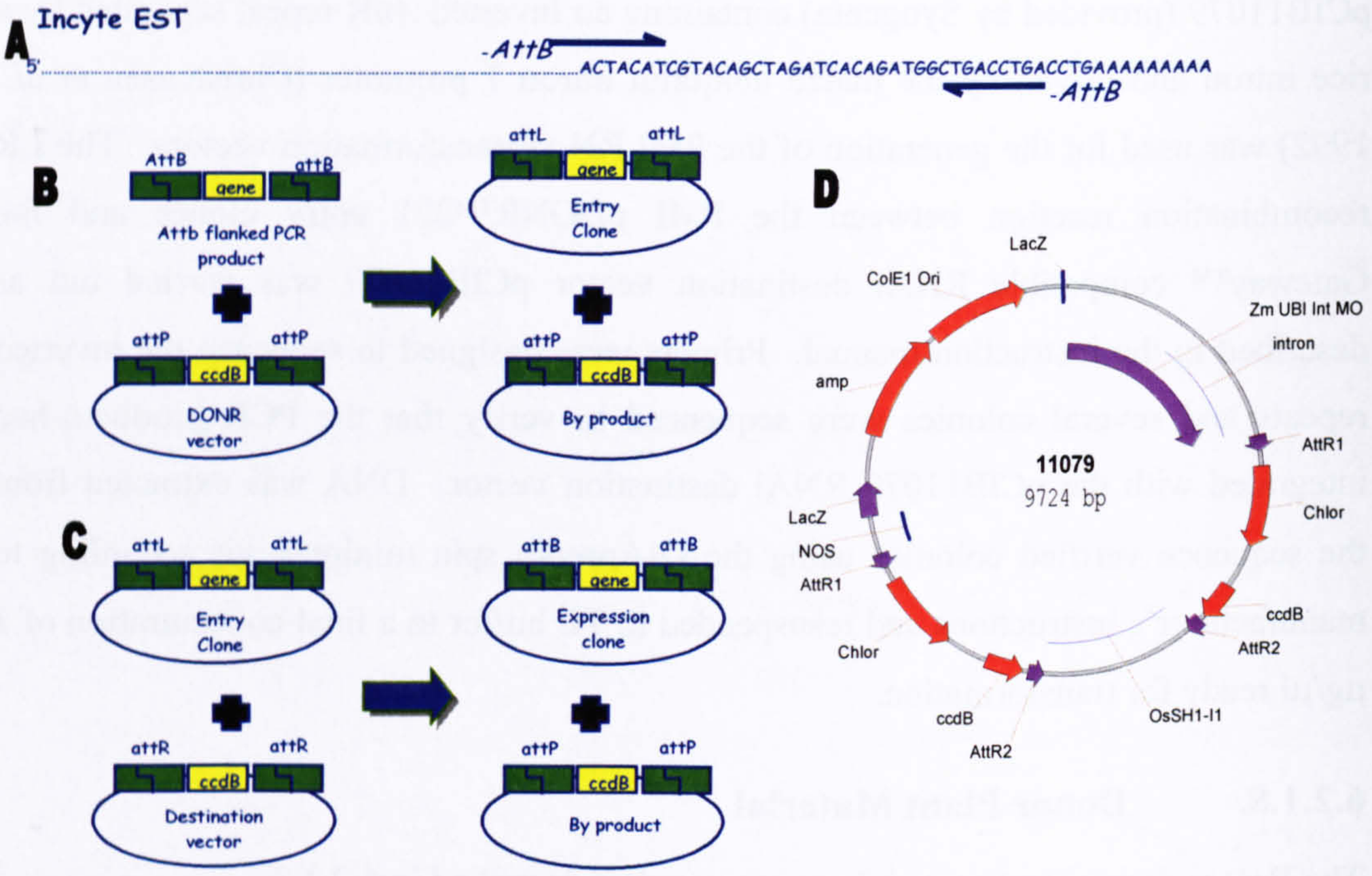


Figure 47. Diagram showing the construction of the RNAi transformation vectors for the shortlisted gene candidates. A) Synthesis of RNAi PCR product using *AttB* linked GSPs. B) BP reactions between the *AttB* linked PCR products and the pDONR™221 entry vector to give the entry clones. C) LR recombination reaction between the pDONR™221 entry clones and the Gateway™ compatible RNAi destination vector pCIB11079 to give the expression clones for transformation. D) Map of the Gateway™ compatible RNAi destination vector pCIB11079 showing the maize ubiquitin promoter, the *AttR* sites for the introduction of the inverted repeat gene sequence, the rice intron spacer and the *nos* terminator sequences.

6.2.2. RNAi Functional Analysis

Following the establishment that co-bombardment of a selectable gene marker/fluorescent protein with an RNAi construct can induce gene silencing in wheat, a series of transformation experiments were setup to test the effect of silencing the endogenous expression of each of the shortlisted candidate genes. The full procedure used by the Syngenta wheat transformation team for generating transgenic wheat was followed for these experiments. The rationale behind this approach was to compare the efficiency of regenerating shoots from calli transformed with a PMI construct against calli co-transformed with a PMI and a gene-specific RNAi construct. Any resulting reduction in the regeneration frequency could therefore be attributed to the silencing of the gene in question implicating it as being essential for SE.

6.2.2.1. RNAi Transformation Vector Preparation

A *pmi* fragment preparation as described in section 6.2.1.6. was used for transformation along with gene specific RNAi transformation vectors. For the generation of the RNAi vectors (Fig. 33.), a PCR product for each of the chosen genes was amplified using the *AttB* linked gene-specific primers (Table 19) and purified plasmid DNA from the respective EST clones as the template. Primers were designed within the coding region of the EST clones using the sequence information in the Phytoseq database. BP reactions between the gene-specific *AttB* linked fragments and the pDONR™221 entry vector were carried out according to the Invitrogen™ Gateway™ instruction manual as described in 6.2.1.2. The LR recombination reaction between the different EST pDONR™221 entry clones and the Gateway™ compatible RNAi destination vector pCIB11079 was also carried out as described in 6.2.1.2. Several colonies were sequenced to verify that the PCR products had integrated with the pCIB11079 RNAi destination vector as described in 6.2.1.2. DNA was extracted from the sequence verified colonies using the QIAprep® spin miniprep kit according to manufacturer's instructions and resuspended in TE buffer to a final concentration of 1 μg/μl ready for transformation.

Table 19: *Attb* linked GSPs used for generating gene specific RNAi vectors.

Gene Name	Forward Primer (5' to 3')	Reverse Primer (5' to 3')	PCR Product (BPs)
DsPTP1	5' CTCCTAAGCTTG GGTCTATTATCGC 3'	5' GGTCGTCCTTG CCTACTTGATG 3'	485
ABRE BF	5' TACTAGGTTCTT ACGGCTCCCCTC 3'	5' GACCTGAACCA GGAACTGGTGAG 3'	337
OSE731	5' CCAGTCTTGTCG GATCTACCGTGGT 3'	5' ACACCCACGAGT TCGAGGTCAAGGG 3'	400
SCARECROW	5' AGAACCAACTC ATCACACCTTGCC 3'	5' CATTGCCGAC CAGGTCATGG 3'	245
SE RELATED	5' AGTACAGACGCT TCACTTCAAAGGG 3'	5' GTATGGTGGAA GAAGCAACATGC 3'	374
SER-THR PK	5' GGGTGTTGGAG CAAGAGGACTT 3'	5' GCAGCTGATCGG TTCTTCACTCTAT 3'	500
INCYTE EST	5' CGACAATCATCC AAAGACCAACTC 3'	5' TCGGATGATG GGCAATGTGG 3'	416
PP2C	5'AGGCTAAGGGGGT ATACAGAAACAC3'	5' AAAGTTGCCC GGCGACAGAT 3'	411
HYPO PRO	5' CAAAGAAAAGA GGGTGGGAATCTG 3'	5' AAATCGAGC TGCGAGCCTTG 3'	500
LRRK	5' TGTTGCCACTC AATTGGTTGTG 3'	5' TGCACTTCTCC AGTTCCTCACTG 3'	533
BZIP	5' TTATCCATATCC AGTGGGTAGGAGC 3'	5' GAGGATGCTC TCCAACAGGGA 3'	592
PICKLE	5' GAGAAGCTCCA TGGAATGCTGA 3'	5' CAACCCACCAGCT CTAGTAGAAAGA 3'	552
Att B	5'GGGGACAAGTTTGTAC AAAAAAGCAGGCTCAA3'	5'GGGGACCACTTTGTA CAAGAAAGCTGGGTA3'	

6.2.2.2. Plant Material

The spring variety Bolero was used as the wheat material for the RNAi gene testing experiments and was prepared for transformation as described in section 6.2.1.3.

6.2.2.3. Transformation

Transformations were carried out as described in 6.2.1.4. using 1μg of DNA for the PMI control transformations and an equal Molar concentration of PMI and PMI RNAi constructs. The bombarded wheat embryos were moved to callus induction (3MS3S) medium (2.2.3.2.) following high osmotic treatment and incubated in the dark at 25°C/40% humidity for three weeks.

6.2.2.4. Selection and Regeneration

PMI selection and regeneration was carried out as described in 6.2.1.10. Shoots that regenerated on MS 1.5M 1S medium were transferred to larger containers (Greiner

Bio-one) with a medium consisting of ½ strength (2.15 g/L) MS salts (Sigma, M 5524), full strength MS vitamins and 15 g/L mannose (referred to as ½ MS 1.5M 1A) and grown for a further 2 weeks. Shoots that continued to grow vigorously were sampled for the presence of the PMI and PMI RNAi constructs by PCR using primers specific to the *pmi* gene and nopaline synthase (*nos*) terminator as described below (6.2.2.5.)

6.2.2.5. Confirmation of Transgenic Lines

Genomic DNA for the PCR assays was isolated from wheat leaf tissue of transgenic and non-transgenic control plants. Approximately 7 x 25 mm of leaf material were sampled directly into 1.2ml 96 deep-well plates (Greiner Bio-One), containing 4mm steel beads, and was frozen and stored at –80°C overnight and then ground in the Geno-grinder (GenoGrinder) at 1100strokes/min for 2 minutes. Genomic DNA was extracted using the Wizard® Magnetic 96 DNA Plant System (Promega, FF3761) according to manufacturer's instruction on a Biomek-FX liquid handling workstation (Beckman-Coulter, Inc., Fullerton, CA). The DNA was eluted in 100 µl of TE (10 mM Tris-HCI, pH 8.0, 1 mM EDTA) plus 0.1% BSA. PCRs were carried out in 96-well microtitre plates (ABGene, AB-0796). For each sample, 5 µl of extracted genomic DNA was combined with 1 Ready-to-Go™ PCR bead (Amersham Pharmacia Biotech, 27-9557) suspended in 25µl of water [giving a final concentration of 2.5 units puReTaq DNA Polymerase, 200µM dNTP (dATP, dCTP, dTTP and dGTP), 10mM Tris-HCl (pH 9.0), 50mM KCl and 1.5mM $MgCl_2$]. The PCR primers were added to each sample at a final concentration of 4ng/µl. The plates were sealed with plastic plate seals (StarLabs, E2796-9794) and the PCR reaction run in the following programme: 94°C for 5 minutes followed by 30 cycles of 94°C for 45 seconds, 65°C for 1 minute, 72°C for 1 ½ minutes. Samples were run on a 0.8% agarose gel to detect for the presence of PCR products.

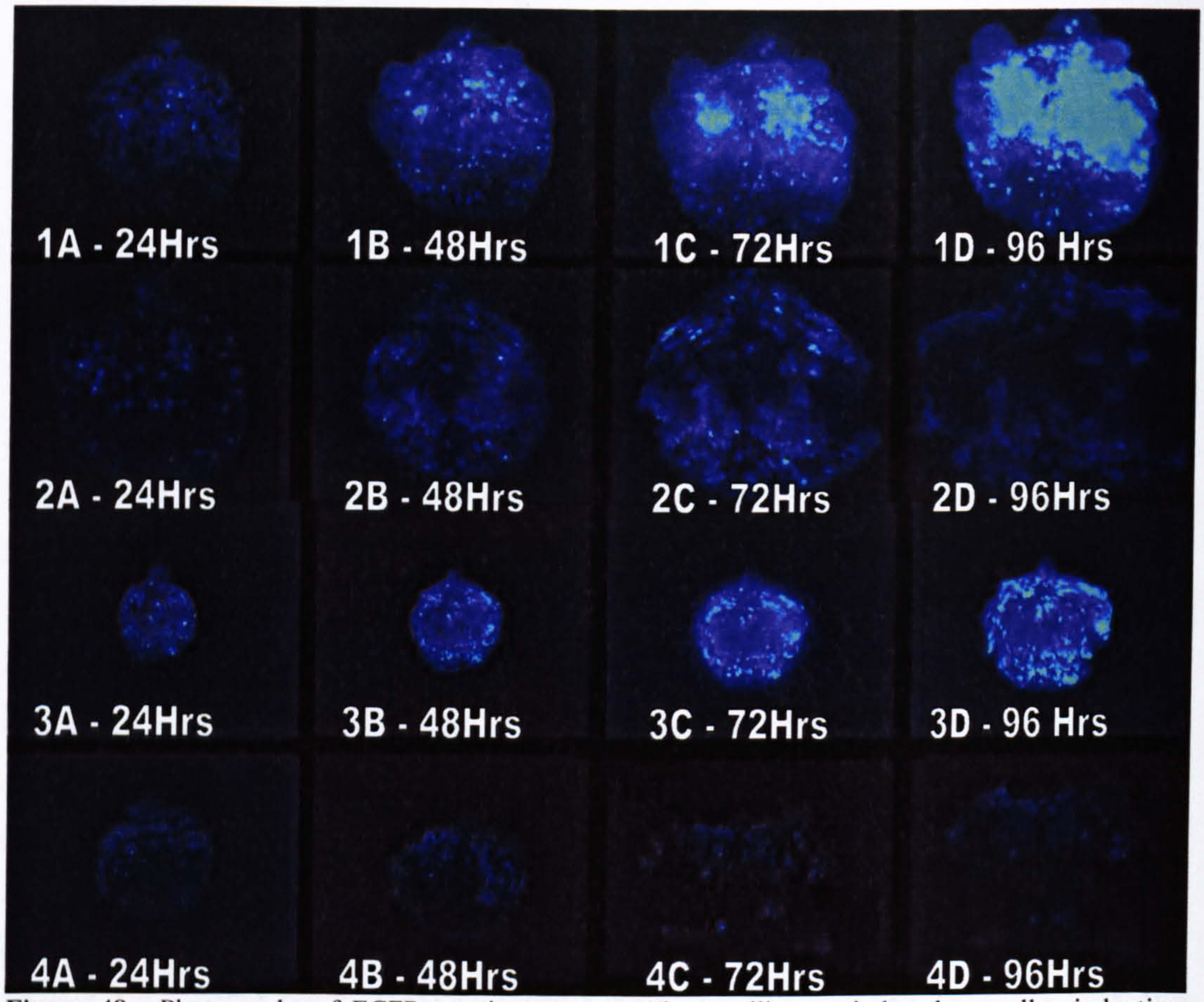


Figure 48. Photographs of ECFP transient assay. Wheat calli were induced on callus induction medium (3MS3S) for one week. 1) Calli bombarded with the ECFP overexpression construct (pCIB11251). 2) Calli co-bombarded with the ECFP overexpression construct (pCIB11251) and ECFP RNAi construct. 3) Replicated experiment of calli bombarded with the ECFP overexpression construct (pCIB11251). 4) Replicated experiment of calli co-bombarded with the ECFP overexpression construct (pCIB11251) and ECFP RNAi construct. Photographs were taken at 24 hourly intervals following transformation as follows: A; 24 Hrs, B; 48 Hrs, C; 72 Hrs and D; 96Hrs.

6.3. RESULTS

6.3.1. RNAi transient Assays

As previously discussed, an important prerequisite for obtaining reliable functional data about the ISH shortlisted genes via an RNAi based approach was to establish the efficacy of this technology in wheat which was unknown at the time. The strategy employed to do this was to set up two separate bombardment experiments as shown in Figure 45. Firstly, a co-bombardment experiment that tested the silencing of a transiently expressed ECFP visual marker was carried out. And secondly, a co-bombardment experiment that tested the silencing of the *pmi* selectable marker on regenerating shoots under mannose selection was also conducted. The latter experiment serves the additional purpose of providing a reliable control data set for undertaking and interpreting the RNAi gene silencing experiments.

6.3.1.1. ECFP Transient Assay

Following gene delivery and osmotic treatment bombarded calli were cultured on callus induction medium (3MS3S) for one week during which time they were evaluated by fluorescence microscopy as described in section 6.2.1.4. The fluorescence images (Fig. 48) were taken of the calli at 24 hourly intervals post bombardment to monitor and record the transient expression levels of the ECFP protein. This allowed for the direct comparison of calli bombarded with the ECFP overexpression construct alone alongside calli that had been co-bombarded with the ECFP overexpression construct (pCIB11251) and the ECFP RNAi construct. It can be seen in Figure 48 that calli bombarded with the linearised pCIB11251 plasmid exhibit strong transient expression of the ECFP protein. The expression of the ECFP protein is detected after 24 hours post-bombardment in a small number of foci (Fig. 34 1A + 3A) and increases in its expression until it peaks at 4 days (96Hrs) post-bombardment where it is strongly expressed in a large number of cells in the calli. In contrast, calli that were co-bombarded with the linearised pCIB11251 plasmid and the ECFP RNAi construct exhibit much weaker transient expression of the ECFP protein. The expression of the ECFP protein is also detected after 24 Hrs post-bombardment but in a smaller number of foci (Fig. 48 2A + 4A). In Figure 48 2A-D the transient expression of ECFP increases much less markedly at 48 and 72 hours and actually

subsides at 96 hours. Furthermore, in Figure 48 4A-D there is little or no increased transient expression of ECFP after 48 hours, a decrease in the expression at 72 hours and complete silencing at 96 hours.

6.3.1.2. PMI Transient Assay

Following gene delivery, osmotic treatment and culture on callus induction medium (3MS3S) for three weeks, embryogenic tissue was separated from the non-embryogenic tissue and cultured on a regeneration/selection medium (NG 1M 0.5S) in the light at 24°C for two weeks, as described in sections 6.2.1.8-6.2.1.10. The number of emerging shoots that were scored and transferred to mannose containing medium (MS 1.5M 1S) can be seen in tables 20 + 21. Shoots which continued to grow in ½ MS 1.5M 1A medium in larger containers for a further 2 weeks were considered to be transgenic and were sampled (6.2.2.5.) to calculate the transformation efficiency as seen in tables 22 + 23.

Table 20: Regeneration efficiencies of calli bombarded with PMI.

Construct	Experiment:	No. Calli Bombarded	No. Shoots	No. shoots per calli
PMI	E 478	100	435	4.34
PMI	E 478	100	485	4.85
PMI	E 451	100	421	4.21
Total		**300**	**1341**	
Mean		100	447	**4.47 ± 0.34**

Table 21: Regeneration efficiencies of calli bombarded with PMI RNAi.

Construct	Experiment:	No. Calli Bombarded	No. Shoots	No. shoots per calli
PMI RNAi	E 478	100	205	2.05
PMI RNAi	E 478	100	195	1.95
PMI RNAi	E 451	100	263	2.63
Total		**300**	**663**	
Mean		100	221	**2.21 ± 0.37**

Tables 20 and 21 show the effect on regeneration efficiencies of calli bombarded with the PMI marker and PMI RNAi constructs. It can be seen that from a total of 300 bombarded calli an average of 4.47 shoots per callus regenerate when bombarded with the PMI marker construct alone. In contrast, when a total of 300 calli are co-bombarded with the PMI marker and PMI RNAi constructs the average number of shoots per callus that regenerate is reduced to 2.21.

Table 22: Transformation efficiencies of calli bombarded with PMI.

Construct	Experiment	No. Calli Bombarded	No. Shoots (PMI +ve)	No. Shoots (Nos +ve)	Transformation efficiency
PMI	E 478	100	85		85
PMI	E 478	100	86		86
PMI	E 451	100	81		81
Total		**300**	**252**		
Mean		100	84		**84% ± 2.65%**

Table 23: Transformation efficiencies of calli bombarded with PMI RNAi.

Construct	Experiment	No. Calli Bombarded	No. Shoots (PMI +ve)	No. Shoots (Nos +ve)	Transformation efficiency
PMI RNAi	E 478	100	37	21	37%
PMI RNAi	E 478	100	28	19	28%
PMI RNAi	E 451	100	30	16	30%
Total		**300**	**95**	**56**	
Mean		100	31.67	18.67	**32.5% ± 4.73**

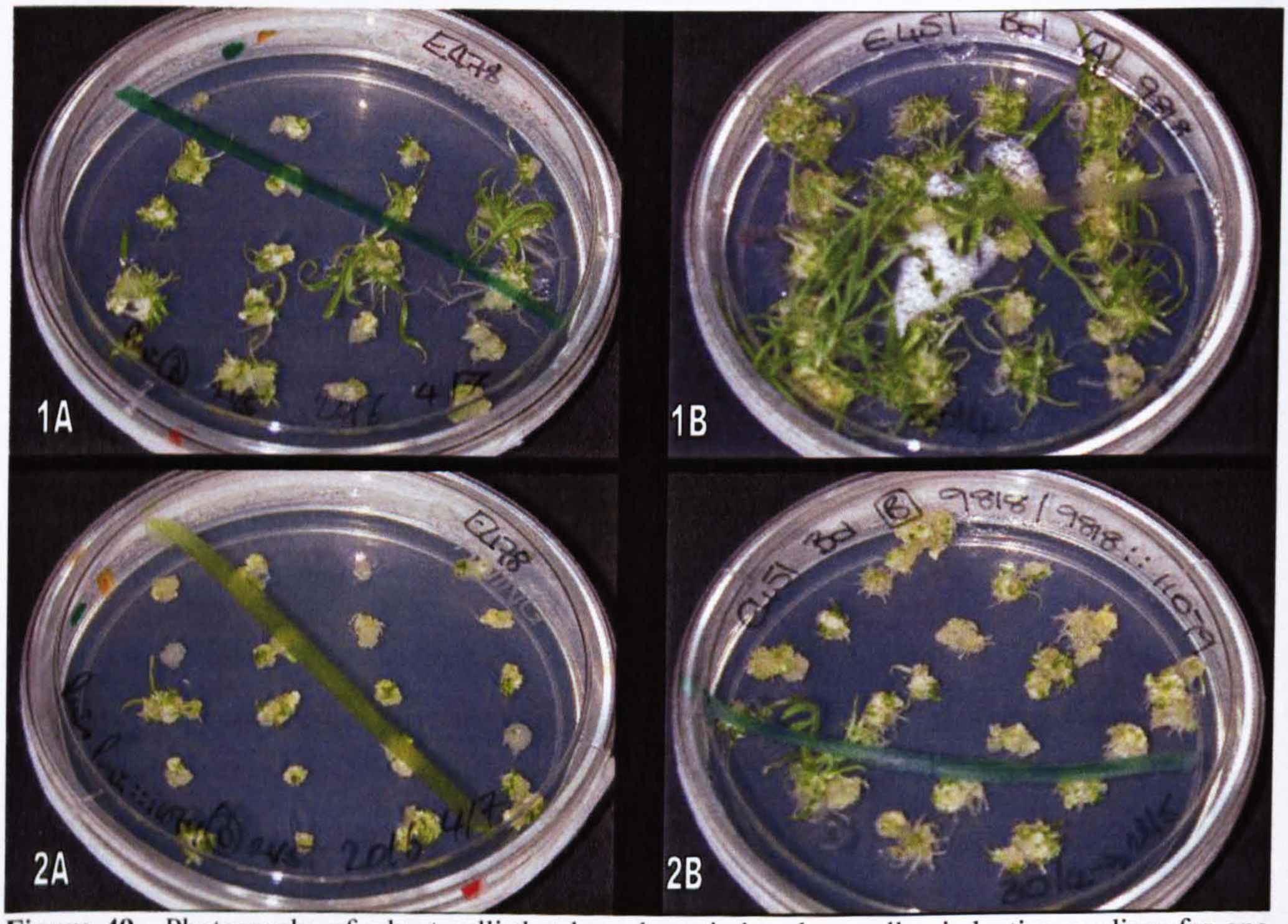


Figure 49. Photographs of wheat calli that have been induced on callus induction medium for one week and transformed with 1) a PMI overexpression construct (1A+1B) 2) co-transformation of a PMI overexpression construct and a PMI RNAi construct. Following a further 3 week callus induction period calli were placed on regeneration (NG) medium under light for 2 weeks (1A and 2A). Calli were transferred to mannose selection medium and photographs taken after 2 weeks (1B and 2B).

Tables 22 and 23 show the effect on the transformation efficiencies of calli bombarded with the PMI marker and PMI RNAi constructs. These results show that from a total of 300 calli bombarded with the PMI marker construct alone, 252 shoots regenerated that tested positive for *pmi*, giving an average transformation efficiency of 84 per cent. By contrast, when a total of 300 calli were co-bombarded with the PMI and PMI RNAi constructs the average number of shoots that tested positive for *pmi* was reduced to 95, giving an average transformation efficiency of 32.5 per cent. Furthermore, a significant number of these *pmi*-positive shoots (39 from 95) lacked any detectable presence of the PMI RNAi construct, effectively resulting in a transformation efficiency of 18.67 per cent. In addition, the effects of bombarding regenerating wheat calli with these constructs can be observed in Figure 49 which shows the results from two separate bombardment experiments. 1A and 1B show the regeneration of shoots from calli that have been bombarded with the PMI construct alone, whereas 2A and 2B show the regeneration of shoots from calli that have been co-bombarded with both the PMI and PMI RNAi constructs. It can be seen that there are a large number of vigorously growing shoots emerging from calli bombarded with the PMI construct. By contrast there is a significant reduction in the number of regenerating shoots in the co-bombarded calli, and those shoots that do regenerate appear less vigorous than those bombarded with PMI alone.

6.3.2. RNAi Gene Testing

Having established that co-bombardment of a selectable gene marker/fluorescent protein with an RNAi construct can induce gene silencing in wheat, a series of transformation experiments was set up to test the effect of silencing the endogenous expression of each of the shortlisted candidate genes. The full procedure used by the Syngenta wheat transformation team for generating transgenic wheat was followed for these experiments. The rationale behind this approach was to compare the efficiency of regenerating shoots from calli transformed with the PMI construct against calli co-transformed with the PMI and each of the shortlisted gene-specific RNAi constructs. Any resulting reduction in the regeneration frequency could therefore be attributed to the silencing of the target gene in question, thereby implicating its functional role in SE. The results for these gene-testing experiments are presented in Tables 24 and 25, as a calculation of the regeneration and transformation efficiencies respectively and there are photographs illustrating the observed phenotypes in Figure 50.

Table 24: Regeneration efficiencies of calli bombarded with gene-specific RNAi constructs.

Construct	Experiment	No. Calli Bombarded	No. Shoots	No. shoots per calli
OSE 731 RNAi	E 489	100	244	2.44
OSE 731 RNAi	E 489	100	225	2.25
OSE 731 RNAi	E 445	100	288	2.88
Total		**300**	**757**	
Mean		100	252.33	**2.52 ± 0.32**
LRRK RNAi	E 489	100	200	2
LRRK RNAi	E 489	100	226	2.26
LRRK RNAi	E.445	100	286	2.86
Total		**300**	**712**	
Mean		100	237.33	**2.37 ± 0.44**
ABREBF RNAi	E 489	100	318	3.18
ABREBF RNAi	E 489	100	314	3.14
ABREBF RNAi	E 434	80	320	4
Total		**280**	**952**	
Mean		93.33	317.3	**3.4 ± 0.49**
SCR RNAi	E 493	100	342	3.42
SCR RNAi	E 493	100	285	2.85
SCR RNAi	E 534	100	332	3.32
Total		**300**	**959**	
Mean		100	319.67	**3.2 ± 0.3**
STPK RNAi	E 489	100	346	3.46
STPK RNAi	E 489	100	402	4.02
STPK RNAi	E 534	100	285	2.85
Total		**300**	**1033**	
Mean		100	344.33	**3.44 ± 0.59**
B.ZIP RNAi	E 502	80	296	3.7
B.ZIP RNAi	E 511	80	327	4.09
Total		**160**	**623**	
Mean		80	311.5	**3.89 ± 0.28**
PICKLE RNAi	E 502	80	442	5.52
PICKLE RNAi	E 511	80	340	4.25
Total		**160**	**782**	
Mean		80	391	**4.89 ± 0.90**
PMI	E 478	100	435	4.34
PMI	E 478	100	485	4.85
PMI	E 451	100	421	4.21
Total		**300**	**1341**	
Mean		100	447	**4.47 ± 0.34**

6.3.2.1. RNAi Regeneration Efficiencies

Table 24 shows the effect of co-transforming each of the gene-specific RNAi constructs on regeneration efficiencies when compared to a PMI control. The bombardment of 300 calli with the PMI construct produced a total number of 1341 shoots at an average of 4.47 shoots per callus. It can be seen that, with the exception of PICKLE, there is a general reduction in the regeneration efficiencies of all the genes tested. The most dramatic of these reductions was observed with LRRK RNAi constructs where the co-bombardment of 300 calli produced a total number of 712 shoots, at an average of 2.37 shoots per callus. This is an average of 2.10 shoots per calli less than when bombarded with PMI alone. This was closely followed by the OSE 731 RNAi construct where the co-bombardment of 300 calli resulted in a total of 757 regenerating shoots, at an average of 2.52 shoots per callus, an average of 1.95 shoots per calli less than when bombarded with PMI alone. Additional reductions in regeneration efficiencies were also observed with the SCR RNAi constructs where the 300 calli that were co-bombarded, resulted in a total number of 959 shoots at an average of 3.2 shoots per callus an average reduction of 1.27 shoots per calli less than when bombarded with PMI alone. Reductions were also observed with the ABREBF RNAi construct where the co-bombardment of 280 calli (a plate of 20 calli was lost due to contamination) resulted in a total of 952 shoots with an average of 3.4 shoots per callus, a reduction of 1.07 shoots per calli when compared with PMI. The co-bombardment of 300 calli with the PMI and STPK RNAi constructs also resulted in a reduction of regeneration efficiency, with a total of 1033 shoots at an average of 3.44 shoots per callus, a reduction of 1.03 shoots per calli when compared with PMI. For the co-bombardment of calli with PMI and B.ZIP or PICKLE RNAi constructs only 160 were recovered for analysis due to contamination. A slight reduction in regeneration efficiency was still observed in the 160 calli that were co-bombarded with PMI and B.ZIP RNAi constructs which resulted in a total of 623 shoots regenerating at an average of 3.89 shoots per callus. A reduction of 0.58 shoots per calli when compared with PMI alone. For the co-bombardment of calli with PMI and PICKLE RNAi constructs a slight increase in regeneration efficiency was observed. Out of the 160 calli co-bombarded a total of 782 shoots at an average of 4.89 shoots per callus were shown to regenerate which is an average of 0.42 shoots per calli more than when bombarded with PMI alone.

Table 25: Transformation efficiencies of calli co-bombarded with gene-specific RNAi Constructs.

Construct	Experiment	No. Calli Bombarded	No. Shoots (PMI +ve)	No. Shoots (Nos +ve)	Transformation efficiency
PMI	E 478	100	85	-	85
PMI	E 478	100	86	-	86
PMI	E 451	100	81	-	81
Total		**300**	**252**	-	
Mean		100	84		**84% ± 2.65%**
PMI RNAi	E 478	100	37	21	37%
PMI RNAi	E 478	100	28	19	28%
PMI RNAi	E 451	100	30	16	30%
Total		**300**	**95**	**56**	
Mean		100	31.67	18.67	**32.5% ± 4.73%**
OSE 731 RNAi	E 489	100	37	0	37%
OSE 731 RNAi	E 489	100	40	0	40%
OSE 731 RNAi	E 445	100	35	0	35%
Total		**300**	**112**	**0**	
Mean		100	37.33	0	**37.3% ± 2.52%**
LRRK RNAi	E 489	100	22	11	22%
LRRK RNAi	E 489	100	32	24	32%
LRRK RNAi	E 445	100	30	11	30%
Total		**300**	**84**	**46**	
Mean		100	28	15.33	**28% ± 5.29%**
ABAREBF RNAi	E 489	100	42	20	42%
ABAREBF RNAi	E 489	100	50	30	50%
ABAREBF RNAi	E 434	100	68	16	68%
Total		**200**	**160**	**66**	
Mean		100	53.33	22	**53% ± 13.31%**
SCR RNAi	E 493	100	64	31	64%
SCR RNAi	E 493	100	48	29	48%
SCR RNAi	E.440	100	43	9	43%
Total		**300**	**155**	**69**	
Mean		100	51.67	23	**52% ± 10.97%**
STPKI RNAi	E 489	100	58	43	58%
STPKI RNAi	E 489	100	46	28	46%
STPKI RNAi	E 440	100	51	28	51%
Total		**300**	**155**	**99**	
Mean		100	51.67	33	**52% ± 6.03%**
B.ZIP RNAi	E 502	80	41	31	41%
B.ZIP RNAi	E 511	80	45	30	45%
Total		**160**	**86**	**61**	
Mean		80	43	30.5	**43% ± 2.83%**
PICKLE RNAi	E 502	80	48	30	48%
PICKLE RNAi	E 511	80	45	20	45%
Total		**160**	**83**	**50**	
Mean			41.5	25	**46.5% ± 2.12%**

6.3.2.2. RNAi Transformation Efficiencies

Table 25 demonstrates the effect on transformation efficiency as a result of co-transforming calli with the PMI and gene-specific RNAi constructs. The transformation of 300 calli with the PMI control resulted in a total of 252 shoots testing *pmi*-positive, giving an average transformation efficiency of 84 per cent. This can be seen to be reduced to an average of 32.5 per cent when a total of 300 calli are co-bombarded with the PMI and PMI RNAi constructs. Consistent with the results of the regeneration frequencies, the most dramatic reduction in transformation efficiencies from the gene-specific RNAi constructs is achieved by the co-transformation of the PMI and LRRK RNAi constructs. Of the 300 calli that were co-transformed, a total of 84 shoots tested positive for *pmi*, 46 of which also tested positive for the presence of the RNAi construct (i.e. *Nos* terminator sequence). This resulted in an average transformation efficiency of 28 per cent, slightly less than for PMI RNAi. Similarly, the co-transformation of 300 calli with the PMI and OSE 731 RNAi constructs resulted in a total of 112 shoots that tested positive for *pmi*, none of which tested positive for the presence of the RNAi construct, to give an overall transformation efficiency of 37.3 per cent, slightly more than for PMI RNAi. The co-transformation of the PMI and AREBF RNAi constructs resulted in a total of 160 shoots testing positive for *pmi* from 300 calli, with 66 of these testing positive for the RNAi construct to give an average transformation efficiency of 53 per cent. This was closely matched by the co-transformation of the PMI and SCR RNAi constructs resulting in a total of 155 shoots from 300 calli testing positive for *pmi*, with 69 of these also testing positive for the RNAi construct, to give an average transformation efficiency of 52 per cent. Similarly, the co-transformation of 300 calli with the PMI and STPK1 RNAi constructs resulted in a total of 155 shoots, of which 99 also tested positive for the RNAi construct, to give an average transformation efficiency of 52 per cent. In addition, the co-transformation of 160 calli with the PMI and B.ZIP RNAi constructs resulted in a total of 86 shoots, of which 61 tested positive for the RNAi construct, to give an average transformation efficiency of 53.75 per cent. And finally the co-transformation of 160 calli with the PICKLE RNAi construct resulted in a total of 83 *pmi* positive shoots, of which 50 tested positive for the RNAi construct, giving an average transformation efficiency of 51.8 per cent.

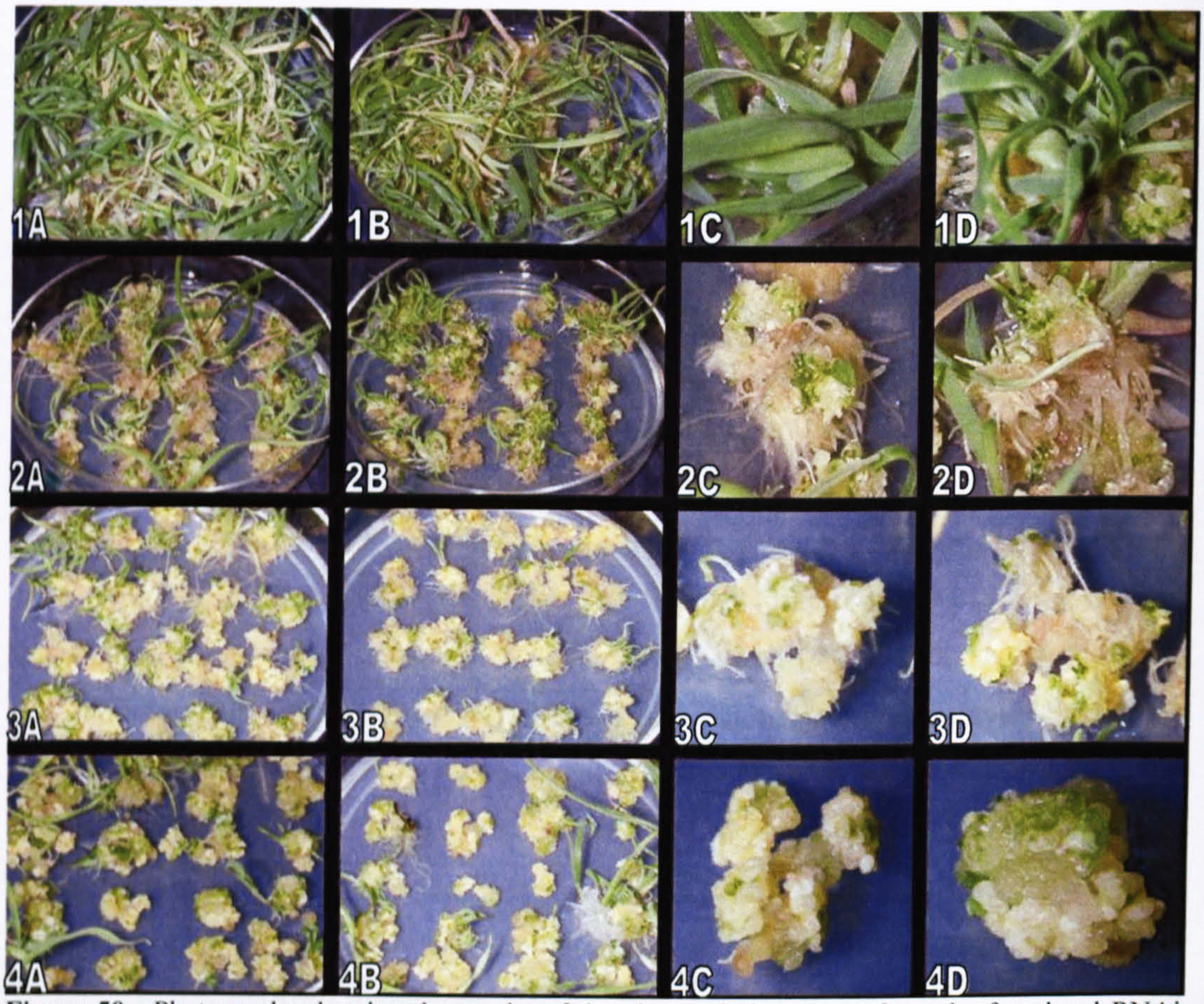


Figure 50. Photographs showing the results of the observed phenotypes from the functional RNAi experiments. Calli were initiated and regenerated following transformation as described in materials and methods. The resulting material was transferred to mannose selection medium and photographed after four weeks. 1) A-D = Regeneration of calli transformed with PMI alone 2) A-D = regeneration of calli co-transformed with PMI and ABREBF RNAi. 3) A-D = regeneration of calli co-transformed with PMI and LRRK RNAi. 4) A-D = regeneration of calli co-transformed with PMI and OSE 731 RNAi.

6.3.2.3. RNAi Phenotypes

The photographs presented in Figure 50 demonstrate the observable effects on regenerating calli that have been transformed with some of the RNAi constructs. The first set of photographs (1A-1B) show the shoot regeneration of calli that have been transformed with the PMI construct alone which can be seen to have well established, vigorously growing shoots. The reduction in shoot regeneration can be seen in each of the examples of the PMI and RNAi co-transformed calli when compared to PMI alone. Photographs 2A-2D show the effect on calli that have been co-transformed with PMI and ABREBF RNAi constructs. It can be seen that there is a significant reduction in the number of shoots that regenerate, and those that do appear to be less vigorous. Close-up images (2C-2D) show the calli to be friable and non-embryogenic in appearance, with a number of short shoot-like structures emerging from the calli. Photographs 3A-3D demonstrates the effect of co-transformation with the PMI and LRRK RNAi constructs. It can be seen that very few shoot structures regenerate from these calli and those that do, appear much less vigorous than the PMI control. Close-up images of these calli (3C-3D) show the lack of adult shoot structures and the callus appears drier and less embryogenic than normal. Finally, photographs 4A-4D show the effect on calli that have been co-transformed with the PMI and OSE 731 RNAi constructs. It can be seen that there is a significant reduction in the number of regenerating shoots when compared with the PMI control (4A and 4B), however, those shoots that do emerge appear fairly healthy and vigorous. Close-up images (4C-4D) of the calli reveal an interesting phenotype in which there are no adult shoot or root structures emerging from the calli at all. Furthermore, the callus has a healthy embryogenic appearance, with a large number of somatic embryo-like structures.

6.4. Discussion

6.4.1. RNAi Transient Assays

As previously discussed, a transient RNAi assay strategy was used to test the effect of RNAi silencing in wheat which consisted of a selectable (PMI) marker and a visual (ECFP) marker co-bombarded with RNAi gene silencing constructs. The results are discussed here.

6.4.1.1. ECFP Transient Assay

The results from the ECFP transient assay provides strong evidence for the effective silencing of ECFP expression by the co-transformed RNAi construct. Figure 48 shows photographs of wheat calli that have been bombarded with an ECFP overexpression construct in parallel with the co-bombardment of the ECFP and ECFP RNAi constructs. It can be seen that wheat calli bombarded with ECFP show an increasingly high transient expression of the fluorescent protein which peaks after four days post bombardment. By contrast, calli that have been co-bombarded with the ECFP and ECFP RNAi constructs show a significantly reduced transient expression of the fluorescent protein. Furthermore, rather than showing an increase in the expression of the fluorescent protein the expression appears to have been reduced by four days post bombardment demonstrating that silencing of the ECFP product is occurring. The low level expression of the fluorescent protein in co-bombarded calli is probably best explained by the existence of some cells that have received the ECFP construct alone which would allow the ECFP gene to be expressed uninhibitedly. Whereas, those cells that have received both constructs would have ECFP expression inhibited by the RNAi construct resulting in fewer cells expressing the ECFP protein. In addition, it is possible that in those cells that receive both constructs the procedure of co-bombarding constructs with constitutive promoters causes residual expression of the ECFP protein due to the inability of the RNAi silencing constructs to fully silence the ECFP gene, resulting in the observed low level expression of ECFP in co-bombarded calli. To further define the molecular action of RNAi silencing, further experiments could be performed in which the RNAi construct is linked with another visible marker, for example, dsRED, to determine whether cells that are expressing ECFP have been co-transformed with the RNAi construct, or not. Nevertheless, these results provide strong evidence that the co-bombardment of the linearised pCIB11251 plasmid and ECFP RNAi expression constructs leads to the effective silencing of the transiently expressed ECFP protein by the ECFP RNAi expression construct in wheat calli.

6.4.1.2. PMI Transient Assay

The results from the PMI transient assay provide further evidence for the effective silencing of a marker gene by a co-transformed RNAi construct in wheat calli. Firstly, it can be seen in Figure 49 that when the PMI overexpression construct is

delivered to wheat calli alone it results in the regeneration of a number of vigorous shoots that continue to grow on mannose selectable media. By contrast, wheat calli that were co-bombarded with the PMI overexpression and RNAi constructs show a significantly reduced number of regenerative shoots, and those which do regenerate appear less vigorous and healthy than the PMI control. The regeneration and transformation efficiencies that were calculated over three separate experiments further support this observation. Firstly, the number of emerging shoots that were scored after regeneration on NG medium shows a 49.4% reduction in the number of shoots per calli, with an average of 4.47 shoots in PMI bombardments and an average of 2.21 shoots in the PMI and PMI RNAi co-bombardments. Taken together these results clearly show that there is a significant reduction in the regeneration of shoots from calli that are co-bombarded with the PMI and PMI RNAi constructs when compared with calli that have been bombarded with the PMI construct alone. In addition, after more stringent selection on full mannose medium RNAi silencing was demonstrated to have a significant effect on the transformation efficiency. This was reduced from 84% in the PMI control to 31.7% in calli that were co-bombarded with the PMI and PMI RNAi constructs. Furthermore, 39 of these 95 *pmi*-positive shoots tested negative for the RNAi construct (*nos* terminator sequence) and are therefore able regenerate in the same way as the PMI control, which if discounted would effectively reduce the transformation efficiency to 18.67%. The existence of shoots that contain both the PMI and PMI RNAi constructs suggests that although there is a significant silencing effect on PMI by the PMI RNAi construct it is not 100% complete. This may be due to the fact that the RNAi constructs have been damaged by the bombardment procedure and are therefore detectable but not fully functional. However, this is unlikely due to the number of transgenics identified as PCR positive. Another explanation, which was also discussed with the ECFP assay, is the possibility that the co-bombardment of the two constructs which are driven by constitutive promoters may result in the inability of the RNAi construct to completely silence the over-expressed *pmi* gene product. An alternative explanation would be that, because the effect of RNAi silencing is principally cell-specific, it is possible that mosaic shoots could regenerate which contain cell types that possess only the PMI construct, other cell types that possess only the PMI RNAi construct, as well as cell types that possess both PMI and PMI RNAi constructs. Therefore cell types that contained only the RNAi constructs would be detectable in regenerating shoots yet would be

ineffective in silencing the PMI transformed cells. This would account for the observation in Figure 49 that although some shoots from the co-bombarded calli were able to regenerate under mannose selection, the shoots that regenerated did not grow as vigorously as the PMI control. Overall, these results clearly show a significant reduction in the transformation efficiencies of calli that are co-bombarded with the PMI and PMI RNAi constructs when compared to calli that are bombarded with the PMI marker alone. However, the silencing does not appear to be complete, as demonstrated by the presence of both constructs in regenerating shoots.

6.4.2. RNAi Gene Testing

Following the satisfactory demonstration that gene silencing in wheat is occurring from the introduction of a hairpin construct, gene-specific RNAi constructs for each of the candidate clones were used to test their effect on SE in wheat calli. It is important to note here that, due to the requirement of callus induction on 2,4-D medium prior to bombardment, the early effects on the action of the targeted genes . may not be inhibited by the introduction of the RNAi silencing constructs. However, because SE takes place over a number of weeks with the continuous production of SEs it would be expected that the introduction of RNAi constructs to the embryogenic calli would still result in an observable phenotypic effect and reduction in regeneration and transformation efficiencies. The results of these experiments can be seen in Tables 23 and 24. Using the PMI control as a reference it can be seen that, except for PICKLE, there is a general reduction in these efficiencies for all of the genes tested. Generally speaking the results for the regeneration and transformation efficiencies appear to fall into two groups. Firstly, calli that were co-transformed with PMI and LRRK RNAi or OSE 731 RNAi constructs have similar transformation efficiencies to those seen for calli co-transformed with PMI and PMI RNAi constructs (28%, 37.3% and 32.5% respectively). Secondly, there are very similar transformation efficiencies observed for the remaining gene-specific RNAi constructs (ABREBF RNAi, SCR RNAi, STPK RNAi B.ZIP RNAi and PICKLE RNAi) which range from 51.8% to 53.75%. Furthermore these results are consistent with those observed for the regeneration efficiencies (with the exception of PICKLE) where LRRK and OSE 731 RNAi constructs have similar lower efficiencies of 2.37 and 2.52 respectively and the remaining gene-specific RNAi constructs ABREBF RNAi, SCR RNAi, STPK RNAi and B.ZIP RNAi range between 3.2 and 3.89.

6.4.2.1. LRRK RNAi Construct

The most dramatic of these reductions was observed in calli that were co-bombarded with the PMI and leucine rich receptor-like kinase (LRRK) RNAi constructs resulting in a 50% reduction in the number of shoots that regenerated on NG medium and a reduction in the transformation efficiency from 84% to 28% when compared with PMI. Of the 84 shoots that grew on mannose selection 46 of these (54.8%) were identified as containing the RNAi silencing construct. This suggests that although silencing of this target gene appears to significantly affect the regeneration and transformation efficiencies it does not completely disrupt the regeneration of adult structures from SEs that contain the RNAi construct. The photographs presented in figure 50 3A-3D demonstrate the significant reduction in the number of observable shoots. Furthermore, the shoots that regenerate can be seen to be much smaller and less vigorous than the PMI control with the calli appearing dry and non-embryogenic. Previous discussions about the possible role of this gene have suggested that it may be a wheat homologue of the SERK gene which is transiently expressed during SE in cells that possess the competence to form embryogenic structures. Due to the requirement of callus induction prior to bombardment, it may be that that this gene has already served its function in early developing SEs to acquire embryogenic competence. This would allow the competent embryos to further develop and regenerate even with the successful silencing of the LRRK target gene. However, the introduction of this construct clearly has an effect on SE and the regeneration and transformation efficiencies. Therefore, it is possible that the RNAi construct only has an effect on the formation and regeneration of secondary SEs which would account for the apparently conflicting observation of reduced transformation efficiencies along with the presence of the RNAi constructs. As discussed earlier in section 6.4.1.2. another possible explanation is that the effect of RNAi silencing is principally cell-specific, and therefore mosaic shoots could regenerate which contain cell types that possess PMI alone, and other cell types that possess both PMI and RNAi constructs.

6.4.2.2. OSE 731 RNAi Construct

Also showing a significant reduction in the regeneration and transformation efficiencies are calli that were co-bombarded with the PMI and OSE731 RNAi constructs. The average number of shoots per calli was reduced by 45.6% from 4.47 to 2.52 and the transformation efficiency reduced from 84% to 37.3%. Photographs

of the OSE 731 RNAi phenotype (Figure 50, 4a-4d) reveals a particularly interesting result whereby the smooth, glossy SE-like structures appear to be continually proliferating and forming without any adult root or shoot structures regenerating from them. The appearance of these calli is comparable to those that have been maintained for a number of weeks on high auxin media. However, these calli have been removed from 2,4-D media and placed on regeneration media for several weeks. This result is particularly exciting when coupled with the transformation efficiency results which show that the shoots that do regenerate only test positive for *pmi*. Not a single shoot from the 112 shoots that were sampled and identified as *pmi* positive were shown to contain the OSE 731 RNAi construct. This gene therefore appears to play a critical role in the transition of somatic embryos into differentiating adult structures with the silencing of this gene resulting in the disruption of somatic embryos to differentiate into adult cells. In addition, this result demonstrates that the transformation efficiency that appears to be 37.3% is in actual fact 0% when the construct is present and indicates that the transformation efficiencies given for each of the genes tested is in fact lower than the results suggest.

6.4.2.3. ABREBF RNAi Construct

The co-bombardment of calli with the PMI and ABREBF RNAi constructs results in a reduction in the average number of shoots that regenerate on NG medium by 23.9% from 4.47 to 3.4 and a reduction in the transformation efficiency from 84% to 53%. Of the total number of shoots that continued to grow on mannose selection 66 of these (41.25%) were identified as containing the RNAi silencing construct, effectively reducing the transformation efficiency to 22%. The photographs presented in figure 50 2A-2D demonstrate the reduction in the number of observable shoots when compared with the PMI control. The shoots that regenerate can be seen to be much less frequent and vigorous than the PMI control, however, the effect is also noticeably less marked than those observed for LRRK and OSE 731. These results suggest that although there is a reduction in the regeneration and transformation efficiencies, the role of this gene does not appear to be critical for the development of SE and the regeneration of shoot structures, in the same way that LRRK and OSE 731 are.

6.4.2.4. SCR RNAi Construct

The co-bombardment of the SCR RNAi construct shows a similar reduction in the regeneration and transformation efficiencies as observed with the ABREBF RNAi construct resulting in the reduction of the regeneration efficiency by 29.4% from 4.47 to 3.2 and results in a reduction of the transformation efficiency from 84% to 52%. Of the total number of shoots that continued to grow on mannose selection 69 of these (44.5%) were identified as containing the RNAi silencing construct, effectively reducing the transformation efficiency to 23%. These results suggest that although there is a reduction in the regeneration and transformation efficiencies the silencing of this gene does not appear to be critical in the development of SE and the regeneration of shoot structures. As discussed for the previous genes this may be as a result of the requirement of callus induction prior to bombardment and the possibility that this gene may have already served its functional role in SE allowing them to further develop and regenerate even with successful silencing of the SCR RNAi target gene. In addition, there is the possibility that mosaic shoots could regenerate which contain cell types that possess PMI alone and other cell types that possess both PMI and RNAi constructs.

6.4.2.5. STPK RNAi Construct

The co-bombardment of the PMI and STPK RNAi constructs also shows a similar reduction in the regeneration and transformation efficiencies as observed with the SCR RNAi and ABREBF RNAi constructs resulting in the reduction of the regeneration efficiency by 23% from 4.47 to 3.44 and a reduction in the transformation efficiency from 84% to 52%. Of the total number of shoots that continued to grow on mannose selection 99 of these (63.9%) were identified as containing the RNAi silencing construct, effectively reducing the transformation efficiency to 33%. These results also suggest that although there is a reduction in the regeneration and transformation efficiencies, the silencing of this gene does not appear to play a critical role in the development of SE and the regeneration of shoot structures. This may be due to the same reasons as provided earlier about the existence of chimeras and the requirement of callus induction prior to bombardment.

6.4.2.6. PICKLE RNAi Construct

The co-bombardment of calli with the PMI and PICKLE RNAi constructs results in a slight increase in the regeneration efficiencies from 4.47 to 4.89 shoots per callus although a reduction in the transformation efficiency from 84% to 51.8% is observed. From the total number of shoots that tested positive for *pmi*, 50 of these (60.24%) also tested positive for the presence of the RNAi silencing construct, effectively reducing the transformation efficiency to 25%. The effect on the regeneration efficiencies of calli co-bombarded with PMI and PICKLE RNAi constructs increases suggesting that the silencing of this target gene results in the enhancement of regeneration efficiencies. However, this is not echoed by the effect on transformation efficiency which suggests a reduction in the ability for somatic embryos to develop and regenerate shoot structures.

6.4.2.7. B.ZIP RNAi Construct

Finally, the bombardment of the B.ZIP RNAi construct shows a slight reduction in the regeneration efficiencies from 4.47 to 3.89 and a reduction in the transformation efficiency from 84% to 53.75%. Of the total number of shoots that continued to grow on mannose selection 61 of these (70.9%) were identified as containing the RNAi silencing construct, effectively reducing the transformation efficiency to 30.5%. These results suggest that the silencing of this gene does not appear to play a critical role in the development of SE and the regeneration of shoot structures. This may be due to the same reasons as provided earlier about the existence of chimeras and the requirement of callus induction prior to bombardment.

6.4.2.8. Non-Specific Gene Targeting Effects

One of the problems facing RNAi experiments is the extent by which the knockdown of gene expression and the resulting phenotype is due to sequence specific gene silencing. It has been found that transfection of long dsRNA molecules (>30 nt) into mammalian cells causes non-specific suppression of gene expression. This suppression has been attributed to an antiviral response, which takes place through one of two pathways. In one pathway, long dsRNAs activate a protein kinase, PKR. Activated PKR, in turn phoshorylates and inactivates the translation initiation factor, eIF2a, leading to repression of translation (Manche *et al.*, 1992). In the other pathway, long dsRNAs activate RNase L, which leads to non-specific RNA

degradation (Minks *et al.*, 1979). Interestingly, dsRNAs less than 30 nt in length do not activate the PKR kinase pathway. This observation prompted researchers to test whether introduction of siRNAs could induce gene-specific silencing in mammalian cells (Elbashir *et al.*, 2001). Indeed, siRNAs introduced by transient transfection were found to effectively induce RNAi in mammalian cultured cells in a sequence-specific manner. However, sequence non-specific effects have been observed with siRNAs at high concentrations which can lead to false positive conclusions regarding the role of the target gene in functional assays. To control for such effects an RNAi sequence directed against a gene target not expressed in host cells, for example GFP, should be used as a negative control. In addition to non-specific effects, sequence specific, off-target signatures of siRNAs have also been observed in gene profiling experiments (Jackson *et al.*, 2003). Using multiple siRNAs to target the same mRNA transcript these researchers found siRNA specific, rather than target specific, expression profiles. The only effective measure against this is good bioinformatic analysis of the target sequence to identify RNA segments that are as unique as possible and the use of multiple RNAi constructs that target different regions of a single gene which would increase the confidence with which the phenotype can be ascribed to silencing of the target gene.

It is difficult to say whether the results generated here have been affected by non-specific or off-target effects without undertaking further experiments that incorporate the use of appropriate controls mentioned above. For example, the RNAi gene testing experiments could include a negative RNAi control containing sequence for a gene which is not present in wheat calli cells, such as the ECFP RNAi construct. This would allow one to ascertain whether or not there was any non-specific silencing effect on transformation and regeneration efficiencies by the introduction of the dsRNA constructs. In addition the use of multiple RNAi constructs to target the OSE 731 gene, for example, would increase the confidence with which the phenotype can be ascribed to silencing of the target gene and not as a result of off-target effects. It would be useful to consider these controls for any further work that may be undertaken. Nevertheless, the co-bombardment of several gene-specific RNAi constructs in parallel helps provide important relative controls with each other. For example, it can be interpreted with confidence that the complete silencing effect as demonstrated by OSE 731 is due to silencing of the target sequence or any possible

off-target sequences and not as a result of general non-specific effects and potentiation of the RISC pathway. If it were due to non-specific effects then transformation with each of the RNAi constructs would have resulted in their complete silencing which has been demonstrated by these results to not be the case.

Chapter Seven

Discussion

7 Chapter Seven – Discussion

Taking a microarray based approach, this projects aim was to utilise this powerful gene expression technology to identify and characterise genes that are regulated during somatic embryogenesis (SE) in wheat. Although many studies have already been conducted to see how cells respond to environmental changes and stress through differential expression patterns of mRNA this is probably the first of its kind to look at SE. Microarrays can be thought of as simply a more powerful substitute for more conventional methods for evaluating mRNA abundance. Early experiments on a relatively small set of genes thought to be important to a particular process were used. However, a key advantage of arrays is that it is not necessary to try and decide what the important genes or mechanisms are in advance and as a result a more complete and less biased view of cellular response can be obtained. Such discovery programmes can often be described as being question driven rather than hypothesis driven in a conventional sense, as microarrays allow researchers to ask big questions and the results often generate large quantities of data. Examining individual arrays and identifying which changes are significant provides one of the main difficulties associated with the application of this technology and this study. Furthermore, comparisons of a number of different arrays make this even more difficult and increases the possibility for correlations to be missed.

7.1. Bobwhite and Savannah Cultivars Provide a Testable System.

In order for the aims of this study to be met it was first necessary to establish and compare two suitable experimental systems. This was achieved through the demonstration (in Chapter Two) that Bobwhite (BW) has a far superior regeneration efficiency (16.82 ± 3.01 shoots per calli) when compared to Savannah (SAV) (2.97 ± 0.78 shoots per calli). This established these two cultivars as suitable experimental systems for evaluating the differential expression of genes associated with SE. It also allowed for the design of an experimental hybridisation strategy that divided the processes of SE into several timepoints to test; 1) the expression of genes involved in the de-differentiation of somatic cells towards an embryogenic state (timepoint 1); 2)

the initiation and development of SEs (timepoint 2 & 3); 3) the response of SEs to the removal of 2,4-D (timepoint 4).

7.2. GEM Experiments Identified Hundreds of Differentially Expressed Genes Associated with SE in Wheat.

Material was collected at these different timepoints from which high quality mRNA was extracted and used to synthesise the cDNA fluorescent probes for each microarray hybridisation. To simplify subsequent analysis, the hybridisations were carried out in both the embryogenic and non-embryogenic samples using untreated samples as a standard reference. This allowed for the subtraction of genes that were differentially expressed in the non-embryogenic cultivar from the set of genes that were differentially expressed in the embryogenic cultivar, which resulted in the identification of 700 differentially expressed genes. Seventy one of these were identified as being differentially expressed during timepoint 1, 145 genes were identified from timepoint 2, 188 identified from timepoint 3, 137 identified from timepoint 4 and an additional 160 identified from the BW timepoint 3 versus timepoint 4 and SAV timepoint 4 versus BW timepoint 4 hybridisations. Because the incidence of false positives is made significantly low through internal controls, it is not necessary to independently confirm every change for the results to be valid and trustworthy. Expected results act effectively as internal controls that provide a certain amount of validation and assurance to any new or unexpected results that are obtained from a microarray hybridisation experiment (Lockhart and Winzeler, 2000). Therefore the presence of a number of expected genes such as auxin response genes, stress response genes, heat shock proteins, cell cycle regulated, chitinases etc. provide a degree of assurance over the validity of the microarray data. In addition, because the cDNA arrays are made on partial sequence information it is possible to include completely uncharacterised genes (Incyte ESTs). The identification of these uncharacterised genes in the cDNA microarray experiments allows for a tentative functional assignment. This may not be much more than a low resolution description indicating that these genes are associated with the process of SE but this allows further attention to focus on a smaller subset of genes which have not previously been considered as obvious candidates in the absence of global gene expression experiments. Furthermore, this also highlights the importance of functional

annotation and curation of existing sequence and functional databases. Results covering tens of thousands of genes or expressed sequence tags (ESTs) will be only partly interpretable given the functional and biological information available at the time.

Probably the most relevant studies to this project of microarray experiments that have been performed to date are by (Menges *et al.*, 2002) and (Raghavan *et al.*, 2004) who both used Affymetrix gene arrays to study global genome responses in *Arabidopsis*. Menges et al (2002), using a recently developed cell synchrony system for *Arabidopsis* carried out an analysis of gene expression that resulted in the identification of 463 candidate genes that showed a cell cycle regulated pattern of expression. Among these genes were those both known and suspected to be cell cycle regulated in plants as well as those involved in a number of other cellular processes such as hormone response, signal transduction, transcriptional control and metabolic regulation. Arabidopsis cell suspension cultures were synchronised by reversibly blocking them in the late G1/early S-phase with aphidicolin, a DNA polymerase inhibitor. The drug was then removed and samples taken every two hours over a nineteen hour period. RNA was extracted from each of the samples and analysed for genes expression. Interestingly a significant number of genes that were identified in their study Table 26 were also identified in this study:

S-Phase	G2 -Phase	M-Phase	G1-Phase
putative protein	putative endochitinase	WD-repeat protein-like protein	MADS
putative trypsin inhibitor	putative protein kinase	cytochrome P450 - like protein	
putative endochitinase		hypothetical protein	
Heat shock transcription factor		putative protein	
Histone H2A		nodulins	
putative cytochrome P450		putative auxin regulated protein	
phosphoenolpyruvate carboxylase		ferritin 1 precursor	
O-methyltransferase		zinc finger protein	
putative phospholipase		putative kinesin heavy chain	
DNA (cytosine-5) Methyltransferase		putative expansin	
putative glutathione peroxidase		DNA binding protein	
EREBP2		putative ethylene response element BP	
HMG1		putative NADH ubiquinone oxidoreductase	
EREBP1		MYB transcription factor	
Heat shock protein		putative protein phosphatase 2C	
putative DNA binding protein		leucine -zipper containing protein	
PCNA			
putative WD-repeat protein			
DnaJ homologue			
17.6 kDa HSP			
putative protein phosphatase 2C			
putative SCARECROW gene regulator			
putative NADPH oxidase			
putative SET domain transcriptional regulator			
putative high mobility group protein			
putative glucosyltransferase			
receptor protein kinase - like protein			
NADPH oxidoreductase			

Table 26: cell cycle regulated gene expression in *Arabidopsis*.

A recent study by (Raghavan *et al.*, 2004) attempted to evaluate gene expression patterns of *Arabidopsis* in response to treatment with 2,4-D. Fourteen day old *Arabidopsis* plantlets were treated with 1mM 2,4-D for one hour and the RNA hybridised to an Affymetrix *Arabidopsis* microarray to access changes in gene expression levels. A total of 233 genes were identified as being differentially expressed belonging to several functional categories involving transcription, metabolism, cellular communication and signal transduction, subcellular localisation, transport facilitation, protein fate and the regulation of or interaction with the cellular environment. The data obtained from these experiments indicated that 2,4-D not only modulates the expression of auxin, ethylene and abscisic acid pathways but also regulate a wide variety of other cellular functions (Raghavan *et al.*, 2004). This links with the results in this study which confirm the modulation of other phytohormones pathways as a result of 2,4-D application.

7.3. In Situ Hybridisation Experiments Confirm the Expression of a Large Number of Candidate Genes in SE.

In order to reduce this large number of genes to a more manageable quantity for further characterisation additional selection criteria were applied by selection genes with a bias towards a) their high differential expression value, b) their putative function and c) their gene expression profiles throughout each of the 4 timepoints (timepoint 1 – 4). This resulted in the selection of a diverse range of shortlisted clones of which 57 were successfully used for the synthesis of digoxigenin labelled RNA probes for the *in situ* hybridisation experiments. This part of the study was carried out to achieve 3 objectives. Firstly, to confirm the expression of the selected set of genes that was identified as being differentially expressed from the microarray analysis. Secondly, to provide further characterisation of the spatial and temporal expression of these genes in embryogenic tissues and thirdly to allow for the further selection of a sub section of the genes tested based on their confirmed differential expression and spatial expression patterns. A large number of the shortlisted genes expression were confirmed by the ISH screen and 10 genes were selected to make RNAi gene silencing constructs to test their biological function in SE.

7.4. Bioinformatics and Clustal Analysis Confirm and Reveal Functional Annotation of Shortlisted Genes

The wheat EST sequences were used to perform recent Blast searches of both the rice and non-redundant databases. From these searches new annotation was inferred for several wheat sequences including, the hypothetical protein sequence which aligns with a PAXNEB protein, an RNA polymerase II elongator protein subunit. Several of the Incyte ESTs (702001685, 701700561 and 701965472) which align with a putative CGI-94 protein, a nuclear transport factor 2 (NTF-2) protein and a putative transposase respectively, and a putative protein that aligns with a WD-40 repeat protein. The trees presented confirm some of these alignments, for example the hypothetical protein aligning with an *A.thaliana ELONGATA1* gene which encodes for a subunit of RNA polymerase II. An Incyte EST which aligns with an *A.thaliana* and *O.sativa* nuclear transport factor2 (NTF2) sequence. The majority of sequences are confirmed by alignment with genes of the same or equivalent function, including the RNAi shortlisted sequences. Of particular interest was the tree generated from the alignments of the wheat sequence previously identified as a Serine-Threonine protein kinase (STPK)-like protein (Fig. 21). The wheat sequence was shown to share common ancestry with a number of male germ cell-associated kinase (MAK) gene sequences in both the plant and animal kingdoms. As previously discussed, members of the Mak subgroup are Cdc2-related serine/threonine protein kinases and the expression of the *MAK* gene has been demonstrated in sugar beet to be correlated with cell division activity or cell differentiation (Fowler *et al.*, 2000). Furthermore, the mammalian homolog of this gene is highly expressed in testicular germ cells during meiosis (Matsushime *et al.*, 1990). This is particularly interesting because not only does this demonstrate a high degree of evolutionary conservation for this kinase across the plant and animal kingdoms but also provides an excellent candidate as a gene involved with the onset of SE. Furthermore, the role of this gene in the animal kingdom suggests that there may even be cells undergoing meiosis during SE which has been suggested previously in carrot cell lines (Nutti-Ronchi *et al.*, 1992). This highlights the need for further studies to investigate the role of this gene in SE and the role of meiosis in SE in the field in general.

7.5. RNAi Silencing Provides Functional Evidence for Genes Involved in SE in Wheat.

An important prerequisite to RNAi gene silencing experiments was to establish the efficacy of this technique in wheat. This was demonstrated by designing a transient assay strategy to test the effect of co-bombarding a selectable or visual marker together with a gene specific RNAi silencing construct. The results from both of these co-transformation experiments clearly demonstrate a gene silencing effect therefore allowing for the proceeding of gene testing by the RNAi constructs with confidence, on each of the in situ hybridisation short listed genes. The effect on regeneration of calli co-bombarded with PMI/gene-specific RNAi constructs showed a reduction in the regeneration and transformation efficiencies of the seven successfully transformed RNAi constructs. Most noticeably were the transformation efficiencies for the OSE731, LRRK and ABREBF constructs that were reduced to 37.3%, 28% and 53% respectively (or 0%, 15.33% and 22% if PMI only shoots are discounted) from 84% achieved by the PMI construct alone. Furthermore, each one of these genes resulted in an observable phenotype that was replicated in each experiment, providing strong evidence that these genes play an important role in SE.

In addition, it has been demonstrated that there are essentially two critical steps, in culture, that contribute to the phenomenon of SE in wheat. Firstly, it has been shown that the treatment of somatic tissue with the auxin 2,4-D results in substantial genetic reprogramming, cellular de-differentiation and embryogenic callus formation. In particular the *in situ* hybridisation pattern of the ABREBF gene is clearly induced after only several hours exposure to 2,4-D and has been demonstrated by the RNAi silencing to be required for proper initiation of somatic embryo structures. Secondly, it has been shown by the ISH results that the OSE 731 gene appears to be significantly up-regulated on removal of embryogenic calli from high auxin media. This correlates with the observation that on the removal of 2,4-D the embryogenic tissue undergoes further genetic events that result in the re-differentiation and regeneration of adult structures. Furthermore, it has been demonstrated that the silencing of this gene results in the arrest of somatic embryos and the inability of embryogenic calli to regenerate even after the removal of 2,4-D. Taken together these results strongly implicate this gene as playing an essential role in the differentiation of SEs and

embryogenic tissue into adult structures by disrupting the ability of the SEs and embryogenic structures to respond to the resulting changes in 2,4-D concentration.

This study has significantly improved our understanding of SE. It has provided for the first time global gene expression profiles of wheat calli undergoing SE and identified hundreds of genes as being associated with SE. Some of these genes have been characterised further by ISH and RNAi gene silencing studies resulting in a number of genes being identified that appear to play a role in the onset of SE. In particular, there is a significant finding that a novel wheat gene with homology to the rice embryo specific protein OSE 731 appears to be involved in the pathway controlling cell proliferation and morphogenesis from somatic embryos towards adult shoots. This exciting result, which should be publishable with a little additional work, should also warrant the possibility of further funding so that additional investigations into the functional and mechanistic role in SE of this gene and others identified in this study can be undertaken.

7.6. Future Directions for the Study of SE in Wheat.

The experiments described here provide evidence that a microarray based, functional genomics approach to identifying genes associated with SE is a worthwhile and successful strategy. It has resulted in the confirmation of a large number of previously identified genes associated with SE. More importantly, there are a large number of genes identified that have been previously unassociated with SE which contribute to the novel findings of this study and highlights the need for additional future studies in this field in general. There are a number of ways in which the work undertaken during this study could be taken further to enhance our understanding of SE. Firstly, with regards to improving the expression data further microarray studies could be undertaken using replicates to control for any effects of biological variation and increase confidence in the resulting expression data. In addition, quantitative PCR experiments could be carried out to help confirm the differential expression patterns and values of those genes that were identified from the microarray experiments. Evidence from such experiments would strongly support the results identified in this study and provide a much stronger link between the microarray and ISH experiments. In addition, it would be useful to undertake further ISH experiments, for the genes that were confirmed by the ISH screen, which compared

the ISH signal between Bobwhite and Savannah samples. This would allow one to conclusively confirm whether or not the expression of the genes tested is specific to Bobwhite and the embryogenic response. Furthermore, with adequate resources it would be of particular interest to undertake a high throughput ISH screen of all 700 of the genes that were identified as being differentially expressed in embryogenic calli. This would undoubtedly reveal more candidates with interesting ISH signalling patterns that could also be characterised further by ISH experiments described above and functional experiments similar to those undertaken in this study.

There are also additional transformation experiments that could be carried out further to clarify the results presented in this study. For example, in order to circumnavigate problems associated with having to bombard tissue that has been exposed to high auxin media for several days and therefore already undergoing SE, you could introduce RNAi silencing and overexpression constructs with inducible promoters. Caryopsi could then be collected from transgenic plants and placed under standard tissue culture conditions and the promoter induced from the outset to assess the effect of overexpression or silencing of particular genes on the early response to auxin exposure. More crucially, additional RNAi silencing experiments could be carried out which incorporate negative controls which can be used to assess whether the introduction of an RNAi construct has non-specific effects on the regeneration and transformation efficiencies. Future RNAi gene testing experiments could include a negative RNAi control containing sequence for a gene which is not present in wheat calli cells, such as the ECFP RNAi construct. This would allow one to ascertain whether or not there was any non-specific silencing effect on transformation and regeneration efficiencies by the introduction of the dsRNA constructs. If the introduction of an ECFP RNAi construct were to give similar phenotypic results and regeneration and transformation efficiencies as observed in the PMI control then it could be concluded that the silencing effects are by the gene-specific RNAi constructs and not the action of the dsRNA construct *per se*. In addition, the use of multiple RNAi constructs to target the gene sequence under investigation would increase the confidence with which the phenotype can be ascribed to silencing of the target gene and not as a result of off-target effects. For example, if multiple RNAi constructs used against the OSE 731 gene replicated the results of this study then it would provide conclusive evidence for the role of this gene in SE. Finally, to get a more

complete understanding of SE it would be of great interest to clone the full length genes of the shortlisted candidates. This would allow for the analysis of overexpression studies on the regeneration and transformation efficiencies and may produce phenotypes that would provide further data about the functional and mechanistic roles of the genes. In addition, it would allow for more complete sequence alignment searches which should in turn reveal further information about the genes functional role and evolutionary conservation amongst other species. Overall this study has identified and characterised several key genes involved in SE. What seems apparent from this and others studies is that it is unlikely there is any single event that is responsible for the process of SE. Rather there appears to be a complex interaction between the exogenous application of auxin, the endogenous regulation of phytohormone responses and a wide variety of cellular responses that contribute to the process of SE, which is only beginning to be properly understood. This study goes some way to improve our understanding of these interactions and responses but also highlights the need for further work to be undertaken in the study of SE.

Bibliography

Bibliography

Abel, S., P. W. Oeller and A. Theologis (1994). "Early auxin-induced genes encode short-lived nuclear proteins." *Proceedings of the National Academy of Sciences of the United States of America* **91**(1): 326-30.

Alexandrova, K. S. and B. V. Conger (2002). "Isolation of two somatic embryogenesis-related genes from orchardgrass (*Dactylis glomerata*)." *Plant Science* **162**: 301–307.

Ambros, V. (2004). "The functions of animal microRNAs." *Nature* **431**: 350 – 355.

Aravind, L. and D. Landsman (1998). "AT-hook motifs identified in a wide variety of DNA-binding proteins." *Nucleic Acids Research* **26**(19): 4413-4421.

Assaad, F. F., Y. Huet, U. Mayer and G. Jürgens (2001). "The cytokinesis gene *KEULE* encodes a *Sec1* protein that binds the syntaxin *KNOLLE*." *Journal of Cell Biology* **152**: 531 - 544.

Bartel, D. P. (2004). "MicroRNAs: genomics, biogenesis, mechanism and function." *Cell* **116**: 281-297.

Barton, M. K. and R. S. Poethig (1993). "Formation of the shoot apical meristem in *Arabidopsis thaliana*: an analysis of development in the wild type and in the *shoot meristemless* mutant." *Development* **119**: 823-831.

Baud, S., Y. Bellec, M. Miquel, C. Bellini, M. Caboche, L. Lepiniec, J. D. Faure and C. Rochat (2004). "*gurke* and *pasticcino3* mutants affected in embryo development are impaired in acetyl-CoA carboxylase." *EMBO Reports* **5**: 515–520.

Baulcombe, D. (2004). "RNA silencing in plants." *Nature* **431**: 356-363.

Benfey, P. N., P. J. Linstead, K. Roberts, J. W. Schiefelbein, M. T. Hauser and R. A. Aeschbacher (1993). "Root development in *Arabidopsis*: four mutants with dramatically altered root morphogenesis." *Development* **119**(1): 57-70.

Benkova, E., M. Michniewicz, M. Sauer, T. Teichmann, D. Seifertova, G. Jürgens and J. Friml (2003). "Local, efflux-dependent auxin gradients as a common module for plant organ formation." *Cell* **115**(5): 591-602.

Bennett, M., A. Marchant , H. G. Green, S. T. May, S. P. Ward, P. A. Milner, A. R. Walker, B. Schulz and K. Feldmann (1996). "Arabidopsis AUX1 gene: a permease-like regulator of root gravitropism." *Science* **273**: 948-950.

Bennett, V. and A. J. Baines (2001). "Spectrin and ankyrin-based pathways: Metazoan inventions for integrating cells Into tissues." *Physiological Reviews* **81**(3): 1353-1392.

Berleth, T. and G. Jürgens (1993). "The role of the *monopteros* gene in organising the basal body region of the *Arabidopsis* embryo." *Development* **118**: 575-587.

Bogre, L., Z. Olah and D. Dudits (1988). "Calcium dependent protein kinase from alfalfa (*Medicago varia*): partial purification and autophosphorylation." *Plant Science* **58**: 135-144.

Bond, U. and M. J. Schlesinger (1987). "Heat-shock proteins and development." *Advanced Genetics* **24**(1).

Botella, J. R., J. M. Arteca, M. Somodevilla and R. N. Arteca (1996). "Calcium-dependent protein kinase gene expression in response to physical and chemical stimuli in mungbean (*Vigna radiata*)." *Plant Molecular Biology Reporter* **30**(6): 1129-37.

Boutilier, K., R. Offringa, V. K. Sharma, H. Kieft, T. Ouellet, L. Zhang, J. Hattori, C. M. Liu, A. A. van Lammeren, B. L. Miki, J. B. Custers and M. M. van Lookeren Campagne (2002). "Ectopic expression of *BABY BOOM* triggers a conversion from vegetative to embryonic growth." *Plant Cell* **14**(8): 1737-49.

Brzobohaty, B., I. Moore, P. Kristoffersen, L. Bako, N. Campos, J. Schell and K. Palme (1993). "Release of active cytokinin by a beta-glucosidase localized to the maize root meristem." *Science* **262**(5136): 1051-4.

Campbell, J. L., N. Y. Klueva, H.-g. Zheng, J. Nieto-Sotelo, T. Ho, H. D. Nguyen and T. Henry (2001). "Cloning of new members of heat shock protein HSP101 gene family in wheat (*Triticum aestivum* (L.) Moench) inducible by heat, dehydration, and *ABA1*." *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1517**(2): 270-277.

Carrera, P., Y. M. Moshkin, S. Gronke, H. H. W. Sillje, E. A. Nigg, H. Jackle and F. Karch (2003). "*Tousled*-like kinase functions with the chromatin assembly pathway regulating nuclear divisions." *Genes and Development* **17**(20): 2578-2590.

Casamitjana-Martinez, E., H. F. Hofius, J. Xu, C. M. Liu, R. Heidstra and B. Scheres (2003). "Root-specific CLE19 ovexpression and the sol1/2 supressors implicate a CLV-like pathway in the control of Arabidopsis root meristem maintenance." *Current Biology*.

Chang, S., J. Pryear and J. Cairney (1993). "A simple and efficient method for isolating RNA from pine trees." *Plant Molecular Biology Reporter* **11**: 113-117.

Chen, J., S. Zhou, Q. Wang, X. Chen, T. Pan and H. Liu (2000). "*Crk1*, a novel Cdc2-related protein kinase, Is required for hyphal development and virulence in *candida albicans*." *Molecular and Cellular Biology* **20**(23): 8696-8708.

Chen, P., L. Fang, J. Lin, H. Tsay, H. Wu and L. Chen (1997). "Isolation of cDNA clones for genes that are specifically expressed in the rice embryo." *Botanical Bulletin of Academia Sinica* **38**: 13-20.

Chen, X. M. (2004). "a microRNA as a translation repressor of *APETALA* in *Arabidopsis* flower development." *Science* **303**: 2022-2025.

Choi, H., J. Hong, J. Ha, J. Kang and S. Y. Kim (2000). "ABFs, a family of ABA-responsive element binding factors." *Journal of Biological Chemistry* **3**(275): 1723-30.

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

Churin, Y., S. Schilling and T. Borner (1999). "A gene family encoding glutathione peroxidase homologues in *Hordeum vulgare* (barley)." *FEBS Letters* **459**(1): 33-38.

Clark, S. E., M. P. Running and E. M. Meyerowitz (1993). "*CLAVATA1*, a regulator of meristem and flower development in *Arabidopsis*." *Development* **119**: 397-418.

Clark, S. E., M. P. Running and E. M. Meyerowitz (1995). "*CLAVATA3* is a specific regulator of shoot and floral meristem development affecting the same processes as *CLAVATA1*." *Development* **121**: 2057-67.

Clark, S. E., R. W. Williams and E. M. Meyerowitz (1997). "The *CLAVATA1* gene encodes a putatuve receptor kinase that controls shoot and floral mersitem size in *Arabidopsis*." *Cell* **89**: 575-85.

Cosgrove, D. J., P. Bedinger and D. M. Durachko (1997). "Group I allergens of grass pollen as cell wall-loosening agents." *Proceedings of the National Academy of Sciences of the United States of America* **94**(12): 6559-6564.

Davies, R. T., D. H. Goetz, J. Lasswell, M. N. Anderson and B. Bartel (1999). "*IAR3* encodes an auxin conjugate hydrolase from *Arabidopsis*." *Plant Cell* **11**(3): 365-376.

De Jong, A. J., J. Cordewener, F. Lo Schiavo, M. Terzi, J. Vandekerckhove, A. Van Kammen and S. C. De Vries (1992). "A carrot somatic embryo mutant is rescued by chitinase." *Plant Cell* **4**(4): 425-33.

Dharmasiri, N., S. Dharmasiri and M. Estelle (2005). "The F-box protein TIR1 is an auxin receptor." *Nature* **435**: 441-445.

Di Laurenzio, L., J. Wysockadiller, J. E. Malamy, L. Pysh, Y. Helariutta, G. Freshour, M. G. Hahn, K. A. Feldmann and P. N. Benfey (1996). "The *SCARECROW* gene regulates an asymmetric cell division that is essential for generating the radial organization of the *Arabidopsis* root." *Cell* **86**: 423-433.

Dodeman, V. L., G. Ducreux and M. Kreis (1997). "Zygotic embryogenesis versus somatic embryogenesis." *Journal of Experimental Botony* **48**(31): 1493-1509.

Drea, S., B. Crawford, D. Leader, L. Dolan, P. Shaw and J. Doonan (unpublished). "A method for high-throughput in situ analysis of mRNA in plants."

Elbashir, S. M., W. Lendeckel and T. Tuschl (2001). "RNA interference is mediated by 21- and 22-nucleotide RNAs " *Genes Dev* **15**: 188-200.

Fairman, M. E., P. A. Maroney, W. Wang, H. A. Bowers, P. Gollnick, T. W. Nilsen and E. Jankowsky (2004). "Protein displacement by DExH/D "RNA Helicases" without duplex unwinding." *Science* **304**(5671): 730-734.

Fire, A., S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver and C. C. Mello (1998). "Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*." *Nature* **391**(6669): 806-811.

Fleming, A. J., S. McQueen-Mason, T. Mandel and C. Kuhlemeier (1997). "Induction of leaf primordia by the cell wall protein expansin." *Science* **276**: 1415-1418.

Fletcher, J. C., U. Brand, M. P. Running, R. Simons and E. M. Meyerowitz (1999). "Signalling of cell fate decisions by *CLAVATA3* in *Arabidopsis* shoot meristems." *Science* **283**: 1911-14.

Floyd, S. K. and J. L. Bowman (2004). "Gene regulation: ancient microRNA target sequences in plant." *Nature* **428**: 485-486.

Fowler, M. R., A. I. Atanassova, M. C. Elliott, N. W. Scott and A. Slater (2000). "Characterization of a Mak subgroup Cdc2-like protein kinase from sugar beet (Beta vulgaris L.)." *Journal of Experimental Botany* **51**(353): 2119-2124.

Friml, J., E. Benková, I. Blilou, J. Wisniewska, T. Hamann, K. Ljung, S. Woody, G. Sandberg, B. Scheres, G. Jürgens and K. Palme (2002). "AtPIN4 Mediates Sink-Driven Auxin Gradients and Root Patterning in Arabidopsis." *Cell* **108**: 661-673.

Friml, J., A. Vieten, M. Sauer, D. Weijers, H. Schwarz, T. Hamann, R. Offringa and G. Jürgens (2003). "Efflux-dependent auxin gradients establish the apical-basal axis of *Arabidopsis*." *Nature* **426**(6963): 147-53.

Frommer, W. B., S. Hummel and J. W. Reismeier (1993). "Expression Cloning in Yeast of a cDNA Encoding a Broad Specificity Amino Acid Permease from Arabidopsis thaliana." *Proceedings of the National Academy of Sciences of the United States of America* **90**: 5944-5948.

Fujimura, T. and A. Komamine (1980). "Mode of action of 2,4-D and zeatin on somatic embryogenesis in a carrot cell suspension culture." *Z. Pflanzenphysiologie* **99**: 1-8.

Gall, J. G. and M. L. Pardue (1969). "Formation and detection of RNA-DNA hybrid molecules in cytological preparations." *Proceedings of the National Academy of Sciences of the United States of America* **63**: 378-383.

Gälweiler, L., C. Guan, A. Müller, E. Wisman, K. Mendgen, A. Yephremov and K. Palme (1998). "Regulation of Polar Auxin Transport by AtPIN1 in Arabidopsis Vascular Tissue." *Science* **282**: 226-2230.

Gehring, C. A., H. R. Irving and R. W. Parish (1990). "Effects of Auxin and Abscisic Acid on Cytosolic Calcium and pH in Plant Cells." *Proceedings of the National Academy of Sciences* **87**: 9645-9649.

Geldner, N., N. Anders, H. Wolters, J. Keicher, W. Kornberger, P. Muller, A. Delbarre, T. Ueda, A. Nakano and G. Jürgens (2003). "The Arabidopsis GNOM ARF-GEF Mediates Endosomal Recycling, Auxin Transport, and Auxin-Dependent Plant Growth." *Cell* **112**: 219-230.

Geldner, N., J. Friml, S. York-Dieter, G. Jürgens and K. Palme (2001). "Auxin transport inhibitors block PIN1 cycling and vesicle trafficking." *Nature* **413**: 425-428.

Grasser, K. D., R. Grimm and C. Ritt (1996). "Maize Chromosomal HMGc. Two closely related structure-specific DNA-binding proteins specify a second type of plant high mobility group box protein." *Journal of Biological Chemistry* **271**(51): 32900-32906.

Gray, W. M., J. C. del Pozo, L. Walker, L. Hobbie, E. Risseeuw, T. Banks, W. L. Crosby, M. Yang, H. Ma and M. Estelle (1999). "Identification of an SCF ubiquitin-ligase complex required for auxin response in *Arabidopsis thaliana*." *Genes and Development* **13**(13): 1678-91.

Gupta, R., Y. Huang, J. J. Kieber and S. Luan (1998). "Identification of a dual- specificity protein phosphatase that inactivates MAP kinase in *Arabidopsis*." *Plant Journal* **16**: 581-590.

Hadfi, K., V. Speth and G. Neuhaus (1998). "Auxin-induced developmental patterns in *Brassica juncea* embryos." *Development* **125**: 879-887.

Haeker, A., R. Gross-Hardt, B. Geiges, A. Sarkar, H. Breuninger, M. Herrmann and T. Laux (2004). "Expression dynamics of *WOX* genes mark cell fate decisions during early embryonic patterning in *Arabidopsis thaliana*." *Development* **131**: 657-668.

Halperin, W. and D. F. Wetherell (1964). "Adventive embryony in tissue cultures of the wild carrot, *Daucus carota*." *American Journal of Botany* **51**: 274-83.

Hamilton, A. J. and D. Baulcombe (1999). "A species of small antisense RNA in post-transcriptional gene silencing in plants." *Science* **286**: 950-952.

Hammond, S. M., E. Bernstein, D. Beach and G. J. Hannon (2000). "An RNA-directed nuclease mediates post transcriptional gene silcening in *Drosophila* cell." *Nature* **404**: 293-296.

Hammond, S. M., A. A. Caudy and G. J. Hannon (2001). "Post-transcriptional gene silencing by double-stranded RNA." *Nature Reviews Genetics* **2**: 110 -119.

Hanstein, J. L. (1870). Die Entwicklung des Keimes der Monokotylen und Dikotylen. A. Marcus. Bonn, Germany.

Hardtke, C. S. and T. Berleth (1998). "The *Arabidopsis* gene *MONOPTEROS* encodes a transcription factor mediating embryo axis formation and vascular development." *EMBO Journal* **17**(5): 1405-11.

Hata, S. (1991). "cDNA cloning of a novel cdc2+/DCD28-related protein kinase from rice." *FEBS Letters* **279**: 149.

Hauser, B. A., J. Q. He, S. Park and C. S. Gasser (2000). "*TSO1* is a novel protein that modulates cytokinesis and cell expansion in *Arabidopsis*." *Development* **127**: 2219-2226.

He, J. X., S. Fujioka, T. C. Li, S. G. Kang, H. Seto, S. Takatsuto, S. Yoshida and J. C. Jang (2003). "Sterols regulate development and gene expression in *Arabidopsis*." *Plant Physiology* **131**(3): 1258-69.

Hecht, V., J. P. Vielle-Calzada, M. V. Hartog, E. D. L. Schmidt, K. Boutilier, U. Grossniklaus and S. C. de Vries (2001). "The *Arabidopsis SOMATIC EMBRYOGENESIS RECEPTOR KINASE 1* gene is expressed in developing ovules and embryos and enhances embryogenic competence in culture." *Plant Physiology* **127**:: 803-816.

Heck, G. R. and T. H. HO (1996). "Gibberellin-repressible gene expression in the barley aleurone layer." *Plant Molecular Biology* **30**(3): 611-23.

Heuer, S., S. Hansen, J. Bantin, R. Brettschneider, E. Kranz, H. Lorz and T. Dresselhaus (2001). "The maize MADS box gene *ZmMADS3* affects node number and spikelet development and Is co-expressed with *ZmMADS1* during flower development, in egg cells, and early embryogenesis." *Plant Physiology* **127**(1): 33-45.

Hirt, H., A. Pay, J. Gyorgyey, L. Bako, K. Nemeth, L. Bogre, R. J. Schweyen, E. Heberle-Bors and D. Dudits (1991). "Complementation of a yeast cell cycle mutant by an alfalfa cDNA encoding a protein kinase homologous to p34cdc2." *Proceedings of the National Academy of Sciences of the United States of America* **88**: 1636-1640.

Houbaviy, H. B., M. F. Murray and P. A. Sharp (2003). "Embryonic stem cell-specific MicroRNAs." *Developmental Cell* **5**(2): 351-8.

Hubbard, C., D. Singleton, M. Rauch, S. Jayasinghe, D. Cafiso and D. Castle (2000). "The secretory carrier membrane protein family: Structure and membrane topology." *Molecular Biology of the Cell* **11**(9): 2933-2947.

Hutchison, K. W., P. B. Singer, C. Diaz-Sala and M. S. Greenwood (1999). "Expansins are conserved in conifers and expressed in response to exogenous auxin." *Plant Physiology* **120**: 827-832.

Huttly, A. K., R. A. Martienssen and D. C. Baulcombe (1988). "Sequence heterogeneity and differential expression of the alpha-Amy2 gene family in wheat." *Molecular and General Genetics* **214**(2): 232-40.

Ikeda, Y., N. Koizumi, T. Kusano and H. Sano (1999). "Sucrose and cytokinin modulation of *WPK4*, a gene encoding a SNF1-related protein kinase from wheat." *Plant Physiology* **121**(3): 813-820.

Ingram, G. C., J. L. Magnard, P. Vergne, C. Dumas and P. M. Rogowsky (1999). "*ZmOCL1*, an *HDGL2* family homeobox gene, is expressed in the outer cell layer throughout maize development." *Plant Molecular Biology* **40**: 343–354.

Itoh, J.-I., K.-I. Nonomura, K. Ikeda , S. Yamaki, Y. Inukai, H. Yamagishi, H. Kitano and Y. Nagato (2005). "Rice plant development: from zygote to spikelet." *Plant and Cell Physiology* **46**(1): 23-47.

Jackson, A. L., S. R. Bartz, J. Schelter, S. V. Kobayashi, J. Burchard, M. Mao, B. Li, G. Cavet and P. S. Linsley (2003). "Expression profiling reveals off-target gene regulation by RNAi." *Nature Biotechnology* **21**: 635 - 637.

Jeong, S., A. E. Trotochaud and S. E. Clark (1999). "The *Arabidopsis CLAVATA2* gene encodes a receptor-like protein required for the stability of the *CLAVATA1* receptor-like kinase." *Plant Cell* **11**: 1925-34.

Jones, H. D., S. Kurup, N. C. B. Peters and M. J. Holdsworth (2000). "Identification and analysis of proteins that interact with the *Avena fatua* homologue of the maize transcription factor *VIVIPAROUS 1*." *Plant Journal* **21**(2): 133-142.

Jorgensen, R. A., P. D. Cluster, J. English, Q. Que and C. A. Napoli (1996). "Chalcone synthase cosuppression phenotypes in petunia flowers: comparison of sense vs. antisense constructs and single-copy vs. complex T-DNA sequences." *Plant Molecular Biology* **5**: 957-73.

Jürgens, G. (2001). "Apical-basal pattern formation in *Arabidopsis* embryogenesis." *EMBO Journal* **20**(14): 3609-3616.

Jürgens, G., R. T. Ruiz and T. Berleth (1994). "Embryonic pattern formation in flowering plants." *Annual Review of Genetics* **28**: 351-371.

Kajiwara, T., M. Furutani, K.-i. Hibara and M. Tasaka (2004). "The *GURKE* Gene Encoding an Acetyl-CoA Carboxylase is Required for Partitioning the Embryo Apex into Three Subregions in *Arabidopsis*." *Plant and Cell Physiology* **45**: 1122-1128.

Kaplan, D. R. and T. J. Cooke (1997). "Fundamental concepts in the embryogenesis of dicotyledons:A morphological interpretation of embryo mutants." *Plant Cell* **9**: 1903-1919.

Ketting, R. F., S. E. Fischer, E. Bernstein, T. Sijen, G. J. Hannon and R. H. Plasterk (2001). "Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*." *Genes and Development* **15**: 2654-2659.

Kimura, S., T. Suzuki, Y. Yanagawa, T. Yamamoto, H. Nakagawa, I. Tanaka, J. Hashimoto and K. Sakaguchi (2001). "Characterization of plant proliferating cell nuclear antigen (PCNA) and flap endonuclease-1 (FEN-1), and their distribution in mitotic and meiotic cell cycles." *Plant Journal* **28**(6): 643-653.

Kirnos, M. D., N. A. Artyukhovskaya, N. I. Aleksandrushkina, V. W. Ashapkin and B. F. Vanyushin (1986). "Effect of phytohormones on replicative and postreplicative methylation of nuclear DNA in the S phase of the cell cycel of cells of the first leaf of etiolated wheat seedlings." *Biokhimiya* **51**: 1875.

Knox, J. P., S. Day and K. Roberts (1989). "A set of surface glycoproteins forms an early marker of cell position, but not cell type, in the root apical meristem of *Daucus carota L.*" *Development* **106**: 47-56.

Knox, J. P., P. J. Linstead, J. Peart, C. Cooper and K. Roberts (1991). "Developmentally regulated epitopes of the cell surface Arabinogalactan proteins and their relation to root tissue pattern formation." *Plant Journal* **1**: 317-326.

Kollipara, K. P., I. N. Saab, R. D. Wych, M. J. Lauer and G. W. Singletary (2002). "Expression profiling of reciprocal maize hybrids divergent for cold germination and desiccation tolerance." *Plant Physiology* **129**(3): 974-992.

Kreuger, M. and G.-J. van Holst (1995). "Arabinogalactan-protein epitopes in somatic embryogenesis of *Daucus carota L.*" *Planta* **197**: 135-141.

Kurek, I., K. Aviezer, N. Erel, E. Herman and A. Breiman (1999). "The wheat peptidyl prolyl cis-trans-isomerase FKBP77 is heat induced and developmentally regulated." *Plant Physiology* **119**(2): 693-704.

Laux, T. and G. Jürgens (1994). "Establishing the body plan of the *Arabidopsis* embryo." *ACTA Botanica Neerlandica* **43**: 247-260.

Laux, T., K. F. X. Mayer, J. Berger and G. Jürgens (1996). "The *WUSCHEL* gene is required for shoot and floral meristem integrity in *Arabidopsis*." *Development* **122**: 87-96.

Leah, R., J. Kigel, I. Svendsen and J. Mundy (1995). "Biochemical and molecular characterization of a barley seed [IMAGE]-glucosidase." *Journal of Biological Chemistry* **270**(26): 15789-15797.

Lei, M. and B. K. Tye (2001). "Initiating DNA synthesis: from recruiting to activating the MCM complex." *Journal of Cell Science* **114**(8): 1447-1454.

Lenhard, M. and T. Laux (2003). "Stem cell homeostasis in the *Arabidopsis* shoot meristem is regulated by intercellular movement of *CLAVATA3* and its sequestration by *CLAVATA1*." *Development* **130**(14): 3163-73.

Lippman, Z., A. Gendrel, M. Black, M. W. Vaughn, N. Dedhia, W. McCombie, K. Lavine, V. Mittal, B. May, K. D. Kasschau, J. C. Carrington, R. W. Doerge, V. Colot and R. Martienssen (2004). "Role of transposable elements in heterochromatin and epigenetic control." *Nature* **430**(6998): 471-476.

Lippman, Z., B. May, C. Yordan, T. Singer and R. Martienssen (2003). "Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification." *PLoS Biology* **1**(3): e67.

Liu, Z., M. P. Running and E. M. Meyerowitz (1997). "*TSO1* functions in cell division during *Arabidopsis* flower development." *Development* **124**(3): 665-672.

Liu, Z. B., T. Ulmasov, X. Shi, G. Hagen and T. J. Guilfoyle (1994). "Soybean GH3 promoter contains multiple auxin-inducible elements." *Plant Cell* **6**(5): 645-57.

Ljung, K., R. P. Bhalerao and G. Sandberg (2001). "Sites and homeostatic control of auxin biosynthesis in *Arabidopsis* during vegetative growth." *Plant Journal* **28**: 465-474.

Lockhart, D. J. and E. A. Winzeler (2000). "Genomics, gene expression and DNA arrays." *Nature* **405**: 827-836.

Long, J. A., E. I. Moan, J. I. Medford and M. K. Barton (1996). "A member of the KNOTTED class of homeodomain proteins encoded by the *STM* gene of *Arabidopsis*." *Nature* **379**(6560): 66-9.

Lotan, T., M. Ohto, K. M. Yee, M. A. West, R. Lo, R. W. Kwong, K. Yamagishi, R. L. Fischer, R. B. Goldberg and J. J. Harada (1998). "*Arabidopsis LEAFY COTYLEDON1* is sufficient to induce embryo development in vegetative cells." *Cell* **93**(7): 1195-205.

Lu, P., R. Porat, J. A. Nadeau and S. D. O'Neill (1996). "Identification of a meristem L1 layer-specific gene in *Arabidopsis* that is expressed during embryonic pattern formation and defines a new class of homeobox genes." *Plant Cell* **8**: 2155–2168.

Lukowitz, W., U. Mayer and G. Jürgens (1996). "Cytokinesis in the *Arabidopsis* embryo involves the syntaxin-related *KNOLLE* gene product." *Cell* **84**(1): 61-71.

Magyar, Z., L. Bako, L. Bogre, D. Dedeoglu, T. Kapros and D. Dudits (1993). "Active cdc2 genes and cell cycle phase specific cdc2-related kinase complexes in hormone stimulated alfalfa cells." *Plant Journal* **4**: 151.

Manche, L., S. R. Green, C. Schmedt and M. B. Mathews (1992). "Interactions between double-stranded RNA regulators and the protein kinase DAI. ." *Molecular and Cellular Biology* **12**: 5238-5248.

Markus, M. A., P. Matsuidera and G. Wagner (1997). "Refined structure of villin 14T and a detailed comparison with other actin-severing domains." *Protein Science* **6**(6): 1197-1209.

Matsushime, H., A. Jinno, N. Takagi and M. Shibuya (1990). "A novel mammalian protein kinase gene (mak) is highly expressed in testicular germ cells at and after meiosis." *Molecular and Cellular Biology* **10**(5): 2261-2268.

Mayer, K. F. X., H. Schoof, A. Haecker, M. Lenhard, G. Jürgens and T. Laux (1998). "Role of *WUSCHEL* in regulating stem cell fate in the *Arabidopsis* shoot meristem." *Cell* **95**: 805-815.

Mayer, U., G. Buttner and G. Jürgens (1993). "Apical-basal pattern formation in the *Arabidopsis* embryo: studies on the role of the *gnom* gene." *Development* **117**: 149-162.

Mayer, U., R. A. Torres Ruiz, T. Berleth, S. Misera and G. Jürgens (1991). "Mutations affecting body organization in the *Arabidopsis* embryo." *Nature* **353**: 402-407.

McClure, B. A., G. Hagen, C. S. Brown, M. A. Gee and T. J. Guilfoyle (1989). "Transcription, organization, and sequence of an auxin-regulated gene cluster in soybean." *Plant Cell* **1**(2): 229-239.

McQueen-Mason, S. and D. J. Cosgrove (1994). "Disruption of hydrogen bonding between wall polymers by proteins that induce plant wall extension." *Proceedings of the National Academy of Sciences of the United States of America* **91**: 6574-6578.

McQueen-Mason, S., D. M. Durachko and D. J. Cosgrove (1992). "Two endogenous proteins that induce cell wall expansion in plants." *Plant Cell* **4**: 1425-1433.

Menges, M., L. Hennig, W. Gruissem and J. A. H. Murray (2002). "Cell cycle regulated gene expression in *Arabidopsis*." *Journal of Biological Chemistry* **44**: 41987-42002.

Mikolajczyk, M., O. S. Awotunde, G. Muszynska, D. F. Klessig and G. Dobrowolska (2000). "Osmotic stress induces rapid activation of a salicylic acid–induced protein kinase and a homolog of protein kinase ASK1 in tobacco cells." *Plant Cell* **12**(1): 165-178.

Minks, M. A., D. K. West, S. Benvin and C. Baglioni (1979). "Structural requirements of double-stranded RNA for the activation of 2'-5'-oligo(A) polymerase and protein kinase of interferon-treated HeLa cells. ." *Journal of Biological Chemistry* **254**: 10180-10183.

Morrish, F. M. and I. K. Vasil (1989). "DNA methylation and embryogenic competence in leaves and callus of Napiergrass (*Pennisetum purpureum Schum*)." *Plant Physiology* **90**(1): 37-40.

Mourelatos, Z., J. Dostie, S. Paushkin, A. Sharma, B. Charroux, L. Abel, J. Rappsilber, M. Mann and G. Dreyfuss (2002). "miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs." *Genes and Development* **16**: 720-728.

Müller, A., C. Guan, L. Gälweiler, P. Tänzler, P. Huijser, A. Marchant, G. Parry, M. Bennett, E. Wisman and K. Palme (1998). "AtPIN2 defines a locus of Arabidopsis for root gravitropism control." *EMBO Journal* **17**: 6903–6911.

Murashige, T. and F. Skoog (1962). "A revised medium for rapid growth and bioassays with tobacco tissue cultures." *Physiology Plantarum* **15**:: 473-497.

Murray, M. G. and J. L. Key (1978). "2,4-Dichlorophenoxyacetic acid-enhanced phosphorylation of soybean nuclear proteins." *Plant Physiology* **61**: 190–198.

Nakajima, K., G. Sena, T. Nawy and P. N. Benfey (2001). "Intercellular movement of the putative transcription factor SHR in root patterning." *Nature* **413**: 307-311.

Nakazono, M., F. Qiu, L. A. Borsuk and P. S. Schnable (2003). "Laser-capture microdissection, a tool for the global analysis of gene expression in specific plant cell types: Identification of genes expressed differentially in epidermal cells or vascular tissues of maize." *Plant Cell* **15**: 583–596.

Nelissen, H., D. Fleury, L. Bruno, P. Robles, L. De Veylder, J. Traas, J. L. Micol, M. Van Montagu, D. Inze and M. Van Lijsebettens (2005). "The elongata mutants identify a functional Elongator complex in plants with a role in cell proliferation during organ growth." *Proceedings of the National Academy of Sciences* **102**(21): 7754-7759.

Nuti-Ronchi, V. N., L. Giorgetti, M. G. Tonelli and G. Martini (1992). "Ploidy reduction and genome segregation in cultured carrot cell lines .II. Somatic meiosis." *Plant and Tissue Organ Culture* **30**: 115-119.

Ogas, J., S. Kaufmann, J. Henderson and C. Somerville (1999). "*PICKLE* is a CHD3 chromatin-remodeling factor that regulates the transition from embryonic to vegetative development in

Arabidopsis." *Proceedings of the National Academy of Sciences of the United States of America* 96(24): 13839-13844.

Okada, K., J. Ueda, M. K. Komaki, C. J. Bell and Y. Shimura (1991). "Requirement of the Auxin Polar Transport System in Early Stages of Arabidopsis Floral Bud Formation." *Plant Cell* **3**: 677-684.

Old, W. and S. B. Primrose (1994). Principles of Gene Manipulation, Blackwell Science.

Olsen, P. H. and V. Ambros (1999). "The *lin-4* regulatory RNA controls developmental timing in *caenorhabditis elegans* by blocking LIN-14 protein synthesis after the initiation of translation." *Developmental Biology* **216**: 671-680.

Palatnik, J. F., E. Allen, X. Wu, C. Schommer, R. Schwab, J. C. Carrington and D. Weigel (2003). "Control of leaf morphogenesis by microRNAs." *Nature* **425**(18): 257-63.

Parry, G., A. Delbarre, A. Marchant, R. Swarup, R. Napier, C. Perrot-Rechenmann and M. J. Bennett (2001). "Novel auxin transport inhibitors phenocopy the auxin influx carrier mutation aux1." *Plant Journal* **25**: 399-406.

Pellegrineschi, A., L. M. Noguera, B. Skovmand, R. M. Brito, L. Velazquez, M. M. Salgado, R. Hernandez, M. Warburton and D. Hoisington (2002). "Identification of highly transformable wheat genotypes for mass production of fertile transgenic plants." *Genome: Génome* **45**(2): 421-430.

Pharmaceuticals, I. (1999). "GEM microarray reproducibility study." *Incyte technical survey*.

Pitto, M., T. Mutoh, M. Kuriyama, A. Ferraretto, P. Palestini and M. Masserini (1998). "Influence of endogenous GM1 ganglioside on TrkB activity in cultured neurons." *FEBS Letters* **439**: 93-96.

Poggeler, S. and U. Kuck (2004). "A WD40 repeat protein regulates fungal cell differentiation and can be replaced functionally by the mammalian homologue striatin." *Eukaryotic Cell* **3**(1): 232-240.

Pysh, L. D., J. W. Wysocka-Diller, C. Camilleri, D. Bouchez and P. N. Benfey (1999). "The GRAS gene family in *Arabidopsis*: sequence characterization and basic expression analysis of the SCARECROW-LIKE genes." *Plant Journal* **18**(1): 111-9.

Raghavan, C., E. Kok Ong, M. J. Dalling and T. W. Stevenson (2004). "Effect of herbicidal application of 2,4-dichlorophenoxyacetic acid in *Arabidopsis*." *Functional and Integrative Genomics*.

Rakszegi, M., C. Tamas, P. Szucs, L. Tamas and Z. Bedo (2001). "Current status of wheat transformation." *Journal of Plant Biotechnology* **3**(2): 67-81.

Redway, F. A., V. Vasil, D. Lu and L. K. Vasil (1990). "Identification of callus types for long term maintenance and regeneration from commercial cultivars of wheat (*Triticum aestivum L.*)." *Theoretical Applications of Genetics* **79**(609): 617.

Reeves, R. and M. S. Nissen (1990). "The A.T-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure." *Journal of Biological Chemistry* **265**(15): 8573-8582.

Reinert, J. (1958). "Morphogenese und ihre kontrolle an gewebekulturen aus carotten." *Naturwissenchaften* **45**:: 344-345.

Reinhardt, D., E. R. Pesce, P. Stieger, T. Mandel, K. Baltensperger, M. Bennett, J. Traas, J. Friml and C. Kuhlemeier (2003). "Regulation of phyllotaxis by polar auxin transport." *Nature* **426**(6964): 255-60.

Reinhart, B. J., F. J. Slack, M. Basson, A. E. Pasquinelli, J. C. Bettinger, A. E. Rougvie, H. R. Horvitz and G. Ruvkun (2000). "The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*." *Nature* **403**(6772): 901-6.

Repp, A., K. Mikami, F. Mittmann and E. Hartmann (2004). "Phosphoinositide-specific phospholipase C is involved in cytokinin and gravity responses in the moss *Physcomitrella patens*." *Plant Journal* **40**(2): 250-259.

Rhoades, M. W., B. J. Reinhart, L. P. Lim, C. B. Burge, B. Bartel and D. P. Bartel (2002). "Prediction of plant microRNA targets." *Cell* **110**(4): 513-520.

Rodriguez, P. L., M. P. Leube and E. Grill (1998). "Molecular cloning in *Arabidopsis thaliana* of a new protein phosphatase 2C (PP2C) with homology to ABI1 and ABI2." *Plant Molecular Biology Reporter* **38**(5): 879-83.

Roe, J. L., J. L. Nemhauser and P. C. Zambryski (1997). "*TOUSLED* participates in apical tissue formation during gynoecium development in *Arabidopsis*." *Plant Cell* **9**(3): 335-353.

Roe, J. L., C. J. Rivin, R. A. Sessions, K. A. Feldmann and P. C. Zambryski (1993). "The *Tousled* gene in *A. thaliana* encodes a protein kinase homolog that is required for leaf and flower development." *Cell* **75**(5): 939-50.

Sabatini, S., R. Heidstra, M. Wildwater and B. Scheres (2003). "*SCARECROW* is involved in positioning the stem cell niche in the *Arabidopsis* root meristem." *Genes and Development* **17**:: 354–358.

Sabelli, P. A., S. R. Burgess, L. Valásek and P. R. Shewry (1999). "Molecular cloning and characterisation of a maize cDNA for a homologue of the large subunit of the eukaryotic initiation factor 3 (eIF3)." *Molecular and General Genetics* **261**(4-5): 820 - 830.

Schaeffer, H. J. and M. J. Weber (1999). "Mitogen-activated protein kinases: Specific messages from ubiquitous messengers." *Molecular and Cellular Biology* **19**(4): 2435-2444.

Schäfer, S. and T. Schmülling (2002). "The CRK1 receptor-like kinase gene of tobacco is negatively regulated by cytokinin." *Plant Molecular Biology Reporter* **50**: 155–166.

Schiavone, M. and T. Cook (1987). "Unusual patterns of embryogenesis in the domesticated carrot: developmental effects of exogenous auxins and auxin transport inhibitors." *Cell Differentiation* **21**: 53-62.

Schmidt, E. D., F. Guzzo, M. A. Toonen and S. C. de Vries (1997). "A leucine-rich repeat containing receptor-like kinase marks somatic plant cells competent to form embryos." *Development* **124**(10): 2049-62.

Schmitz, J., R. Franzen, T. H. Ngyuen, F. Garcia-Maroto, C. Pozzi, F. Salamini and W. Rohde (2000). "Cloning, mapping and expression analysis of barley MADS-box genes." *Plant Molecular Biology* **42**(6): 899-913.

Schoof, H., M. Lenhard, A. Haecker, K. F. X. Mayer, G. Jürgens and T. Laux (2000). "The stem cell population of *Arabidopsis* shoot meristems is maintained by a regulatory loop between the *CLAVATA* and *WUSCHEL* genes." *Cell* **100**: 635-644.

Schrick, K., U. Mayer, A. Horrichs, C. Kuhnt, C. Bellini, J. Dangl, J. Schmidt and G. Jürgens (2000). "*FACKEL* is a sterol C-14 reductase required for organized cell division and expansion in *Arabidopsis* embryogenesis." *Genes and Development* **14**(12): 1471-84.

Sessions, A., J. L. Nemhauser, A. McColl, J. L. Roe, K. A. Feldmann and P. C. Zambryski (1997). "*ETTIN* patterns the *Arabidopsis* floral meristem and reproductive organs." *Development* **124**(22): 4481-91.

Shah, O. J., J. A. Iniguez-Lluhi, A. Romanelli, S. R. Kimball and L. S. Jefferson (2002). "The activated glucocorticoid receptor modulates presumptive autoregulation of ribosomal protein S6 protein kinase, p70 S6K." *Journal of Biological Chemistry* **277**(4): 2525-33.

Sheppard, L. A., A. M. Brunner, K. V. Krutovskii, W. H. Rottmann, J. S. Skinner, S. S. Vollmer and S. H. Strauss (2000). "*A DEFICIENS* homolog from the dioecious tree black cottonwood is expressed in female and male floral meristems of the two-whorled, unisexual flowers." *Plant Physiology* **124**(2): 627-640.

Shevell, D. E., W. M. Leu, C. S. Gillmor, G. Xia, K. A. Feldmann and N. H. Chua (1994). "*EMB30* is essential for normal cell division, cell expansion, and cell adhesion in *Arabidopsis* and encodes a protein that has similarity to *Sec7*." *Cell* 77(7): 1051-62.

Simon, R., M. I. Igeno and G. Coupland (1996). "Activation of floral meristem identity genes in *Arabidopsis*." *Nature* 384(6604): 59-62.

Singer, J. D., M. Gurian-West, B. Clurman and J. M. Roberts (1999). "Cullin-3 targets cyclin E for ubiquitination and controls S phase in mammalian cells." *Genes and Development* **13**(18): 2375-2387.

Smith, N. A., S. P. Singh, M. Wang, P. A. Stoutjesdijk, A. G. Green and P. M. Waterhouse (2000). "Gene expression: Total silencing by intron-spliced hairpin RNAs." *Nature* 407(6802): 319-320.

Soares, M. B. and M. F. Bonaldo (1998). Construction and screening of normalized cDNA libraries. New York, Cold Spring Harbor Laboratory Press.

Song, W., S. Koh, M. Czako, L. Marton, E. Drenkard, J. M. Becker and G. Stacey (1997). "Anti-sense expression of the peptide transport gene, AtPTR2B, delays flowering and arrests seed development in transgenic *Arabidopsis p*lants." *Plant Physiology* **114**: 927-935.

Springer, N. M., C. A. Napoli, D. A. Selinger, R. Pandey, K. C. Cone, V. L. Chandler, H. F. Kaeppler and S. M. Kaeppler (2003). "Comparative analysis of SET domain proteins in Maize and *Arabidopsis* reveals multiple duplications preceding the divergence of monocots and dicots." *Plant Physiology* **132**(2): 907-925.

Stacey, M. G., S. Koh, J. M. Becker and G. Stacey (2002). "AtOPT3, a member of the oligopeptide transporter family, is essential for embryo development in *Arabidopsis*." *Plant Cell* **14**: 2799-2811.

Stavolone, L., M. Kononova, S. Pauli, A. Ragozzino, P. de Haan, S. Milligan, K. Lawton and T. Hohn (2003). "Cestrum yellow leaf curling virus (CmYLCV) promoter: a new strong constitutive promoter for heterologous gene expression in a wide variety of crops." *Plant Molecular Biology* **53**(5): 663-673.

Steinmann, T., N. Geldner, M. Grebe, S. Mangold, C. L. Jackson, S. Paris, L. Gälweiler, K. Palme and G. Jurgens (1999). "Coordinated polar localization of auxin efflux carrier *PIN1* by *GNOM* ARF GEF." *Science* **286**: 316-318.

Steward, F. C., M. O. Mapes and K. Mears (1958). "Growth and organized development of cultured cells: II. Organization in cultures grown from freely suspended cells." *American Journal of Botany* **45**: 705-708.

Stone, J. M., M. A. Collinge, R. D. Smith, M. A. Horn and J. C. Walker (1994). "Interaction of a protein phosphatase with an Arabidopsis serine-threonine receptor kinase." *Science* **266**(5186): 793-5.

Stone, S. L., L. W. Kwong, K. M. Yee, J. Pelletier, L. Lepiniec, R. L. Fischer, R. B. Goldberg and J. J. Harada (2001). "*LEAFY COTYLEDON2* encodes a B3 domain transcription factor that induces embryo development." *Proceedings of the National Academy of Sciences of the United States of America* 98(20): 11806-11.

Sung, Z. R. and R. Okimoto (1981). "Embryonic proteins in somatic embryos of carrot." *Proceedings of the National Academy of Sciences of the United States of America* 78: 3683-3687.

Suzuka, I., S. Hata, M. Matsuoka, S. Kosugi and J. Hashimoto (1991). "Highly conserved structure of proliferating cell nuclear antigen (DNA polymerase delta auxiliary protein) gene in plants." *European Journal of Biochemistry* **2**: 571-5.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin and D. G. Higgins (1997). "The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools." *Nucleic Acids Research* **24**: 4877-4882.

Tijsterman, M., R. F. Ketting and R. H. A. Plasterk (2002). "The genetics of RNA silencing." *Annual Review of Genetics* **36**(1): 489-519.

Torres-Ruiz, R. A., A. Lohner and G. Jürgens (1996). "The *GURKE* gene is required for normal organization of the apical region in the *Arabidopsis* embryo." *Plant Journal* **10**: 1005–1016.

Tsiantis, M., R. Schneeberger, J. F. Golz, M. Freeling and J. A. Langdale (1999). "The maize *rough sheath2* gene and leaf development programs in monocot and dicot plants." *Science* **284**: 154–156.

Tsuda, K., S. Tsvetanov, S. Takumi, N. Mori, A. Atanassov and C. Nakamura (2000). "New members of a cold-responsive group-3 Lea/Rab-related Cor gene family from common wheat (*Triticum aestivum L.*)." *Genes and Genetic Systems* **75**: 179-188.

Tsuruga, H., N. Yabuta, S. Hosoya, K. Tamura, Y. Endo and H. Nojima (1997). "HsMCM6: a new member of the human MCM/P1 family encodes a protein homologous to fission yeast *Mis5*." *Genes to Cells* **2**(6): 381-99.

Turner, P. R., M. P. Sheetz and L. A. Jaffe (1984). "Fertilization increases the polyphosphoinositide content of sea urchin eggs." *Nature* **310**(5976): 414-5.

Ulmasov, T., G. Hagen and T. J. Guilfoyle (1997). "*ARF1*, a transcription factor that binds to auxin response elements." *Science* **276**(5320): 1865-8.

Ulmasov, T., J. Murfett, G. Hagen and T. J. Guilfoyle (1997). "Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements." *Plant Cell* **9**(11): 1963-71.

Van den Berg, C., V. Willemsen, G. Hendriks, P. Weisbeek and B. Scheres (1997). "Short range control of cell differentiation in the *Arabidopsis* root meristem." *Nature* **390**: 287-89.

Van Hengel, A. J., Z. Tadesse, P. Immerzeel, H. Schols, A. van Kammen and S. C. de Vries (2001). "N-acetylglucosamine and glucosamine-containing arabinogalactan proteins control somatic embryogenesis." *Plant Physiology* **125**(4): 1880-90.

Van Hengel, A. J., A. Van Kammen and S. C. De Vries (2002). "A relationship between seed development, Arabinogalactan-proteins (AGPs) and the AGP mediated promotion of somatic embryogenesis." *Physiologia Plantarum* **114**(4): 637-644.

Varner, J. E. and L.-S. Lin (1989). "Plant cell wall architecture." *Cell* **56**: 231-239.

Verdel, A., S. Jia, S. Gerber, T. Sugiyama, S. Gygi, S. I. Grewal and D. Moazed (2004). "RNAi-mediated targeting of heterochromatin by the RITS complex." *Science* **303**(5658): 672 - 6.

Waterhouse, P. M., M. B. Wang and T. Lough (2001). "Gene silencing as an adaptive defence against viruses." *Nature* **411**: 834-842.

Weeks, I. T., O. D. Anderson and A. E. Blechl (1993). "Rapid production of multiple independent lines of fertile transgenic wheat (*Triticum aestivum*)." *Plant Physiology* **102**: 1077-1084.

Weijers, D. and G. Jurgens (2004). "Funneling auxin action: specificity in signal transduction." *Current Opinions in Plant Biology* **7**(6): 687-93.

Wernicke, W. and L. Milkovits (1987). "Effect of auxin on the mitotic cell cycle in cultured leaf segments at different stages of development in wheat." *Physiologia Plantarum* **69**(9): 16-22.

Wessenger, M., S. Heimes, L. Reidel and H. L. Sanger (1994). "RNA directed de novo methylation of genetic sequences in plants." *Cell* **76**: 567-576.

West, M., K. M. Yee, J. Danao, J. L. Zimmerman, R. L. Fischer, R. B. Goldberg and J. J. Harada (1994). "*LEAFY COTYLEDON1* Is an essential regulator of late embryogenesis and cotyledon identity in *Arabidopsis*." *Plant Cell* **6**(12): 1731-1745.

Wilde, H. D., W. S. Nelson, J. Booij, S. C. De Vries and T. L. Thomas (1988). "Gene expression programs in embryogenic and non-embryogenic carrot cultures." *Planta* **176**: 205.

Willemsen, V. and B. Scheres (2004). "Mechanisms of pattern formation in plant embryogenesis." *Annual Review of Genetics* **38**: 587-614.

Wingard, S. A. (1928). "Hosts and symptoms of ring spot, a virus disease of plants." *Journal of Agricultural Research* **37**: 127-153.

Xie, Z., L. K. Johansen, A. M. Gustafson, K. D. Kasschau, A. D. Lellis, D. Zilberman, S. E. Jacobsen and J. C. Carrington (2004). "Genetic and functional diversification of small RNA pathways in plants." *PLoS Biology* **2**(642-652).

Yang, T. and B. W. Poovaiah (2000). "Molecular and Biochemical Evidence for the Involvement of Calcium/Calmodulin in Auxin Action." *Journal of Biological Chemistry* **275**(5): 3137-3143.

Yang, Y. H. and T. Speed (2002). "Design issues for cDNA microarray experiments." *Nature Reviews Genetics* **3**: 579-588.

Yoo, J. H., M. S. Cheong, C. Y. Park, B. C. Moon, M. C. Kim, Y. H. Kang, H. C. Park, M. S. Choi, J. H. Lee, W. Y. Jung, H. W. Yoon, W. S. Chung and M. J. Cho (2004). "Regulation of the dual specificity protein phosphatase, DsPTP1, through interactions with Calmodulin." *Journal of Biological Chemistry* **279**(2): 848-58.

Yu, D., C. Chen and Z. Chen (2001). "Evidence for an important role of WRKY DNA binding proteins in the regulation of *NPR1* gene expression." *Plant Cell* **13**(7): 1527-1540.

Yu, F. X., H. Q. Sun, P. A. Janmey and H. L. Yin (1992). "Identification of a polyphosphoinositide-binding sequence in an actin monomer-binding domain of gelsolin." *Journal of Biological Chemistry* **267**(21): 14616-14621.

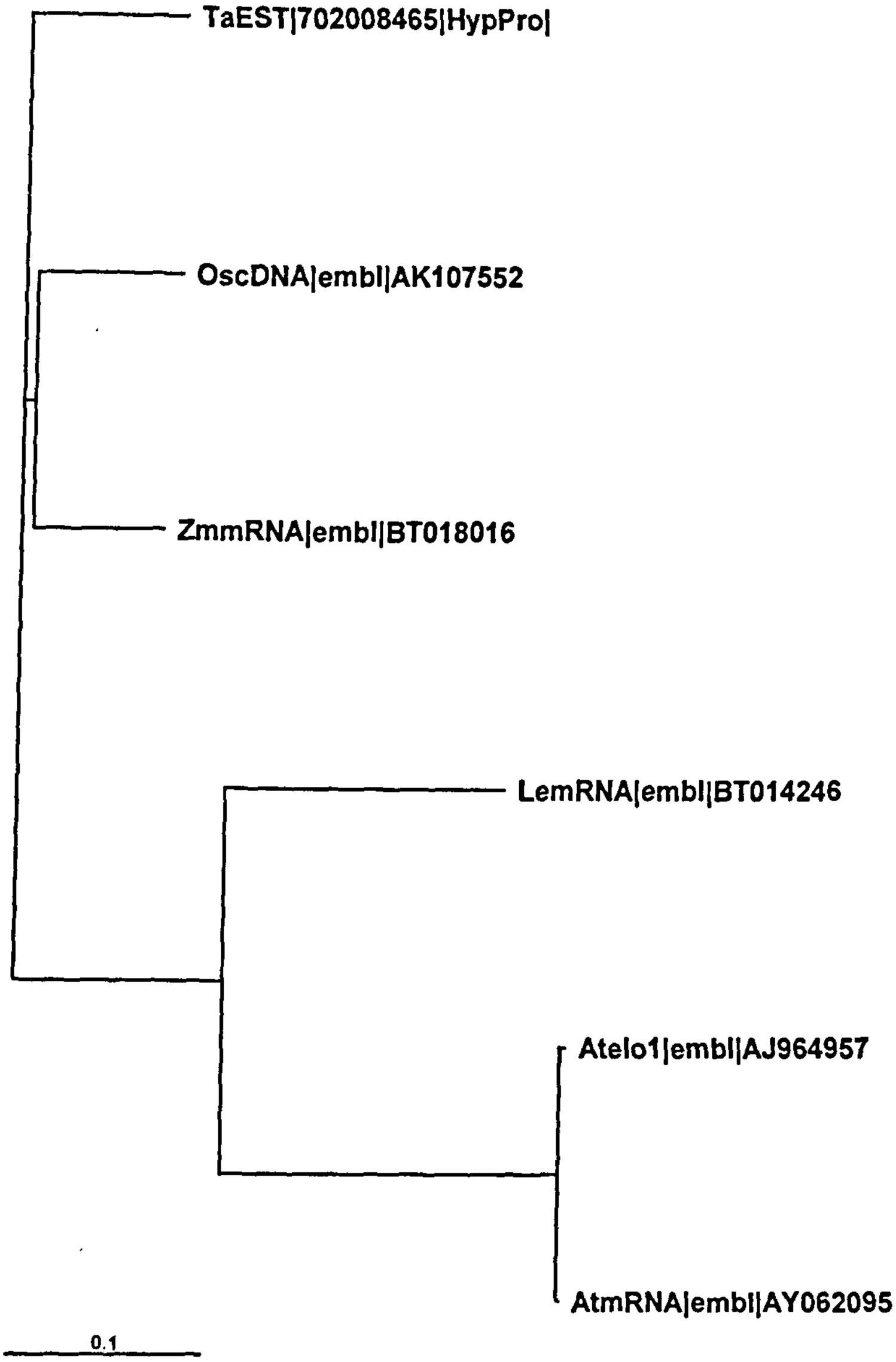
Zheng, L., Liu, S. Batalov, D. Zhou, A. Orth, S. Ding and P. G. Schultz (2004). "An approach to genomewide screens of expressed small interfering RNAs in mammalian cells." *Proceedings of the National Academy of Sciences of the United States of America* **101**(1): 135-140.

Zimmerman, J. L. (1993). "Somatic embryogenesis: A model for early development in higher plants." *Plant Cell* **5**: 1411-1423.

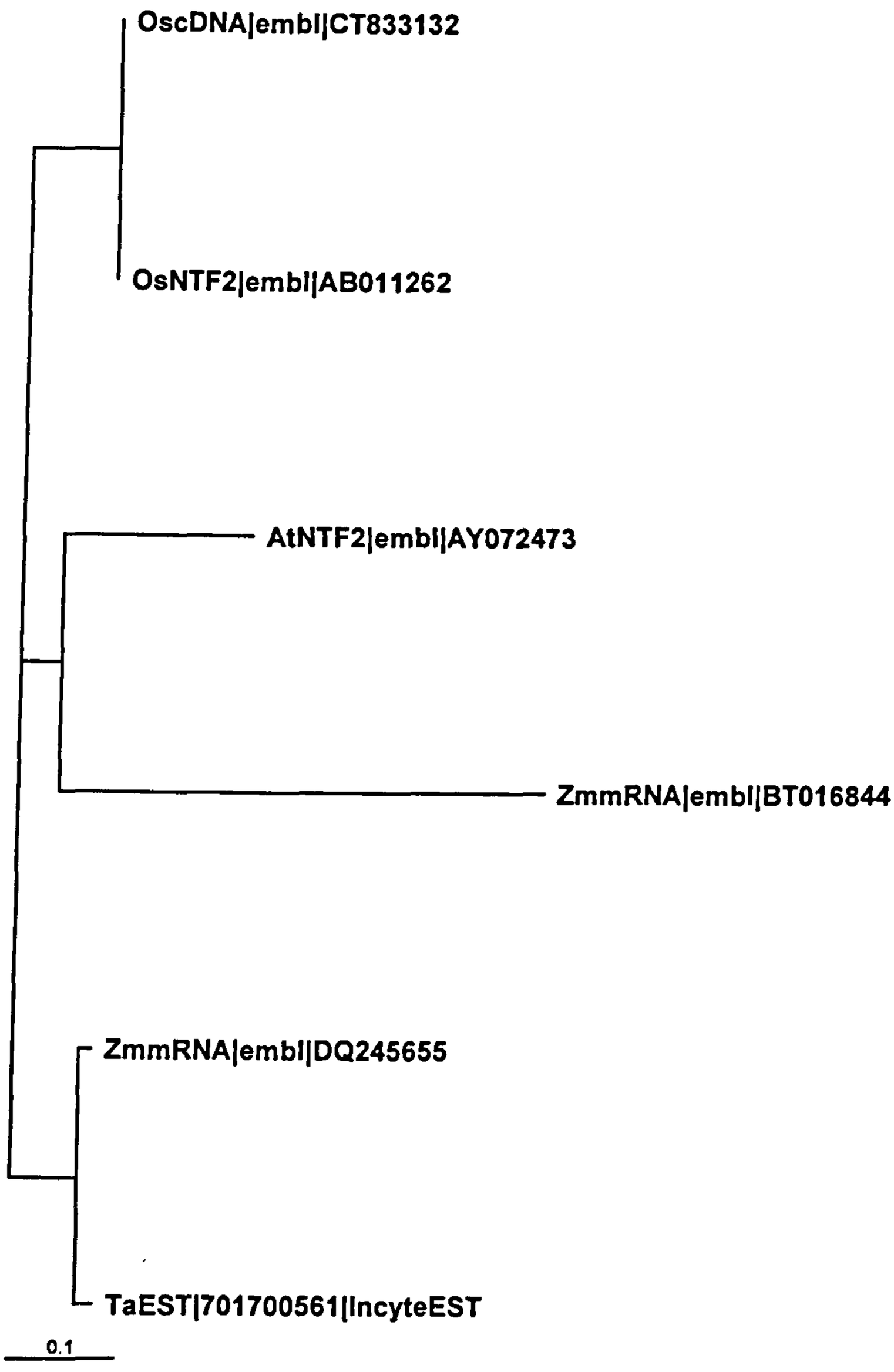
Zuo, J., Q.-W. Niu, G. Frugis and N.-H. Chua (2002). "The *WUSCHEL* gene promotes vegetative-to-embryonic transition in *Arabidopsis*." *Plant Journal* **30**(3): 349-359.

Appendix

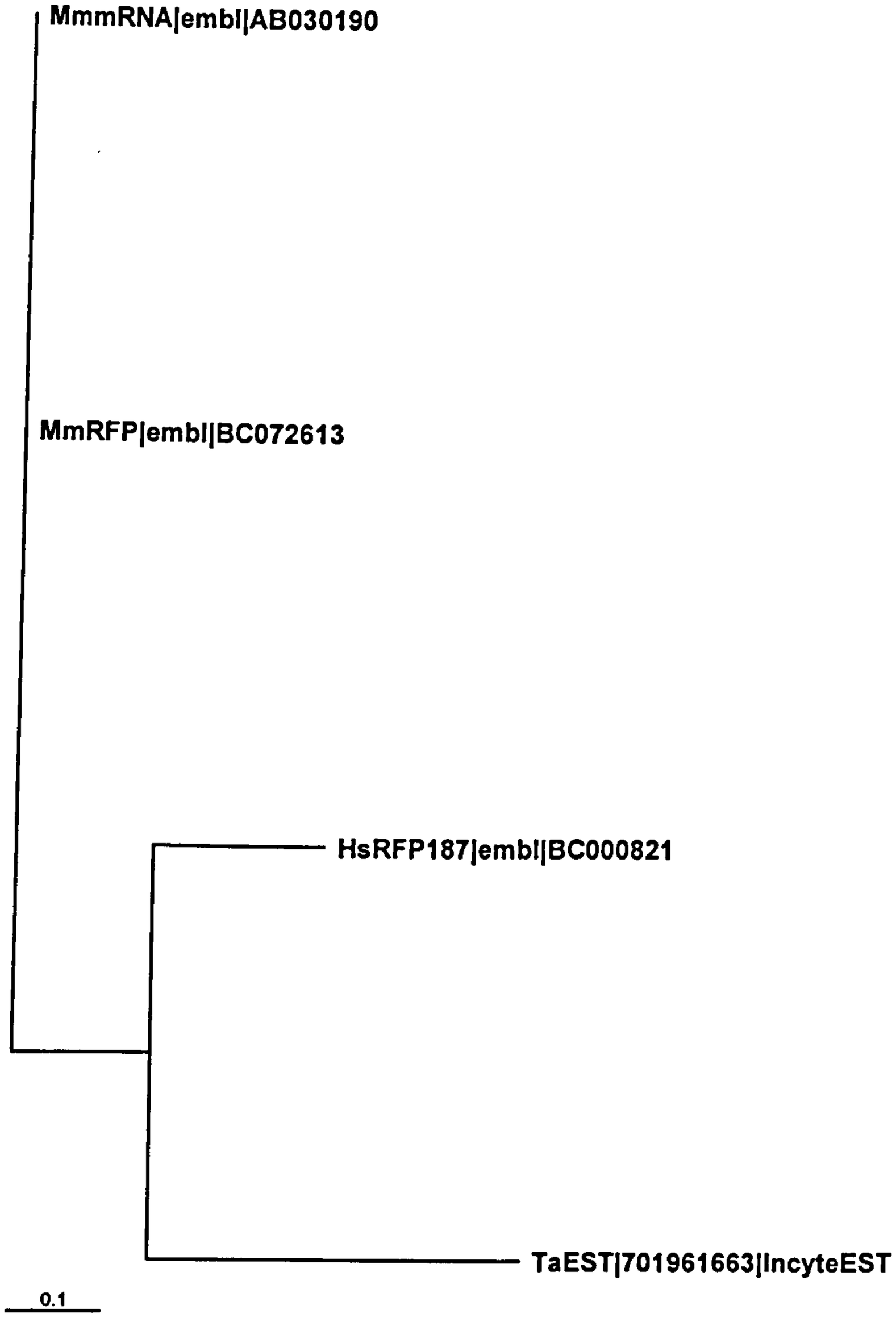
Phylogenetic Trees Inferring New Functional Annotation



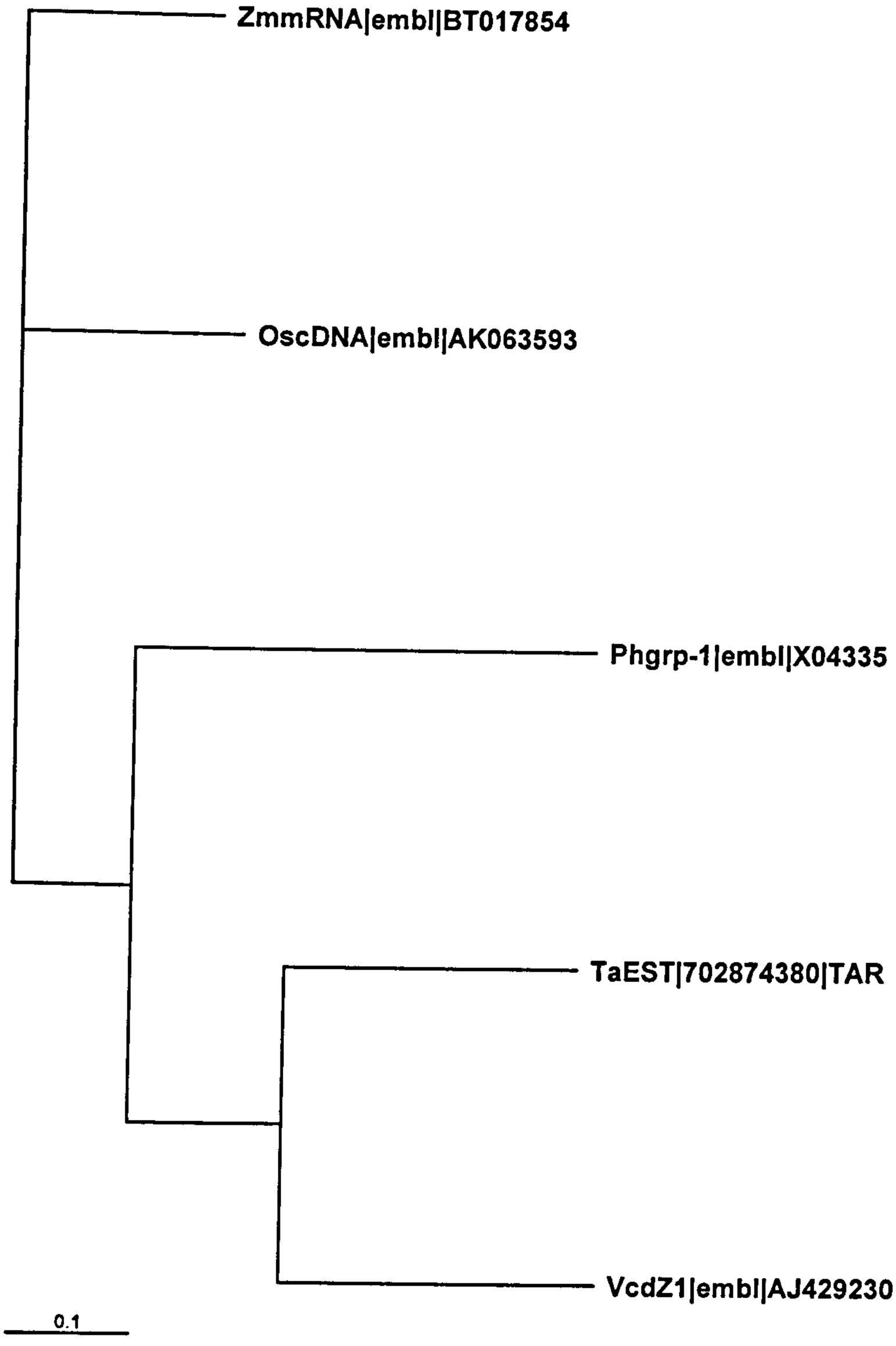
Phylogenetic tree for TaEST|702008465|HypPro produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* Le=*Lycopersicon esculentum;* At=*Arabidopsis thaliana;* EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function; elo1=*ELONGATA1* gene. Scale bar indicates the number of nucleotide replacements per site.



Phylogenetic tree for TaEST|701700561|IncyteEST produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* At=*Arabidopsis thaliana;* NTF2=Nuclear Transport Factor 2; EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function.
Scale bar indicates the number of nucleotide replacements per site.

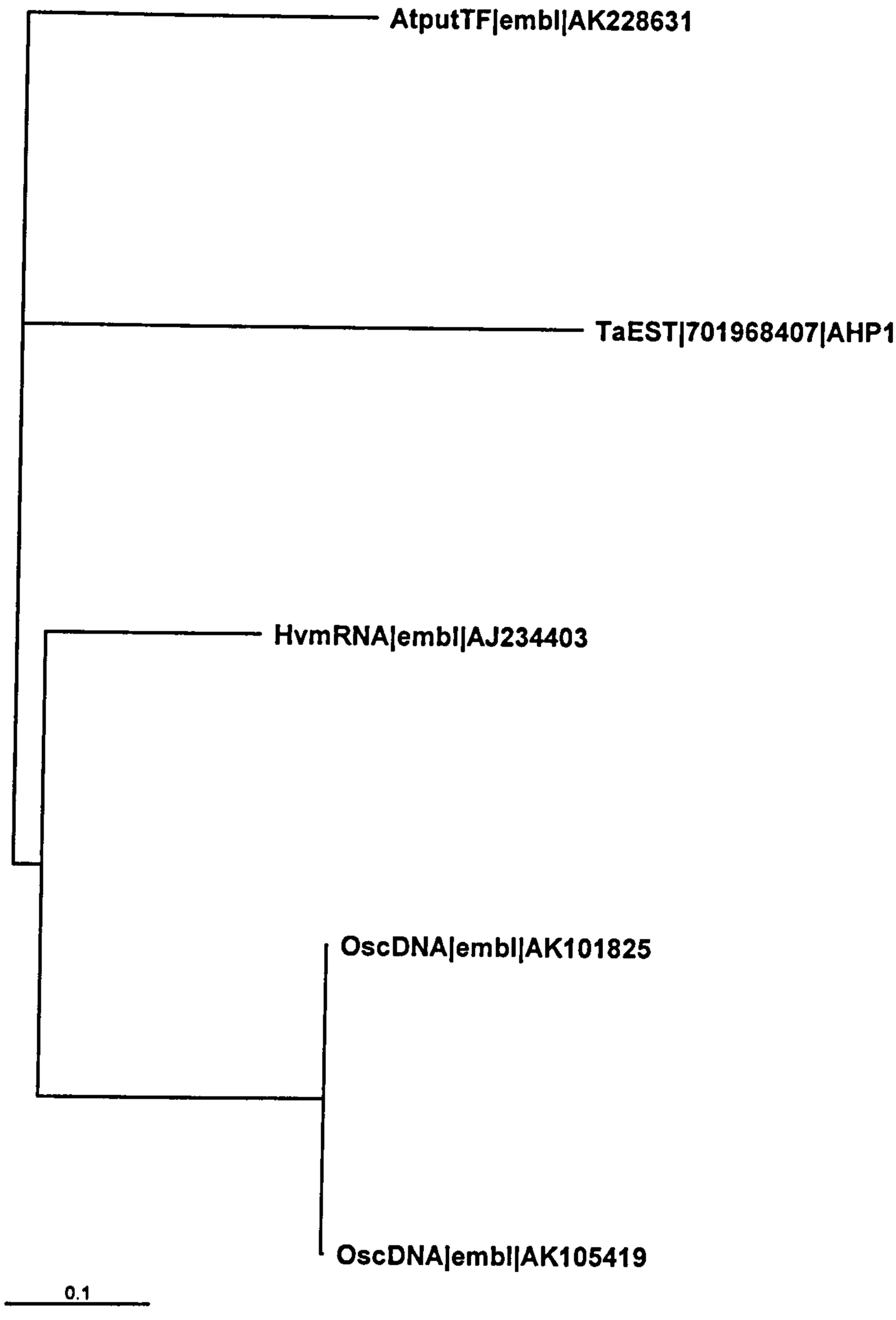


Phylogenetic tree for TaEST|701961663|IncyteEST produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Mm=*Mus musculus;* Hs=*Homo Sapiens;* RFP=ring finger protein; EST=Expressed sequence tag; mRNA=mRNA of no known function. Scale bar indicates the number of nucleotide replacements per site.



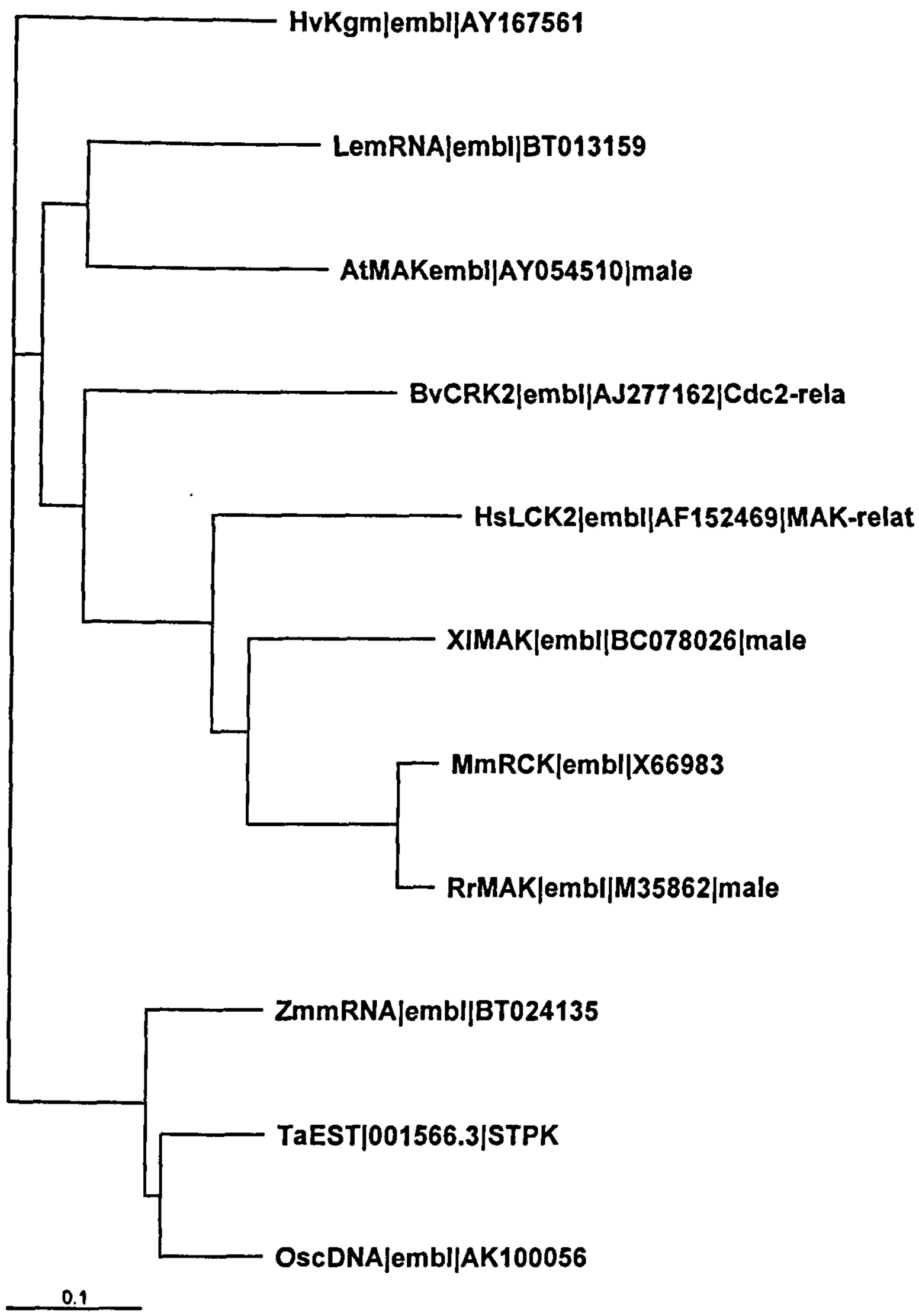
Phylogenetic tree for TaEST|702874380|TAR produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* Ph=*Petunia hybrida;* Vc=*Volvox carteri;* TAR=TAR RNA loop binding protein; dZ1=pherophorin-dz1 protein; grp-1= glycine-rich protein 1; EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

Phylogenetic tree for TaEST|701966045|CDRP produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* Ll=*Lilium longiflorum;* GlsA= Gonidia forming protein GlsA; CDRP=Cell Division Related Protein; EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

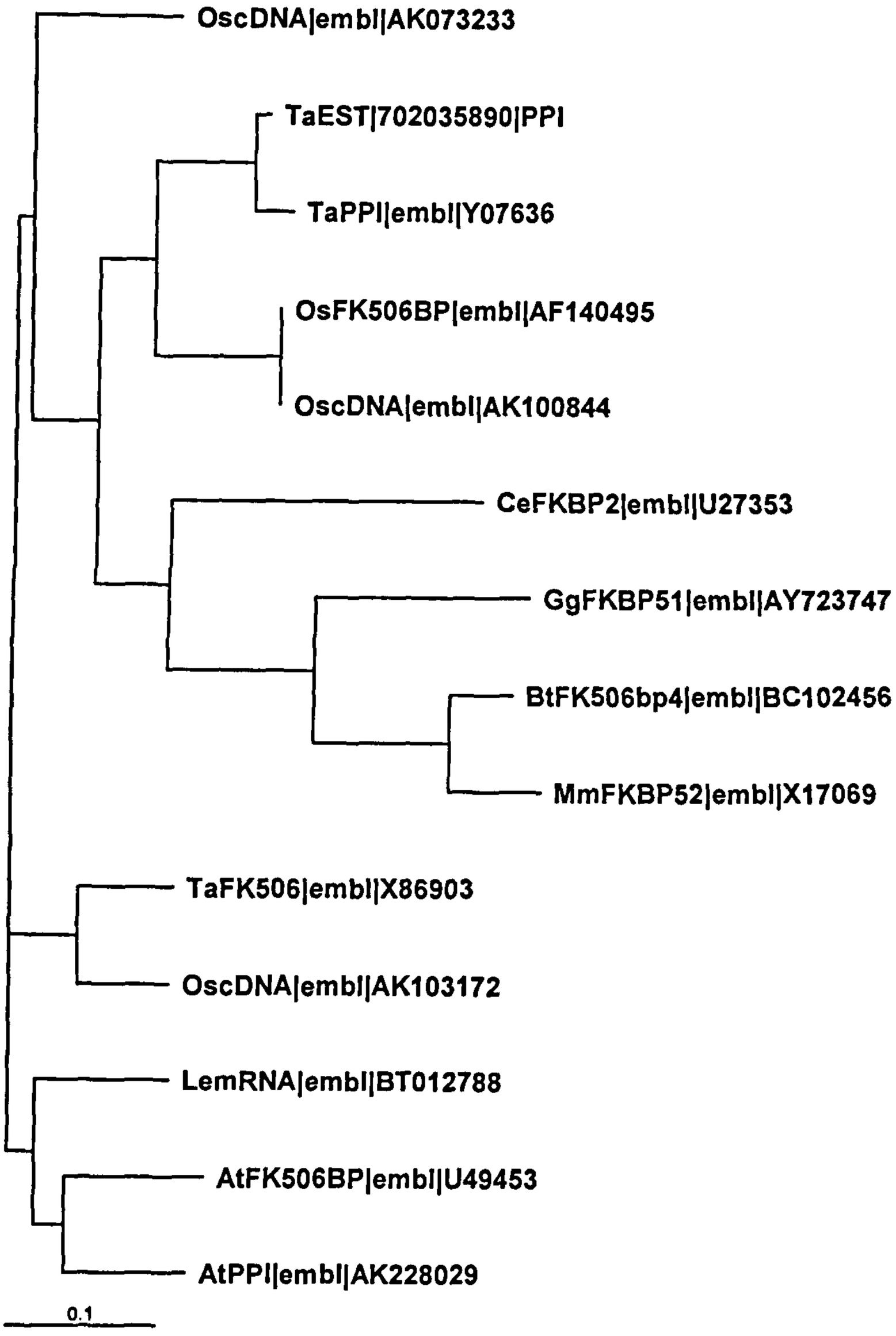


Phylogenetic tree for TaEST|701968407|AHP1 produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Hv=*Hordeum vulgare;* At=*Arabidopsis thaliana;* AHP1=Alkyl HydroPeroxide reductase; PutTF=Putative Transcription Factor; EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

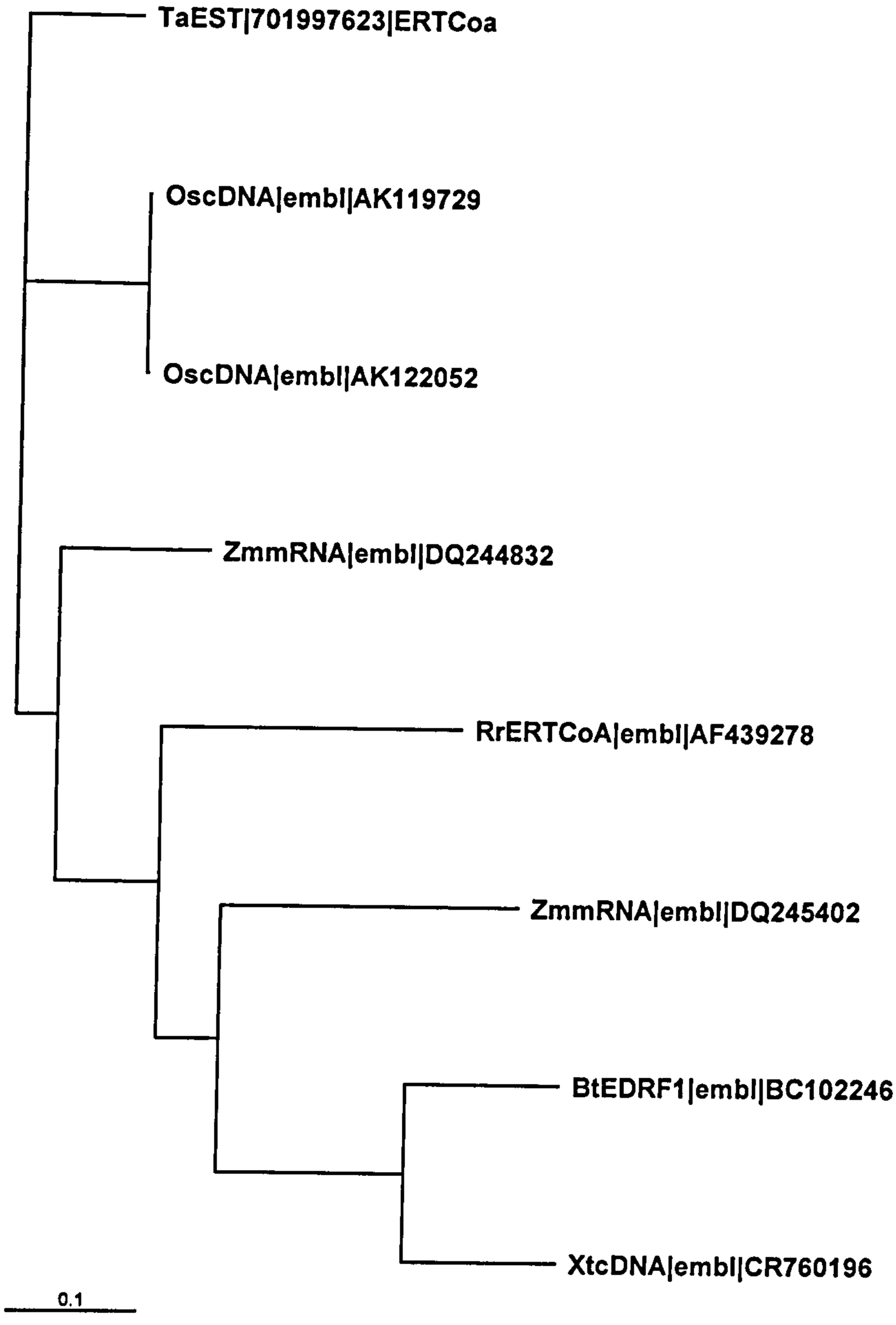
Phylogenetic Trees Confirming Existing Functional Annotation



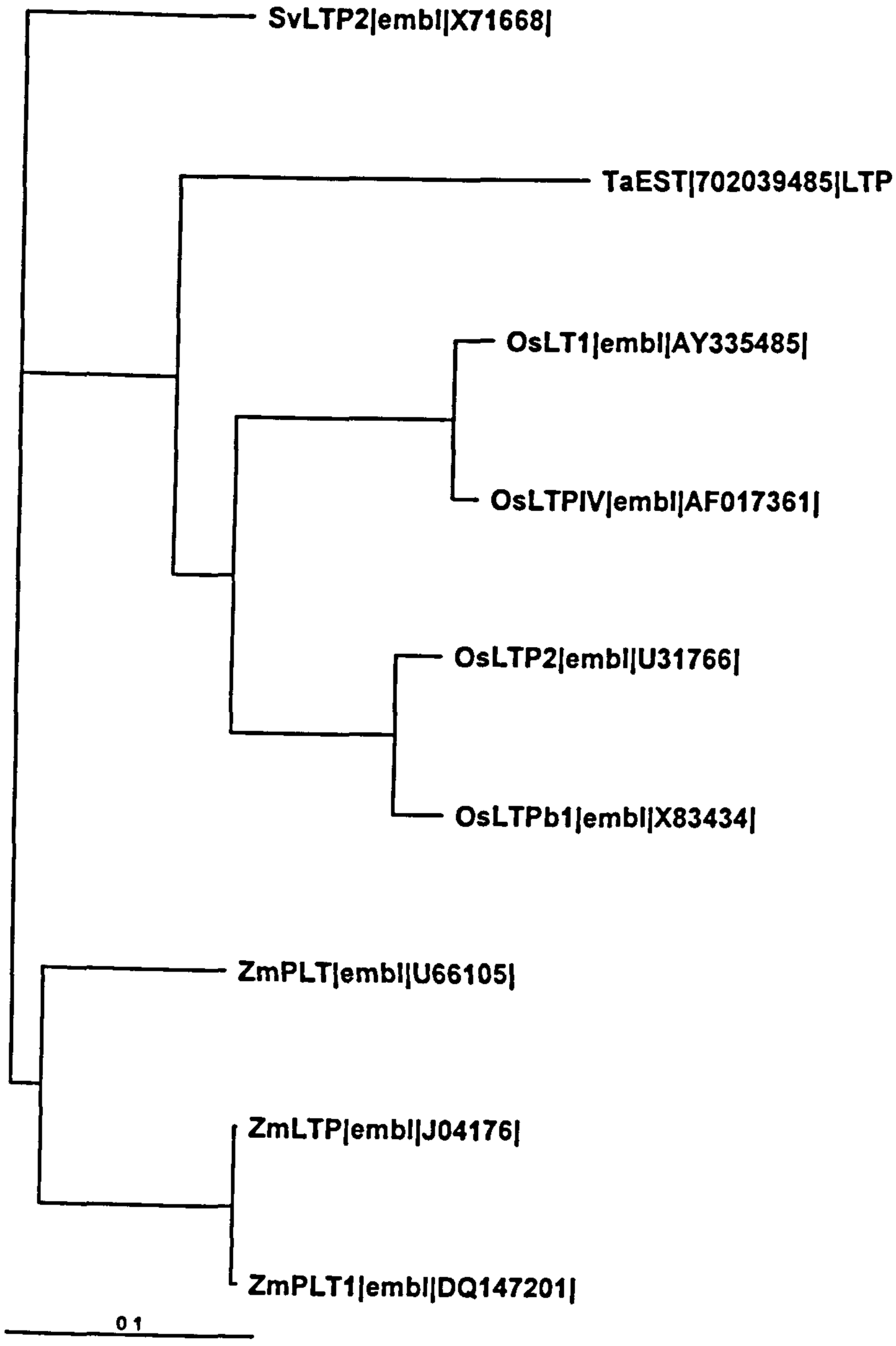
Phylogenetic tree for TaEST|001566.3|STPK produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* Hv=*Hordeum vulgare;* At=*Arabidopsis thaliana;* Bv=*Beta vulgaris;* Le=*Lycopersicon esculentum;* Mm=*Mus musculus;* Xl=*Xenopus laevis;* Hs=*Homo Sapiens;* Dm=*Drosophila melanogaster;* Rr=*Rattus rattus*; GM=GAMYB-binding protein; MAK=male germ cell-associated kinase; CRK= Cdc2-related serine/threonine protein kinases; LCK=laryngeal cancer kinase; RCK=male germ cell protein kinase; EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.



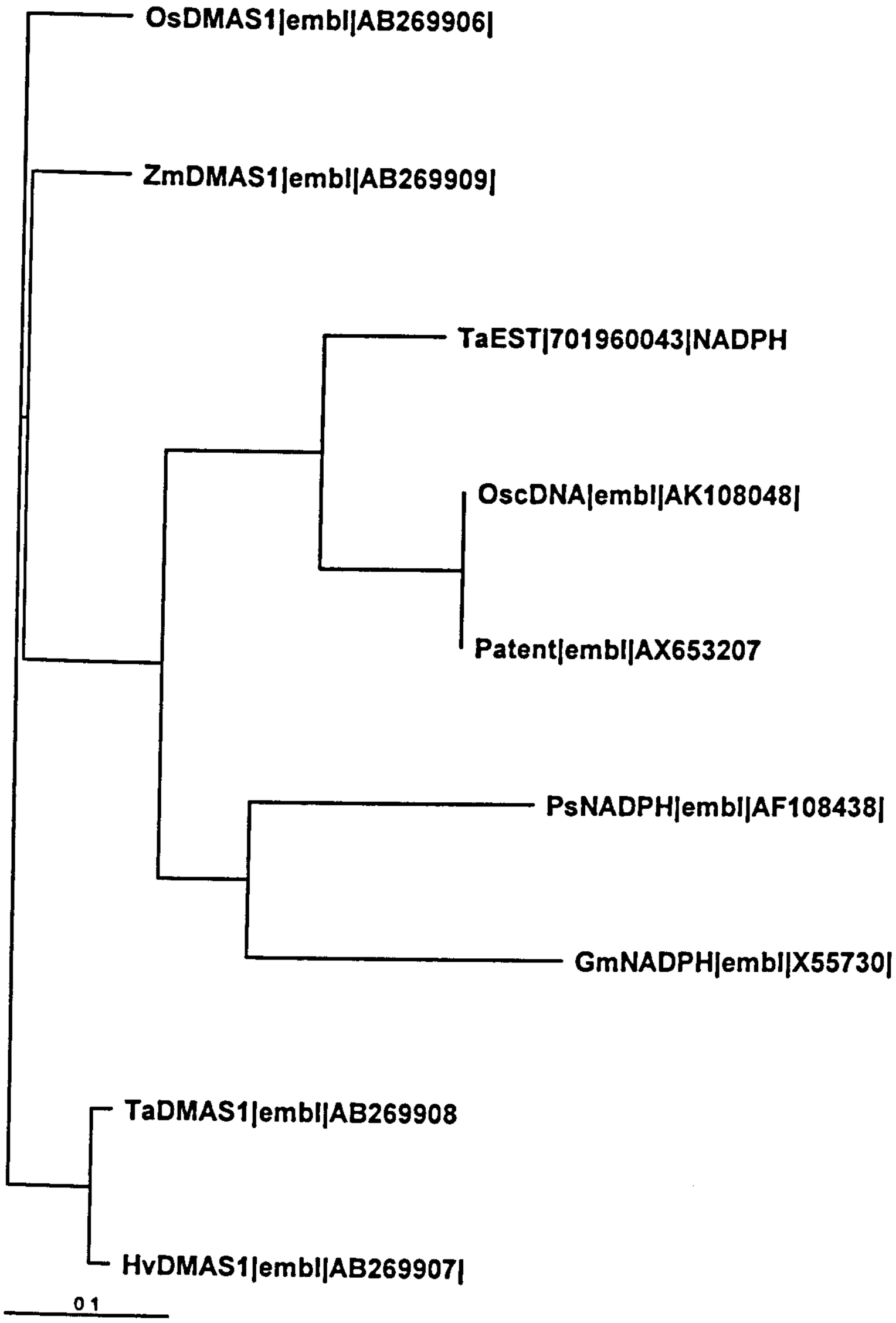
Phylogenetic tree for TaEST|702035890|PPI produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* At=*Arabidopsis thaliana;* Le=*Lycopersicon esculentum;* Mm=*Mus musculus;* Gg=*Gallus gallus;* Ce=*Caenorhabditis elegans;* Bt=*Bos Taurus;* PPI=Peptidyl-Prolyl Isomerase; FK506BP=FK506 Binding Protein; EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.



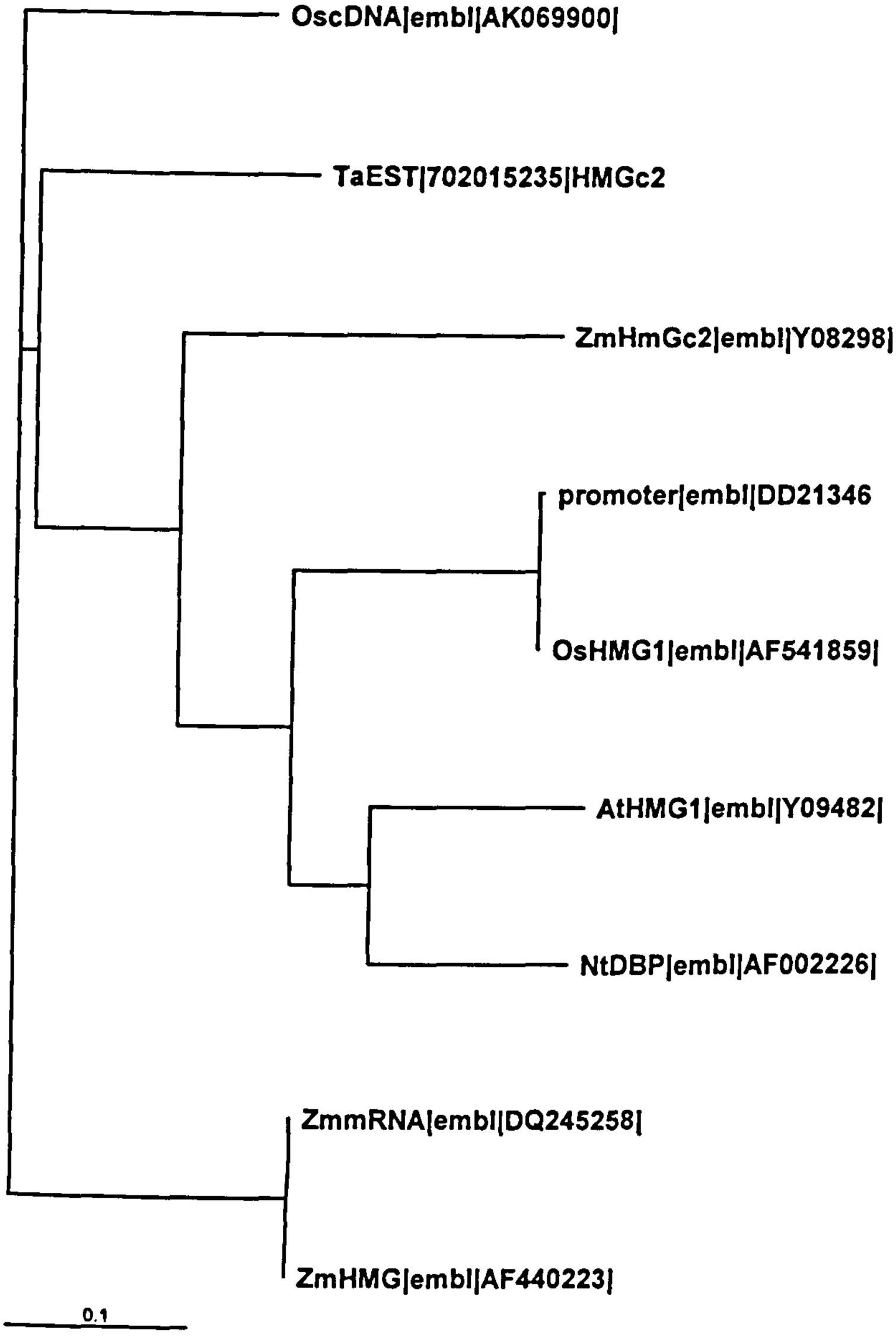
Phylogenetic tree for TaEST|701997623|ERTCoa produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* Rr=*Retama raetam;* Bt=*Bos Taurus;* Xt=Xenopus Tropicalis; ERTCoA=ethylene-responsive transciptional coactivator-like; EDRF=*Endothelial differentiation-related factor;* EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.



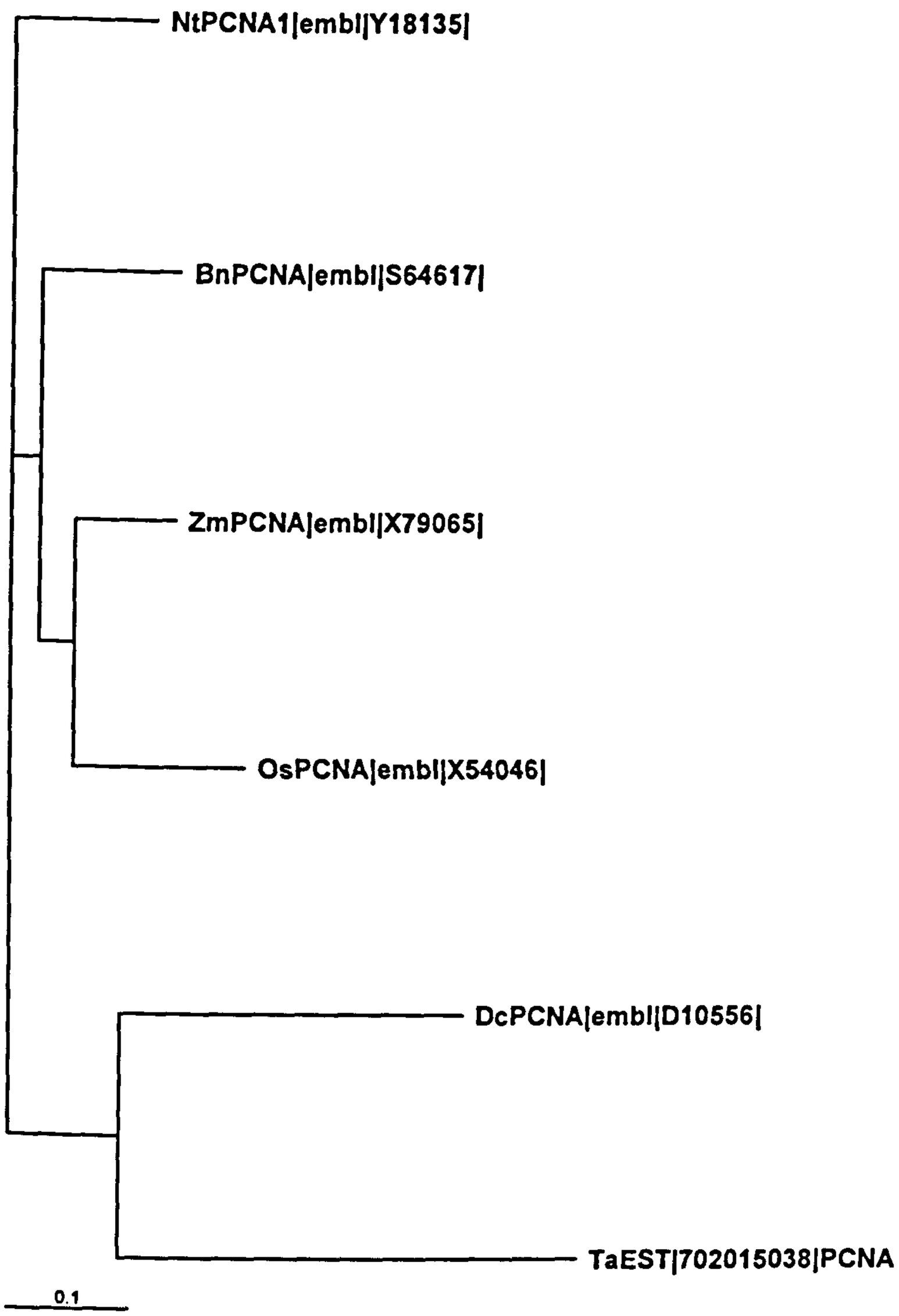
Phylogenetic tree for TaEST|702039485|LTP produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Sv=*Sorghum vulgare;* Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays*; LTP=Lipid Transfer Protein; PLT=Phospholipid Transfer Protein; EST=Expressed sequence tag. Scale bar indicates the number of nucleotide replacements per site.



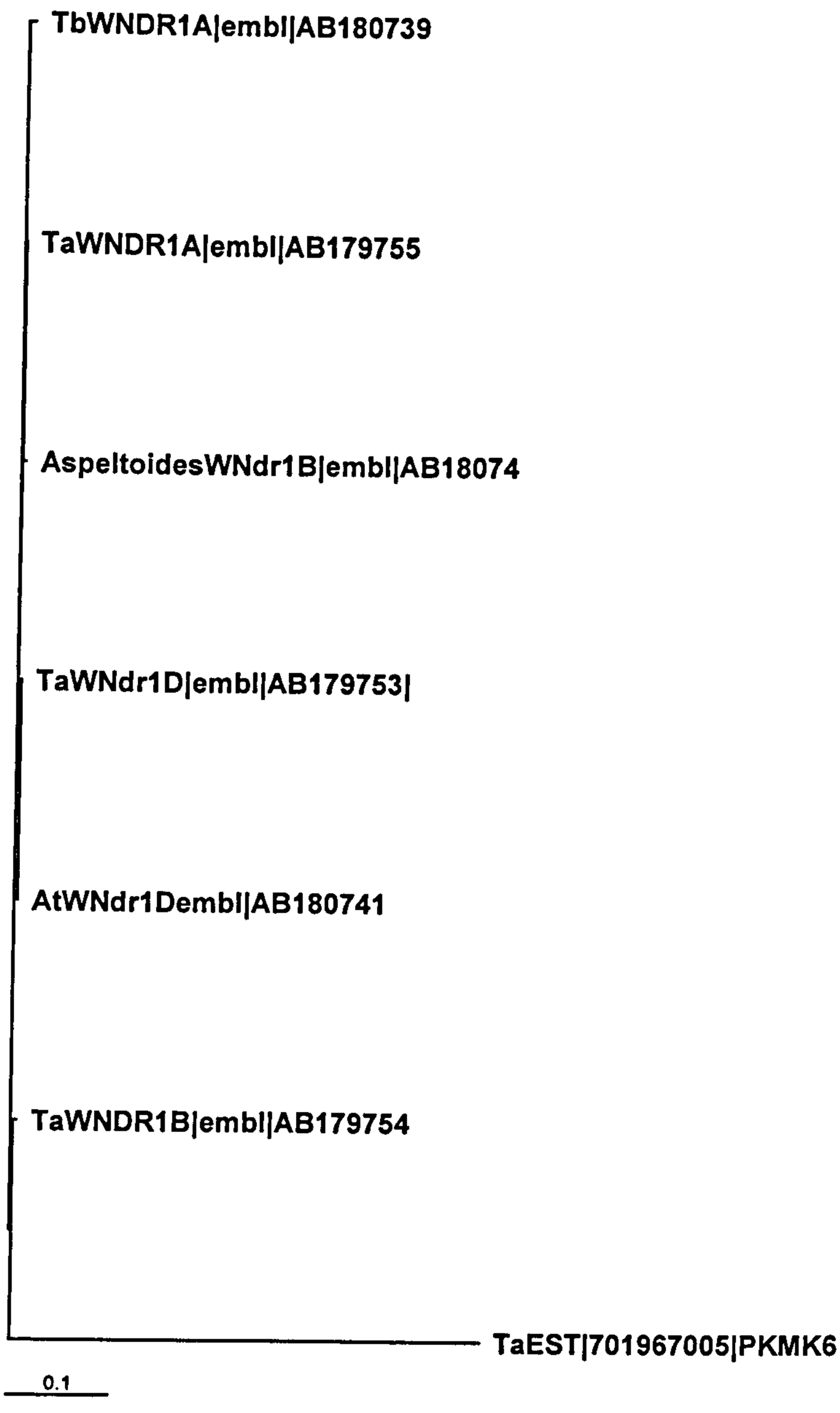
Phylogenetic tree for TaEST|701960043|NADPH produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* Hv=*Hordeum vulgare;* Gm=*Glycine max*; Ps=*Papaver somniferum*; NADPH= nicotinamide adenine dinucleotide phosphate; DMAS1=deoxymugineic acid synthase 1; EST=Expressed sequence tag; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.



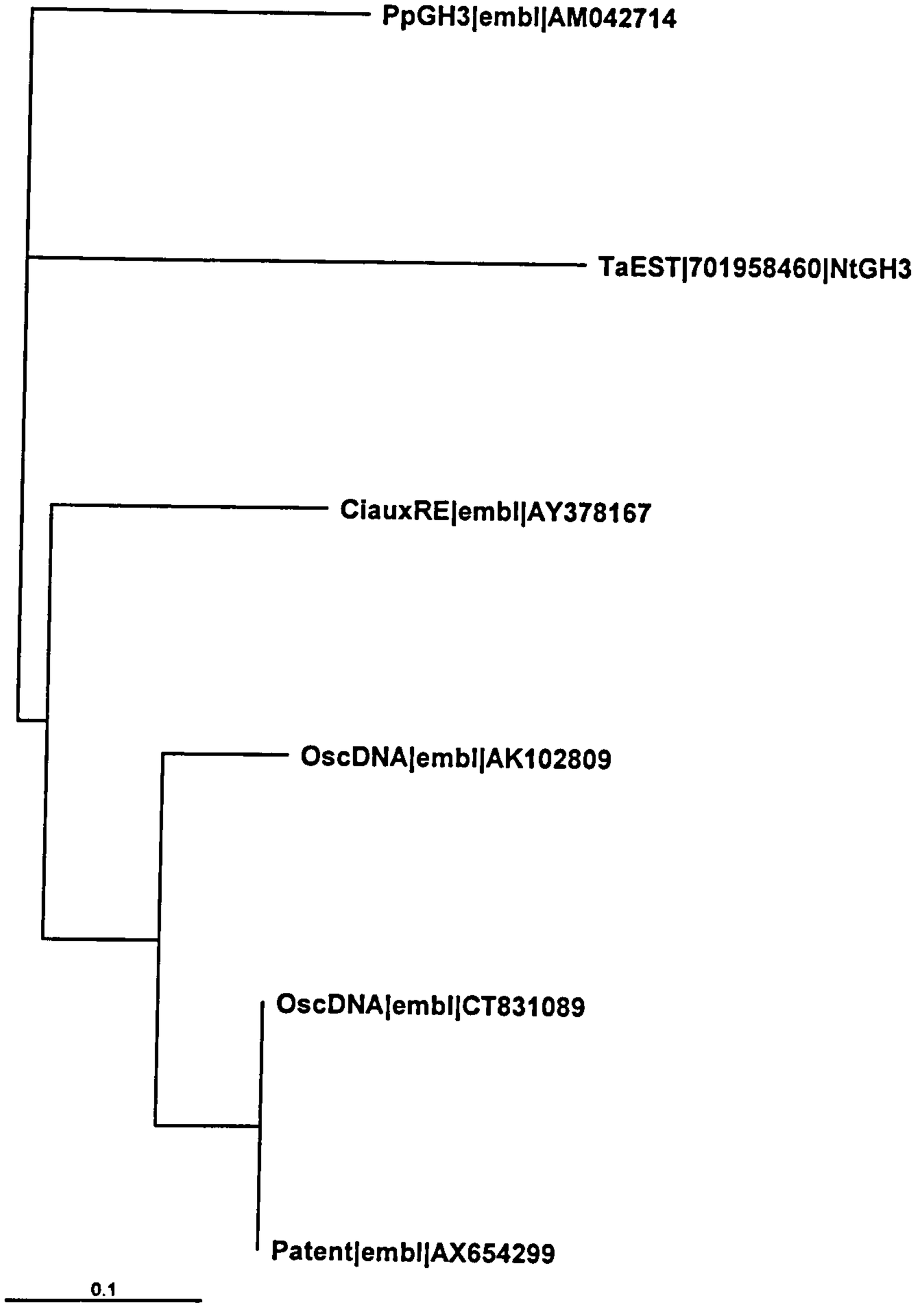
Phylogenetic tree for TaEST|702015235|HMGc2 produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* At=*Arabidopsis thaliana*; Nt=*Nicotiana tabacum;* HMG= High mobility group protein; DBP=DNA Binding Protein; EST=Expressed sequence tag; cDNA=cDNA of no known function; mRNA=mRNA of no known function. Scale bar indicates the number of nucleotide replacements per site.



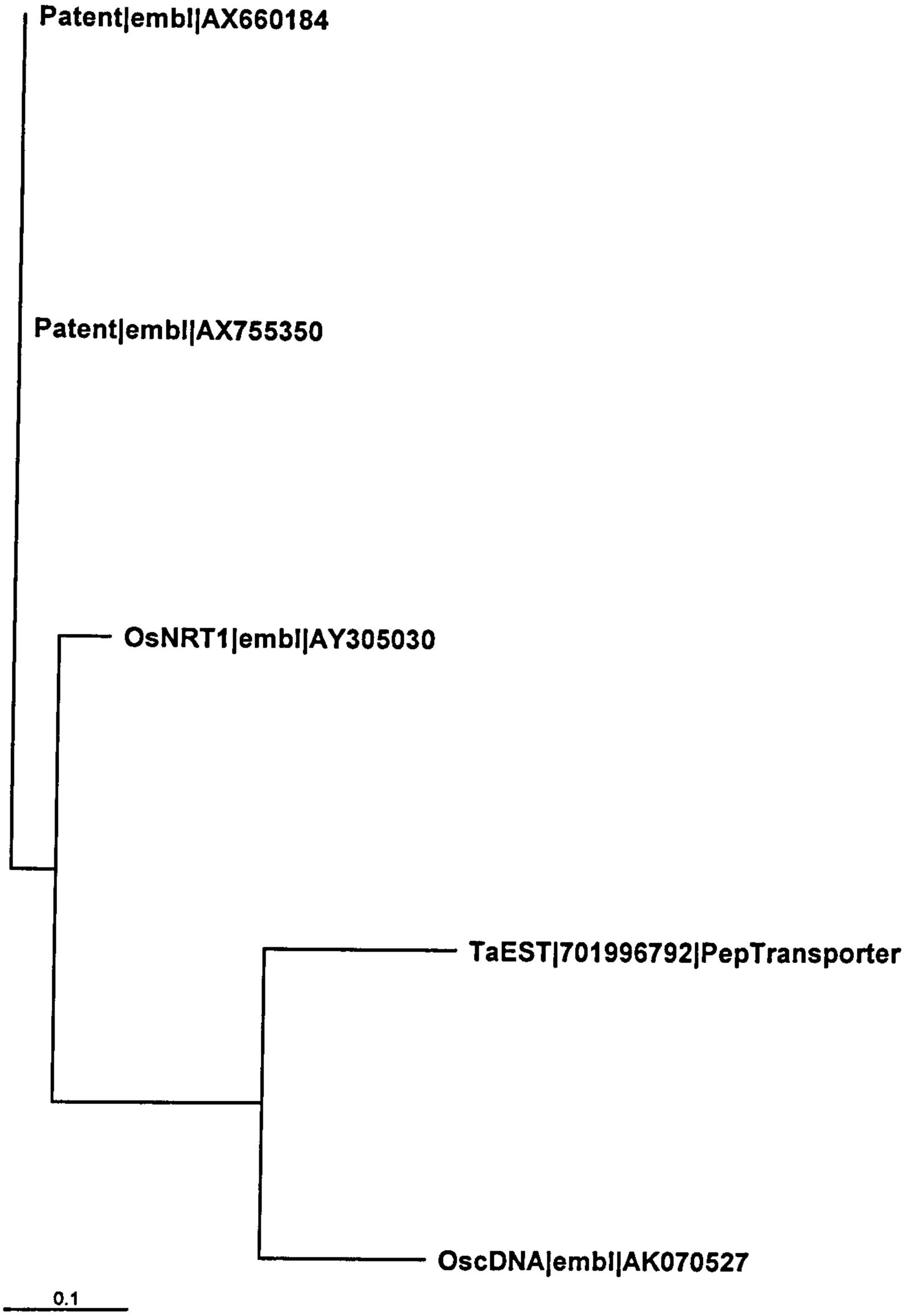
Phylogenetic tree for TaEST|702015038|PCNA produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* Dc=*Daucus carota;* Bn=*Brassica napus;* Nt=*Nicotiana tabacum;* PCNA= Proliferating cell nuclear antigen; EST=Expressed sequence tag. Scale bar indicates the number of nucleotide replacements per site.



Phylogenetic tree for TaEST|701967005|PKMK6 produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Tb=*Triticum baeoticum;* At=*Arabidopsis thaliana;* Aspeltoides=*Aeogilops speltoides;* PKMK6=Protein Kinase MK6; WNDR=NDR Protein kinase; EST=Expressed sequence tag. Scale bar indicates the number of nucleotide replacements per site.

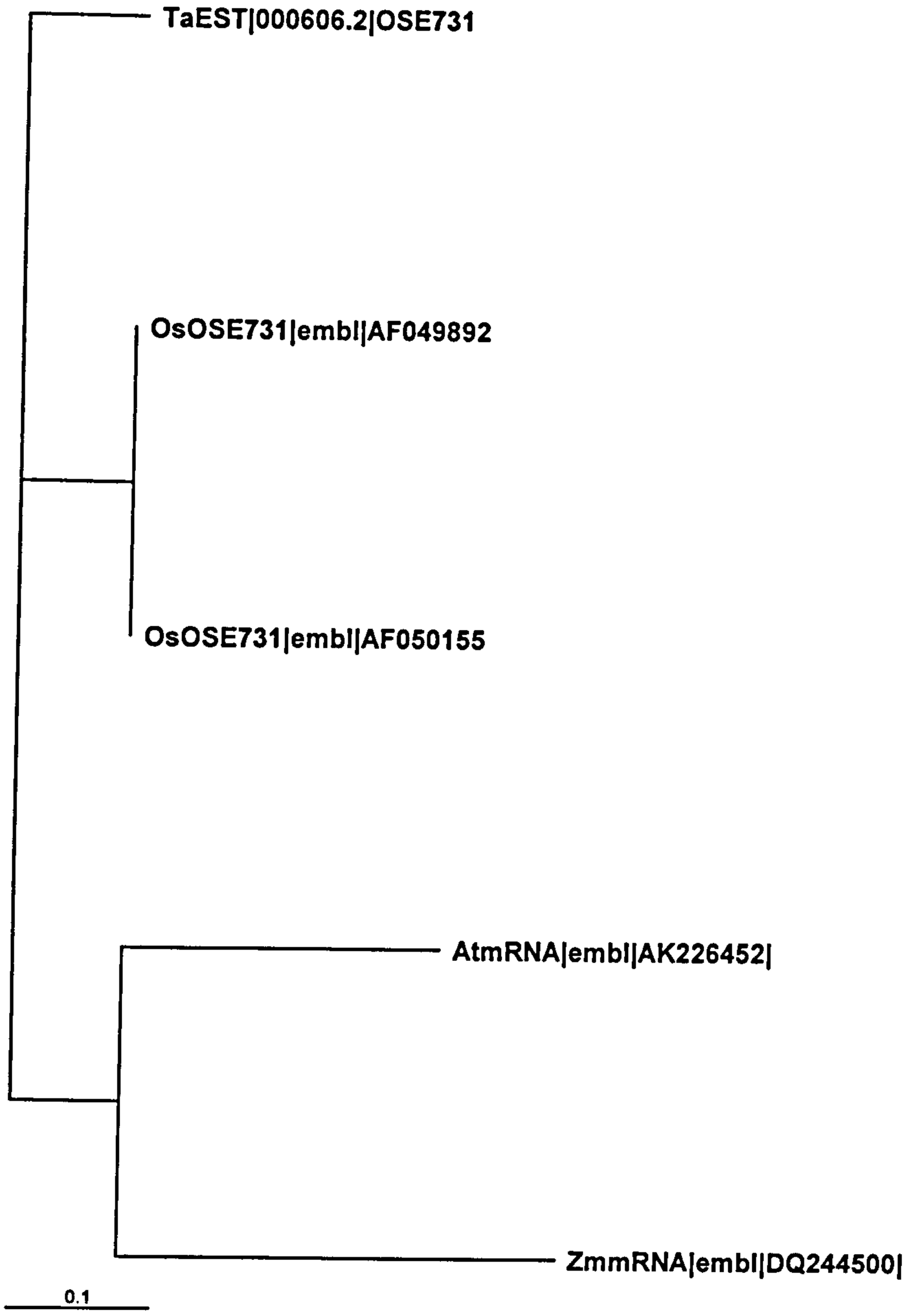


Phylogenetic tree for TaEST|701958460|NtGH3 produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Ci=*Cichorium intybus;* Pp=Pinus pinistar; Nt=*Nicotiana tabacum;* GH3=Auxin responsive GH3 protein; AuxRE=Auxin responsive element; cDNA=cDNA of no known function EST=Expressed sequence tag. Scale bar indicates the number of nucleotide replacements per site.

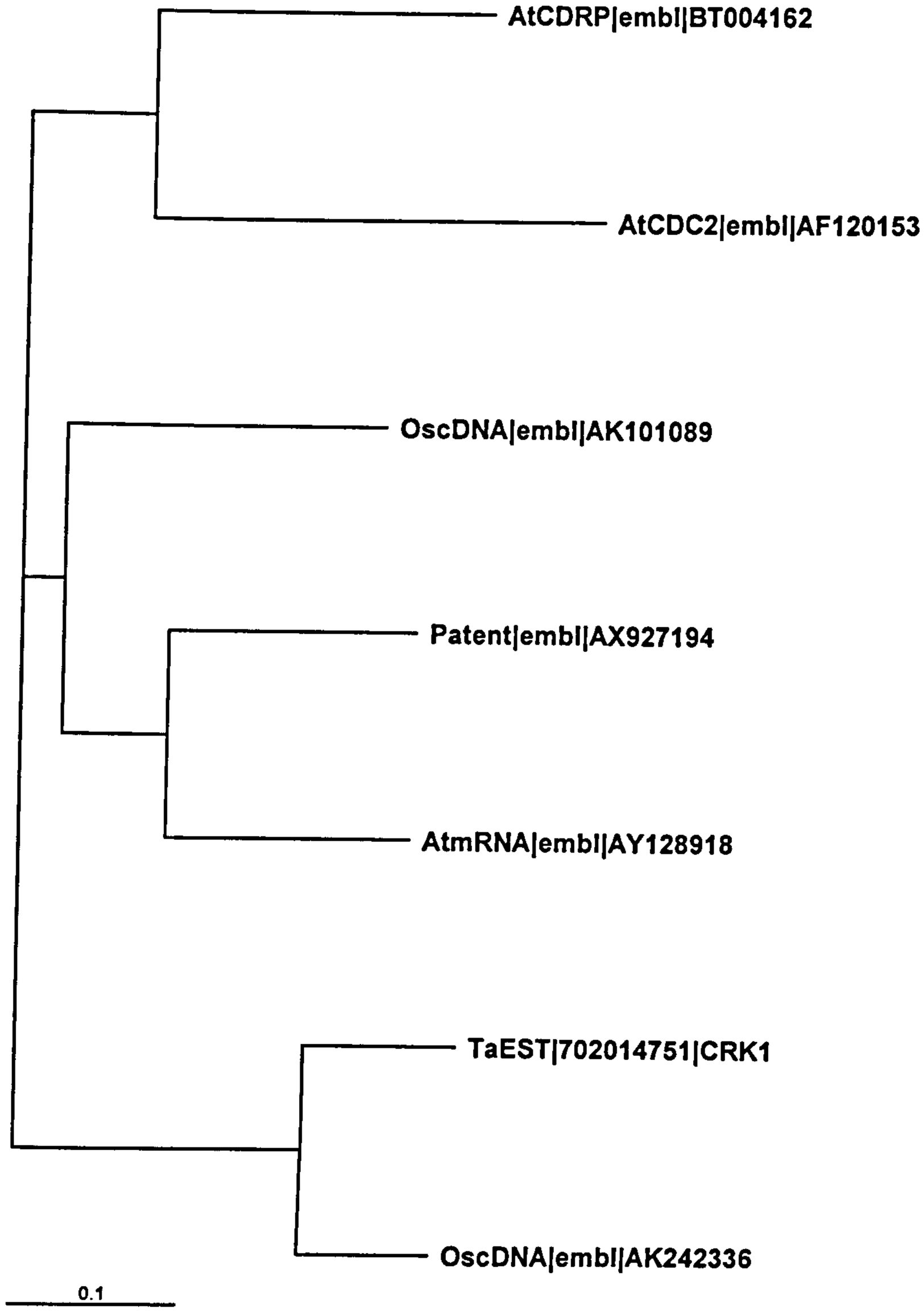


Phylogenetic tree for TaEST|701996792|PepTransporter produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* NRT1=Nitrate Transporter 1; EST=Expressed sequence tag; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

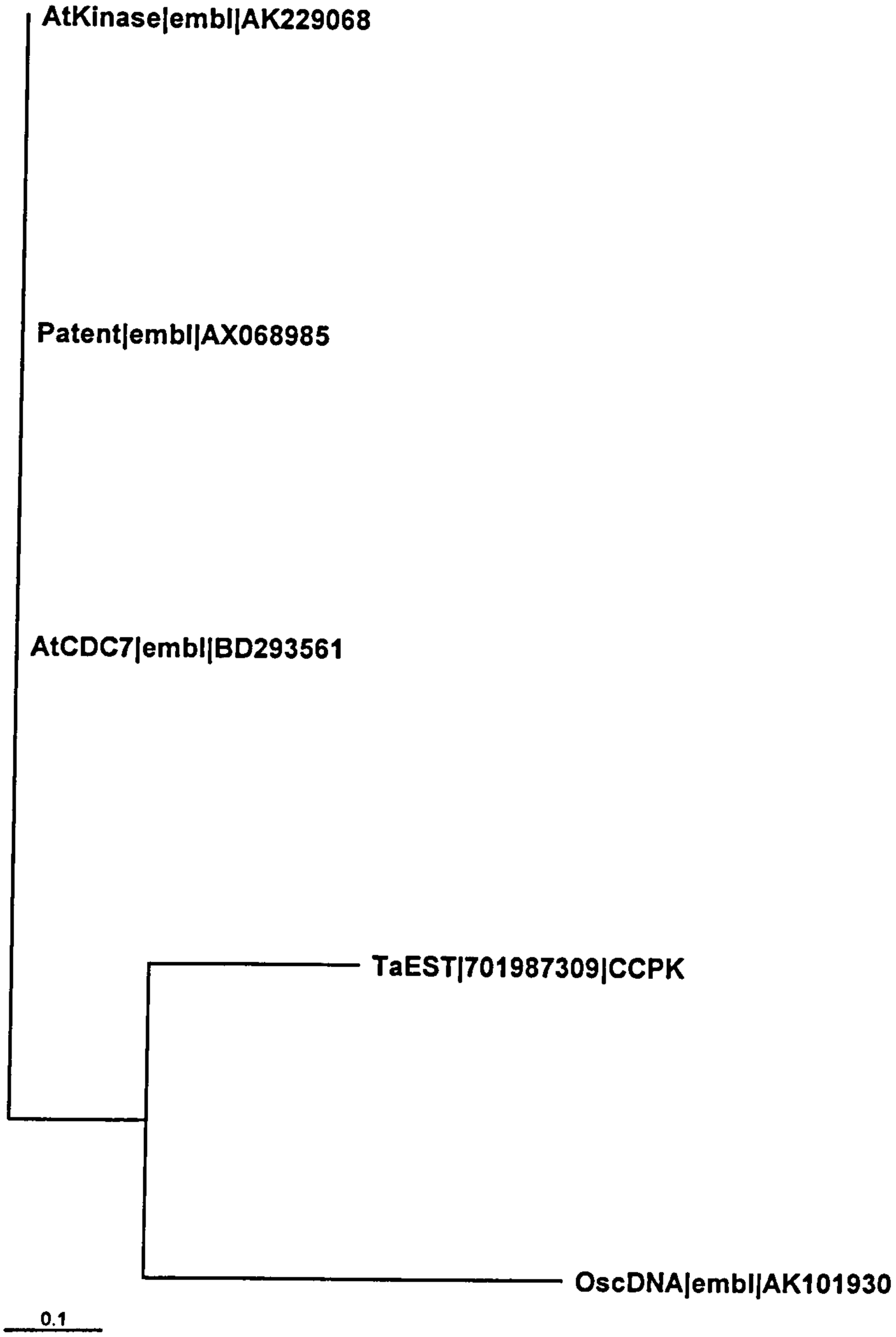
Phylogenetic tree for TaEST|015920.4|SCR produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* At=*Arabidopsis thaliana;* Le=*Lycopersicon esculentum;* SCR=Scarecrow; SCL=Scarecrow-like; GRAS= GAI, RGA, SCR (GRAS) family; EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.



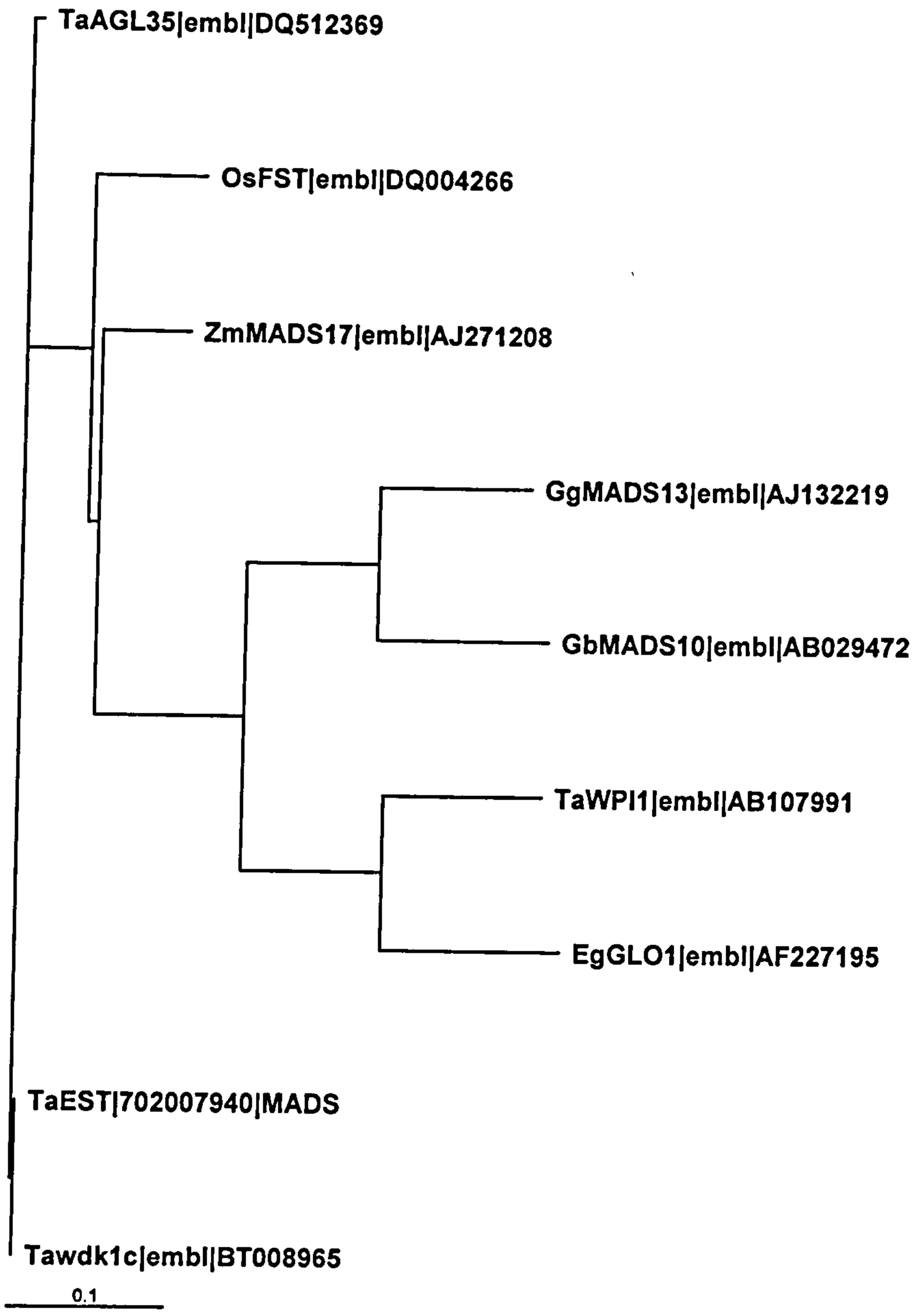
Phylogenetic tree for TaEST|000606.2|OSE731 produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* At=*Arabidopsis thaliana;* OSE731=*Oryza sativa* embryo specific 731 protein; EST=Expressed sequence tag; mRNA=mRNA of no known function. Scale bar indicates the number of nucleotide replacements per site.



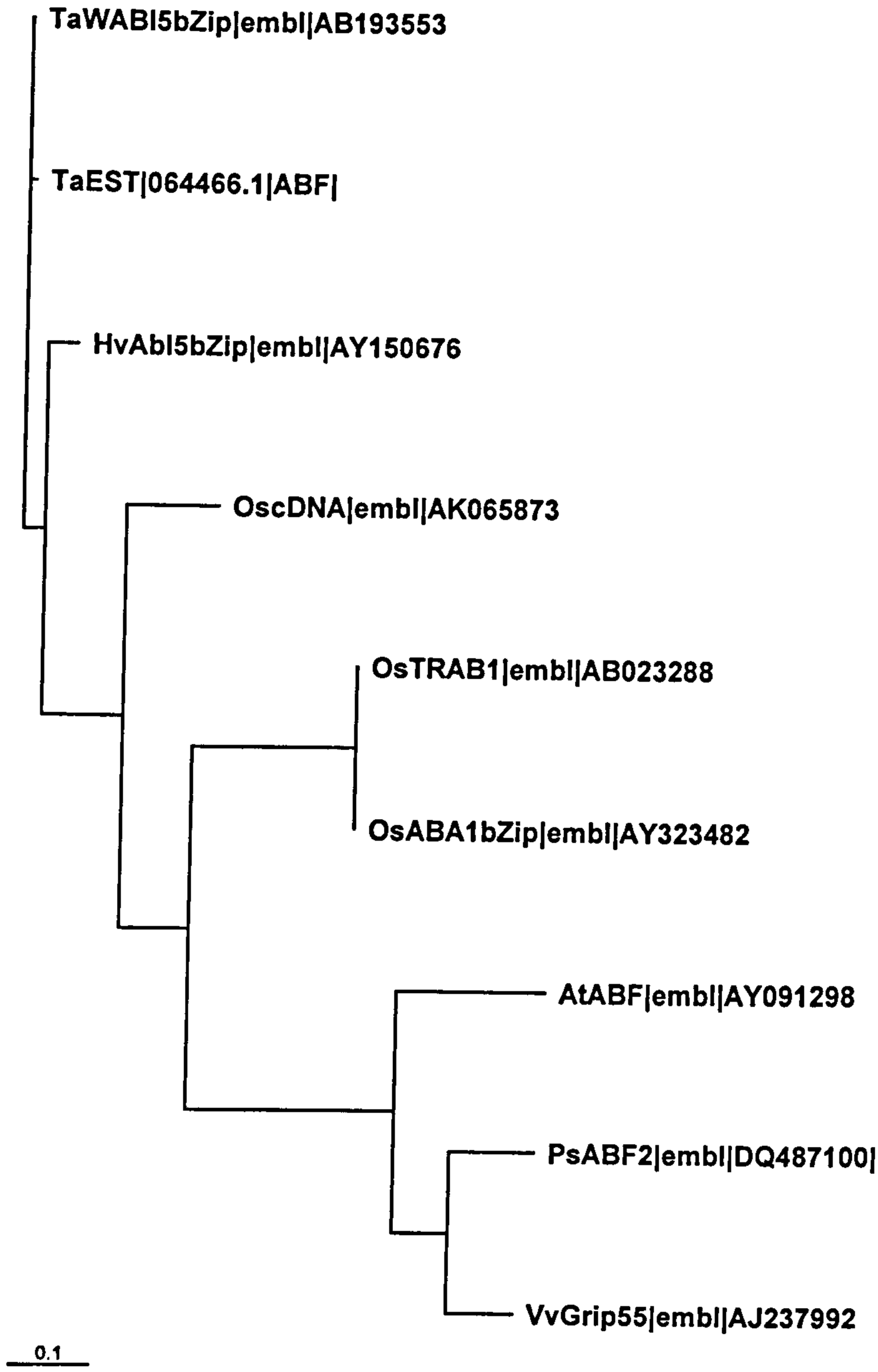
Phylogenetic tree for TaEST|702014751|CRK1 produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* At=*Arabidopsis thaliana;* CRK1= Cdk Related Kinase 1; CDRP=Cell Division Related Protein; CDC2=Cell Division Cycle 2; EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.



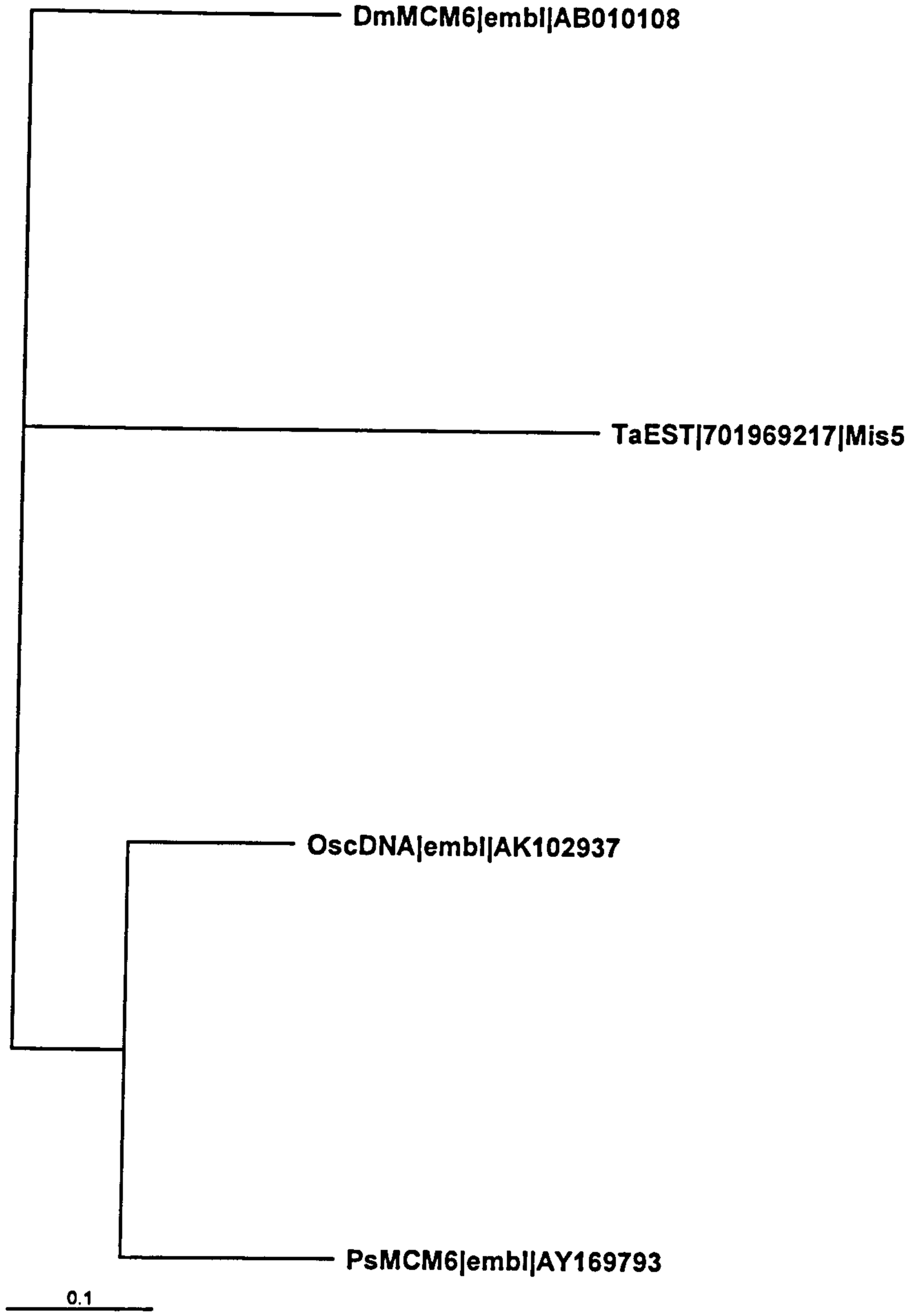
Phylogenetic tree for TaEST|701987309|CCPK produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* At=*Arabidopsis thaliana;* CCPK=Cell Cycle Protein Kinase; CDC7=Cell Division Cycle 7; Kinase=Kinase-like protein; EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.



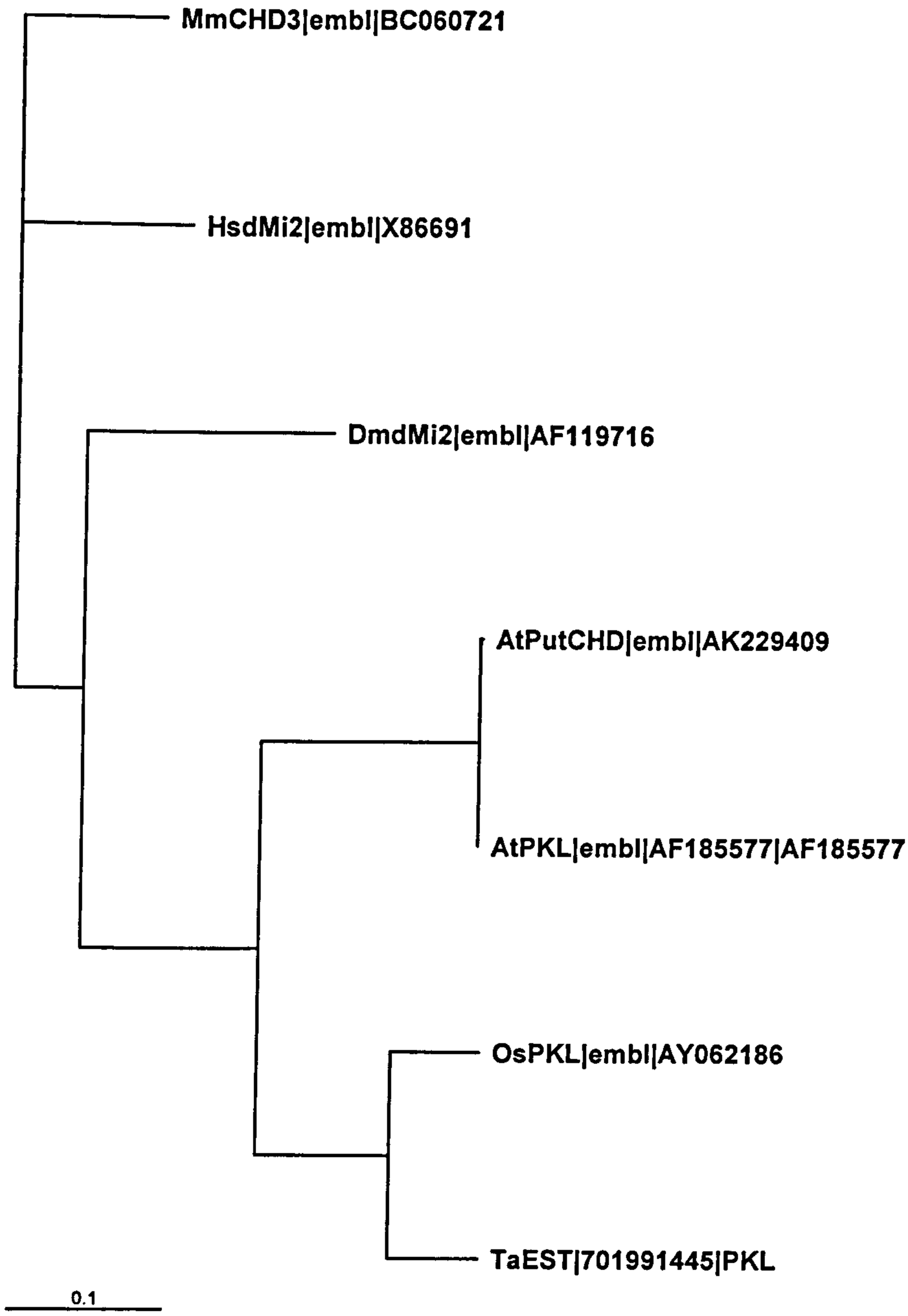
Phylogenetic tree for TaEST|702007940|MADS produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* Eg=*Elaeis guineensis*; Gg=*Gnetum gnemon*; Gb=*Gingko biloba*; MADS= MCM1, AGAMOUS, DEFICIENS and SRF (MADS) family; AGL= Agamous-like; FST=Female Sterility; WPI=Wheat Pistillata1; WDK1c=Wheat Developing Kernel 1c; GLO1= Globby 1; EST=Expressed sequence tag. Scale bar indicates the number of nucleotide replacements per site.



Phylogenetic tree for TaEST|064466.1|ABF produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Hv=*Hordeum vulgare;* At=*Arabidopsis thaliana;* Ps=*Populus suaveolens*; Vv= *Vitis vinifera*; ABI=ABSCISIC ACID-INSENSITIVE; TRAB=ABA-regulated transcription factor; ABA=ABA deficient; ABF=abscisic acid responsive elements-binding factor; Grip=ripening-related bZIP protein; EST=Expressed sequence tag; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

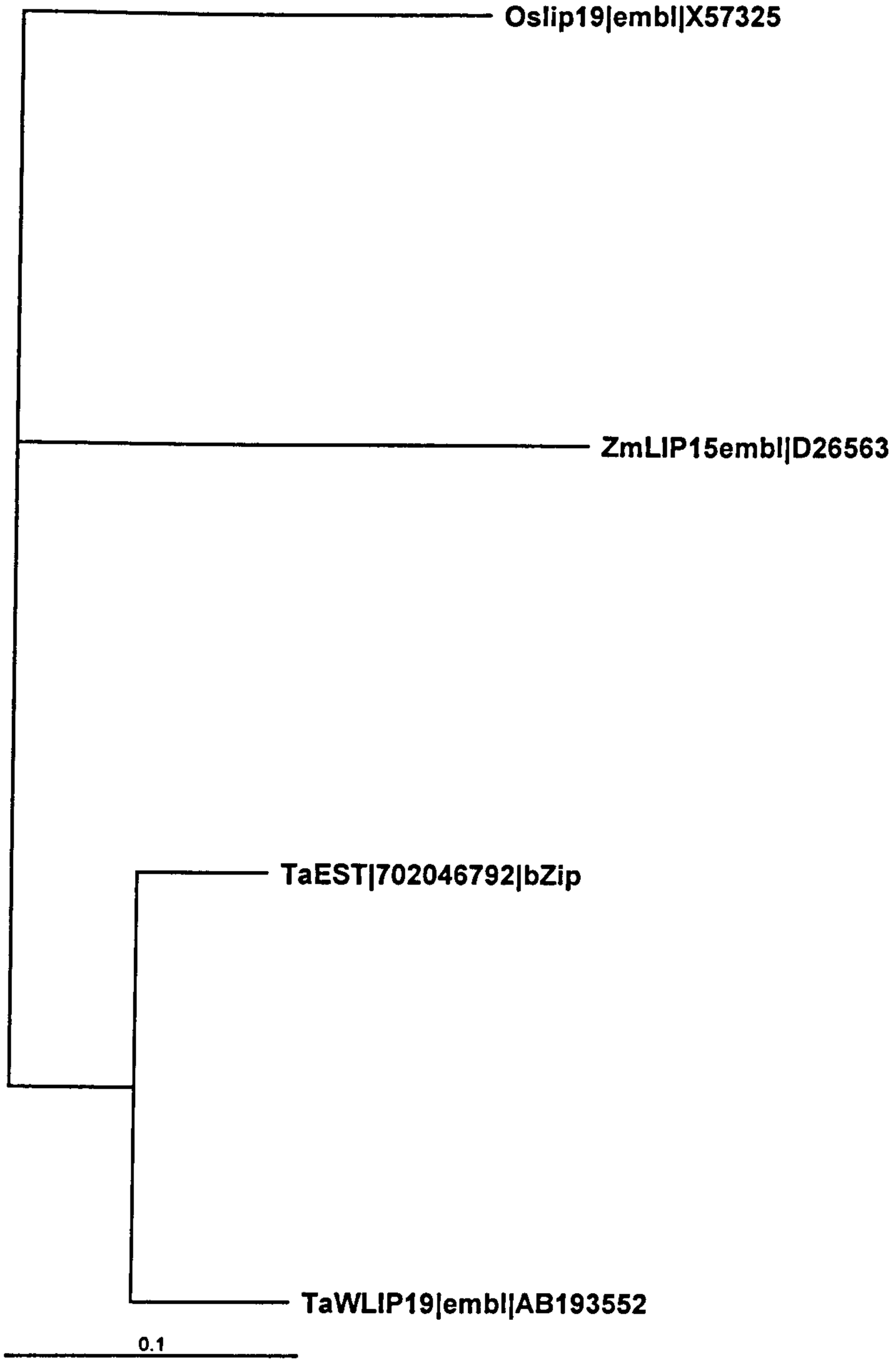


Phylogenetic tree for TaEST|701969217|Mis5 produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Dm=*Drosophila melanogaster;* Ps=*Pisum sativum*; MCM6=Mini-chromosome maintenance protein; Mis5=MCM6 homolog; EST=Expressed sequence tag; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

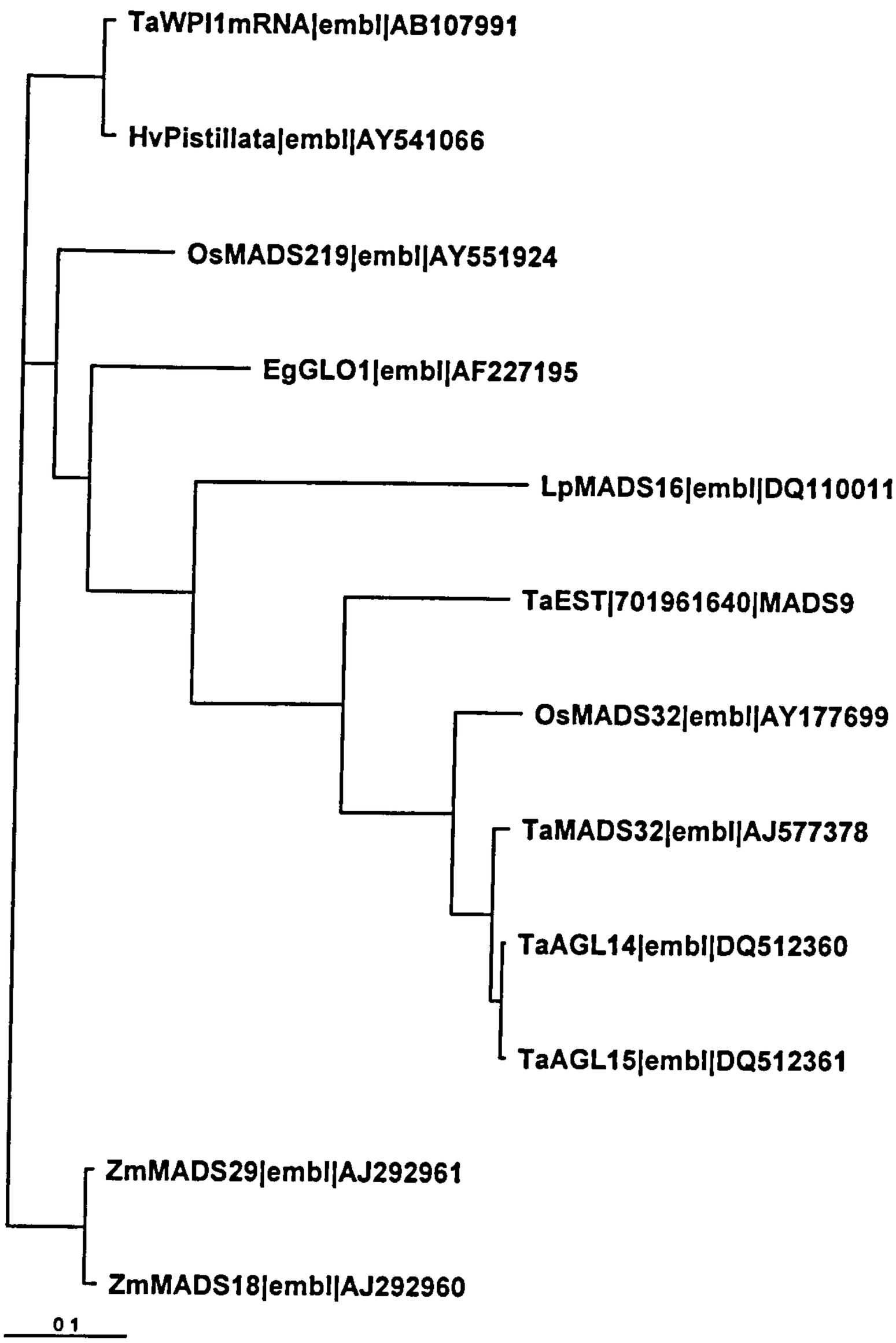


Phylogenetic tree for TaEST|701991445|PKL produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* At=*Arabidopsis thaliana;* Mm=*Mus musculus;* Hs=*Homo Sapiens;* Dm=*Drosophila melanogaster;* PKL=Pickle; CHD3=Chromodomain Helicase DNA binding protein 3; dmi2= Mi2 helicase protein; PutCHD=putative Chromodomain Helicase DNA binding protein; EST=Expressed sequence tag. Scale bar indicates the number of nucleotide replacements per site.

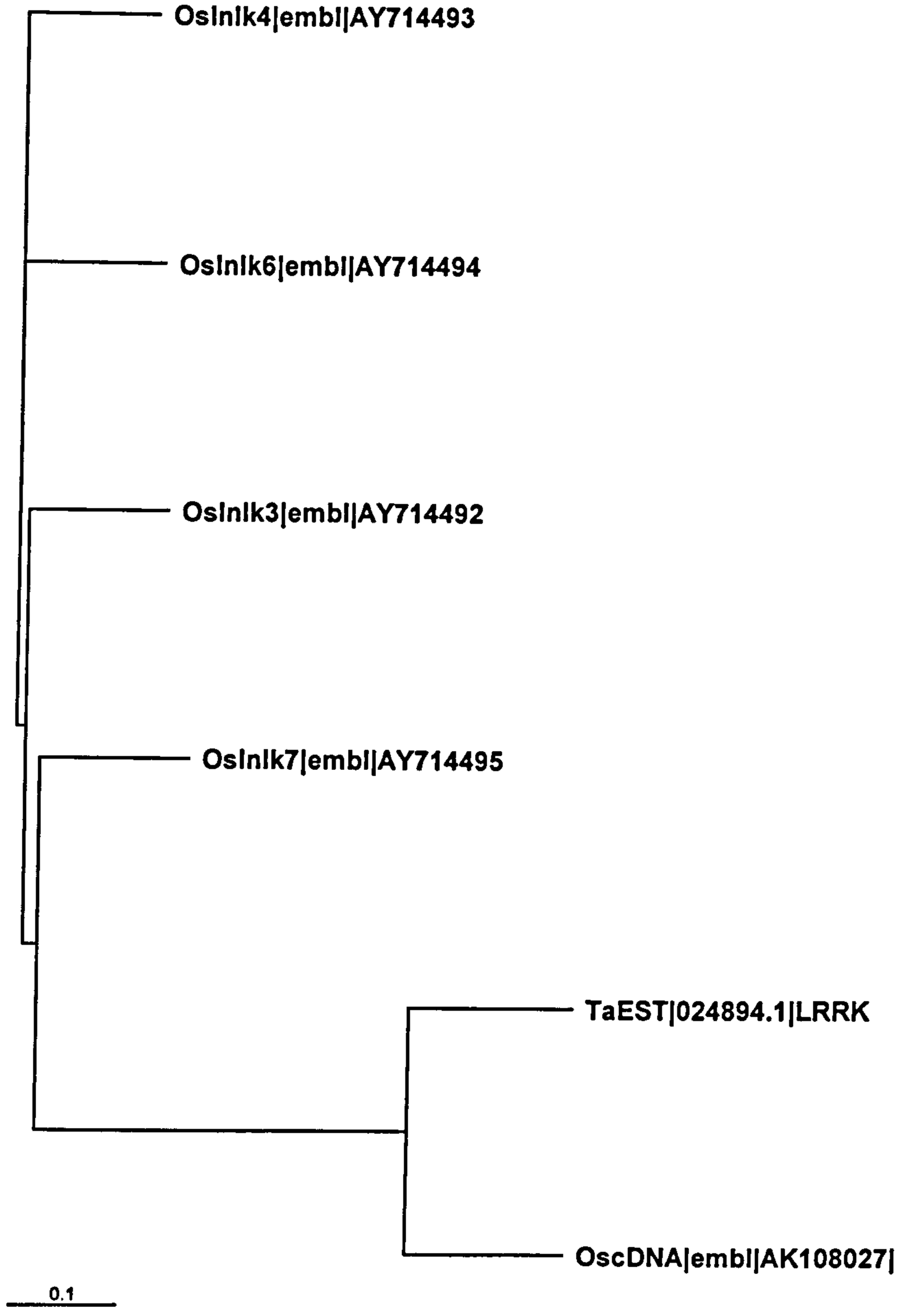
Phylogenetic tree for TaEST|022121.3|DgSERLTD produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* Dg=*Dactylis glomerata*; SERLTD=Somatic Embryogenesis Related protein; EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.



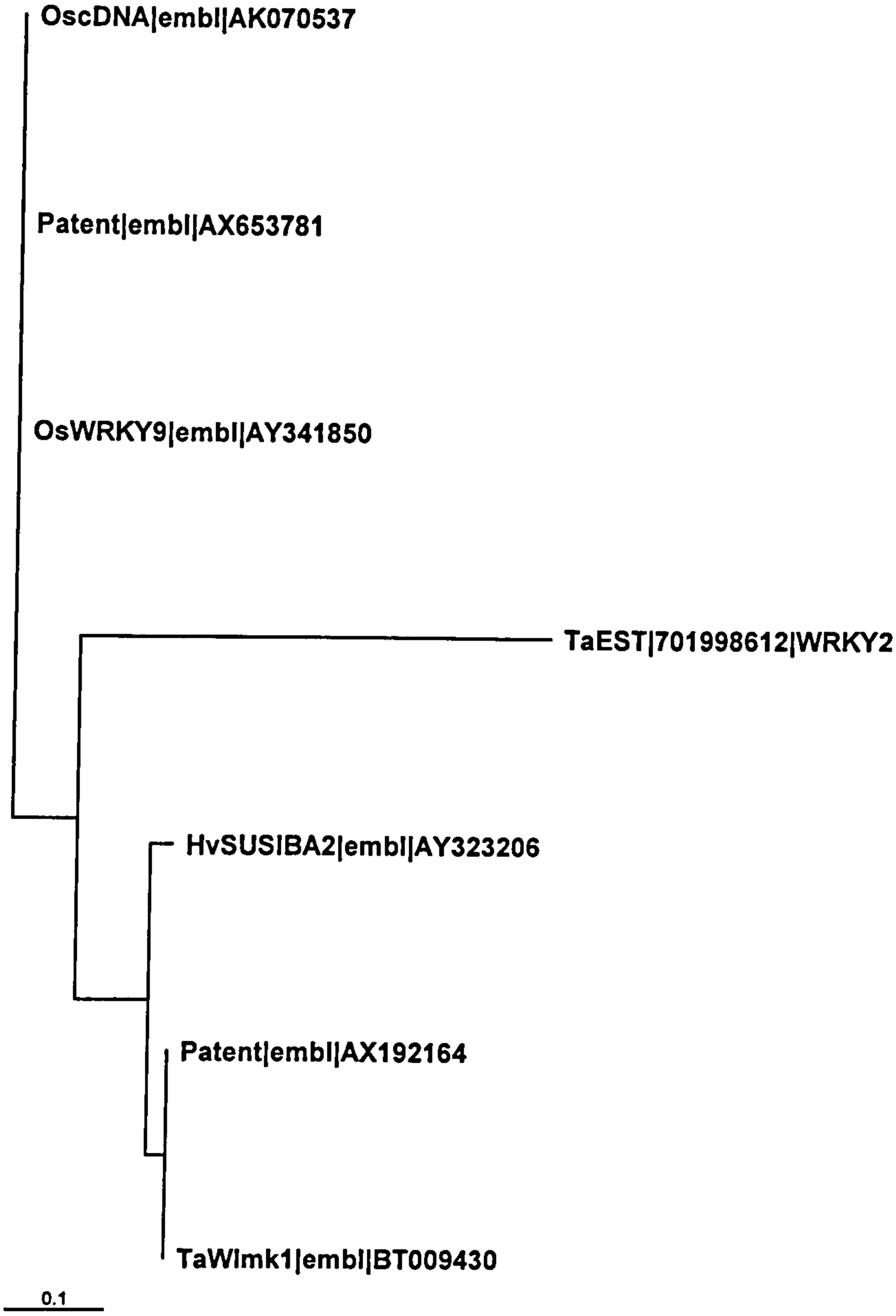
Phylogenetic tree for TaEST|702046792|bZip produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* LIP=basic/leucine zipper protein; bZip= basic/leucine zipper protein; EST=Expressed sequence tag. Scale bar indicates the number of nucleotide replacements per site.



Phylogenetic tree for TaEST|701961640|MADS9 produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* Hv=*Hordeum vulgare;* Eg=*Elaeis guineensis*; Lp=*Lolium perenne;* MADS= MCM1, AGAMOUS, DEFICIENS and SRF (MADS) family; AGL= Agamous-like; WPI=Wheat Pistillata 1; GLO1=Globby1; EST=Expressed sequence tag. Scale bar indicates the number of nucleotide replacements per site.

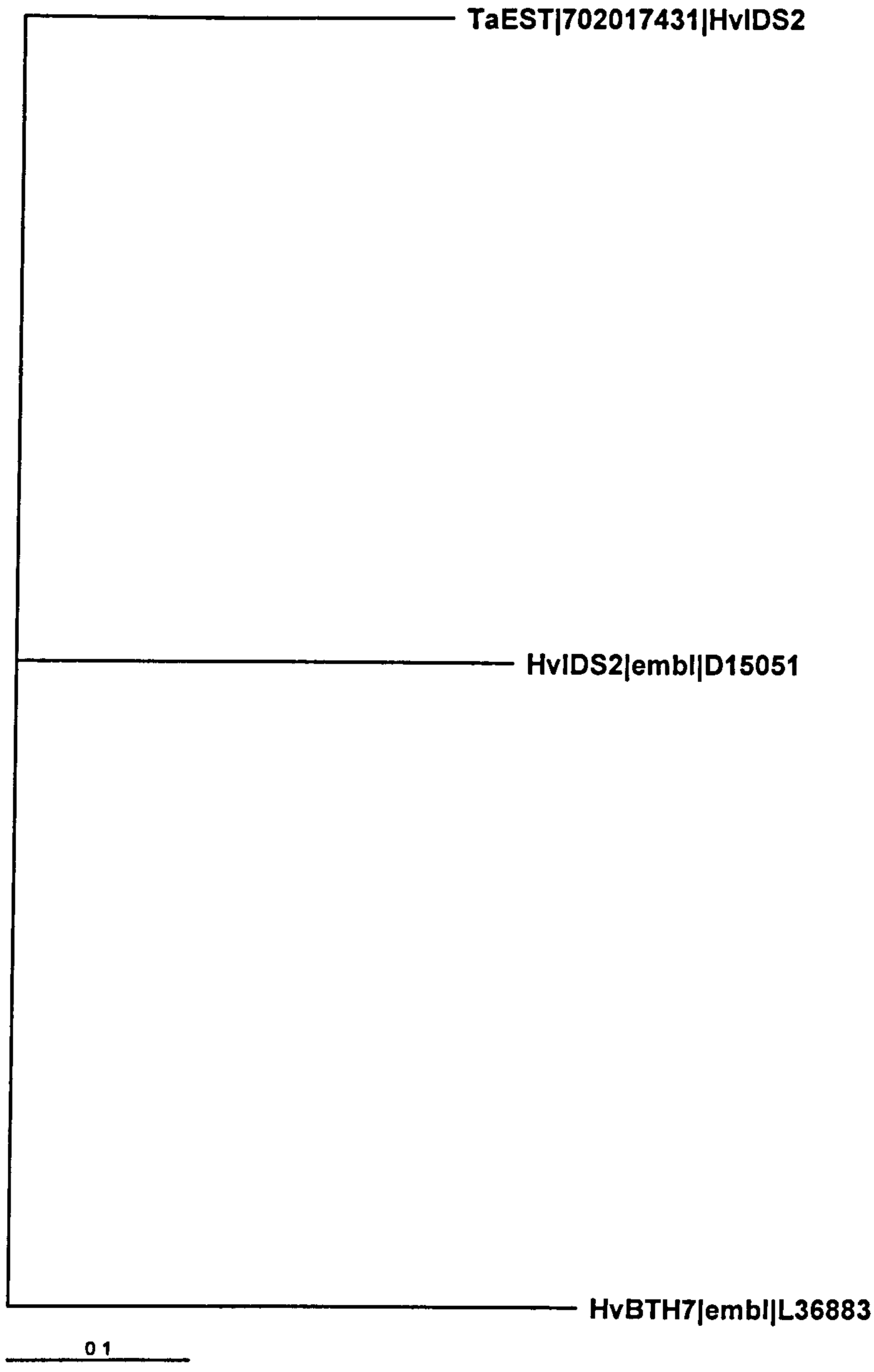


Phylogenetic tree for TaEST|024894.1|LRRK produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* LRRK=Leucine Rich Receptor Kinase; Inlk=putative leucine-rich repeat receptor-like kinase; EST=Expressed sequence tag; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.



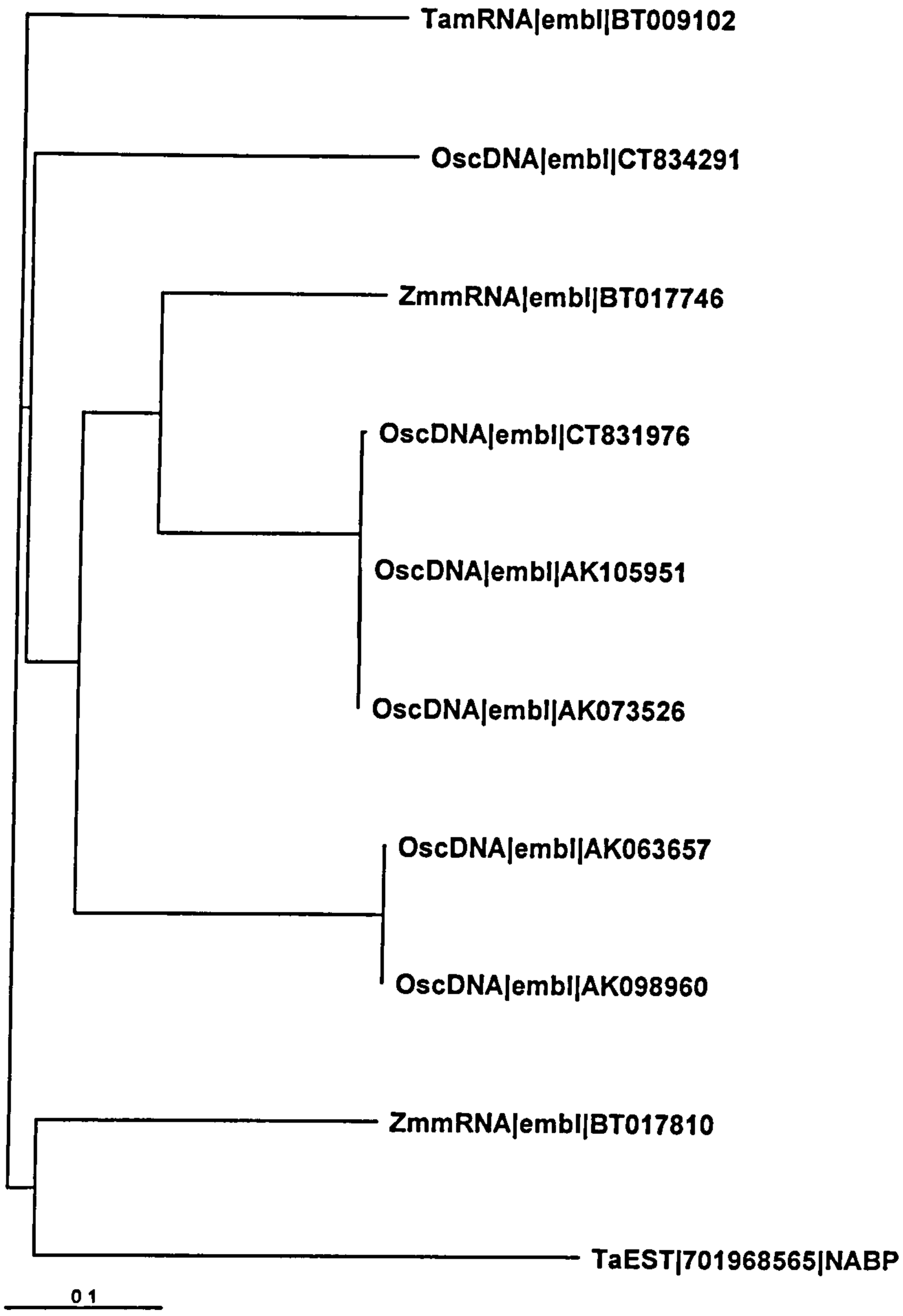
Phylogenetic tree for TaEST|701998612|WRKY2 produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Hv=*Hordeum vulgare;* WRKY=WRKY DNA binding protein; SUSIBA=SUSIBA WRKY Transcription Factor; Wlmk1=wheat clone of unknown function; EST=Expressed sequence tag; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

Phylogenetic tree for TaEST|701991275|OsSCAMP produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* Le=*Lycopersicon esculentum;* Ps=*Pisum sativum;* SCAMP=Secretory Carrier Membrane Protein; EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

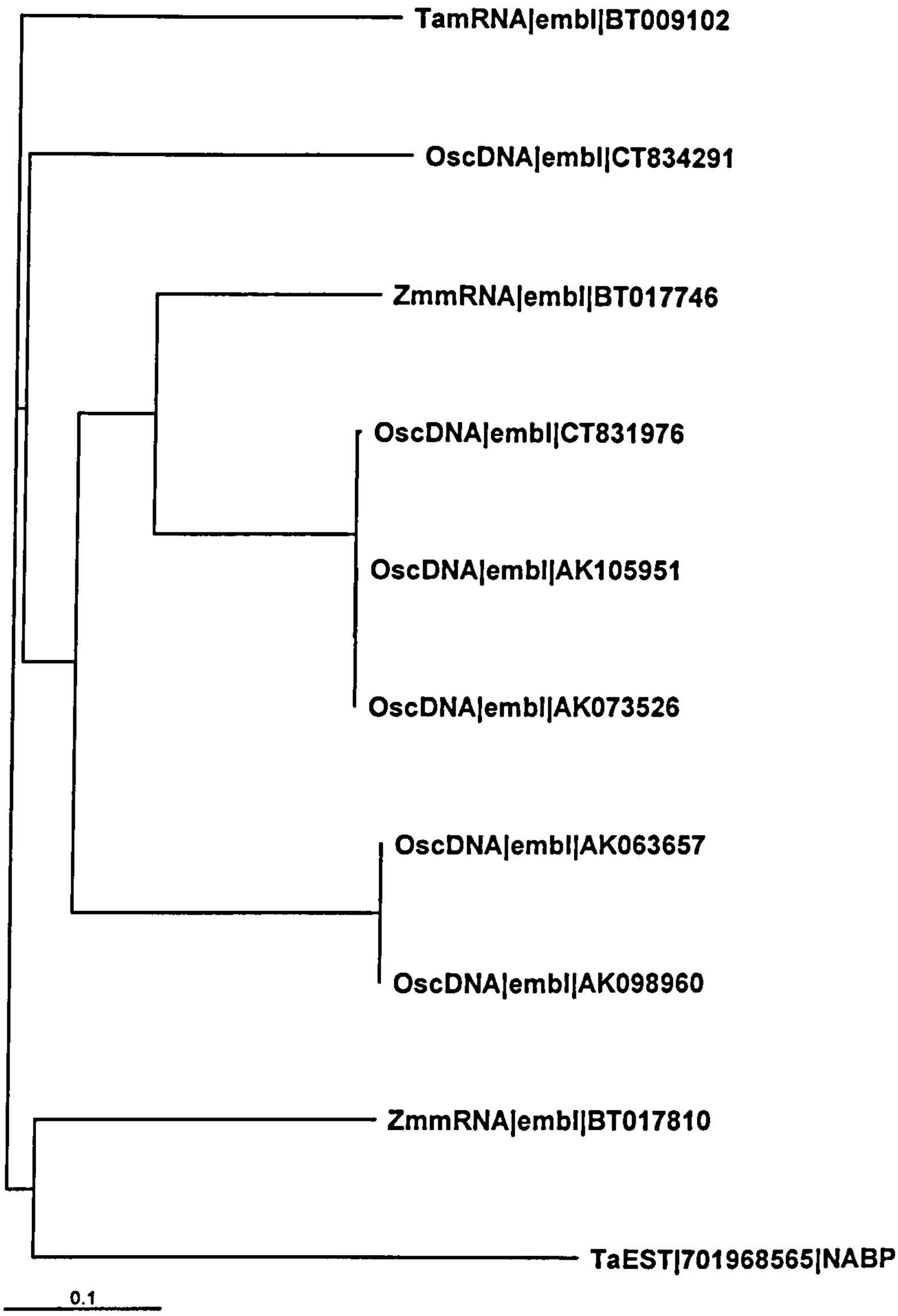


Phylogenetic tree for TaEST|702017431|HvIDS2 produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Hv=*Hordeum vulgare;* IDS2=Dioxygenase ids2; BTH=Beta Thionin; EST=Expressed sequence tag. Scale bar indicates the number of nucleotide replacements per site.

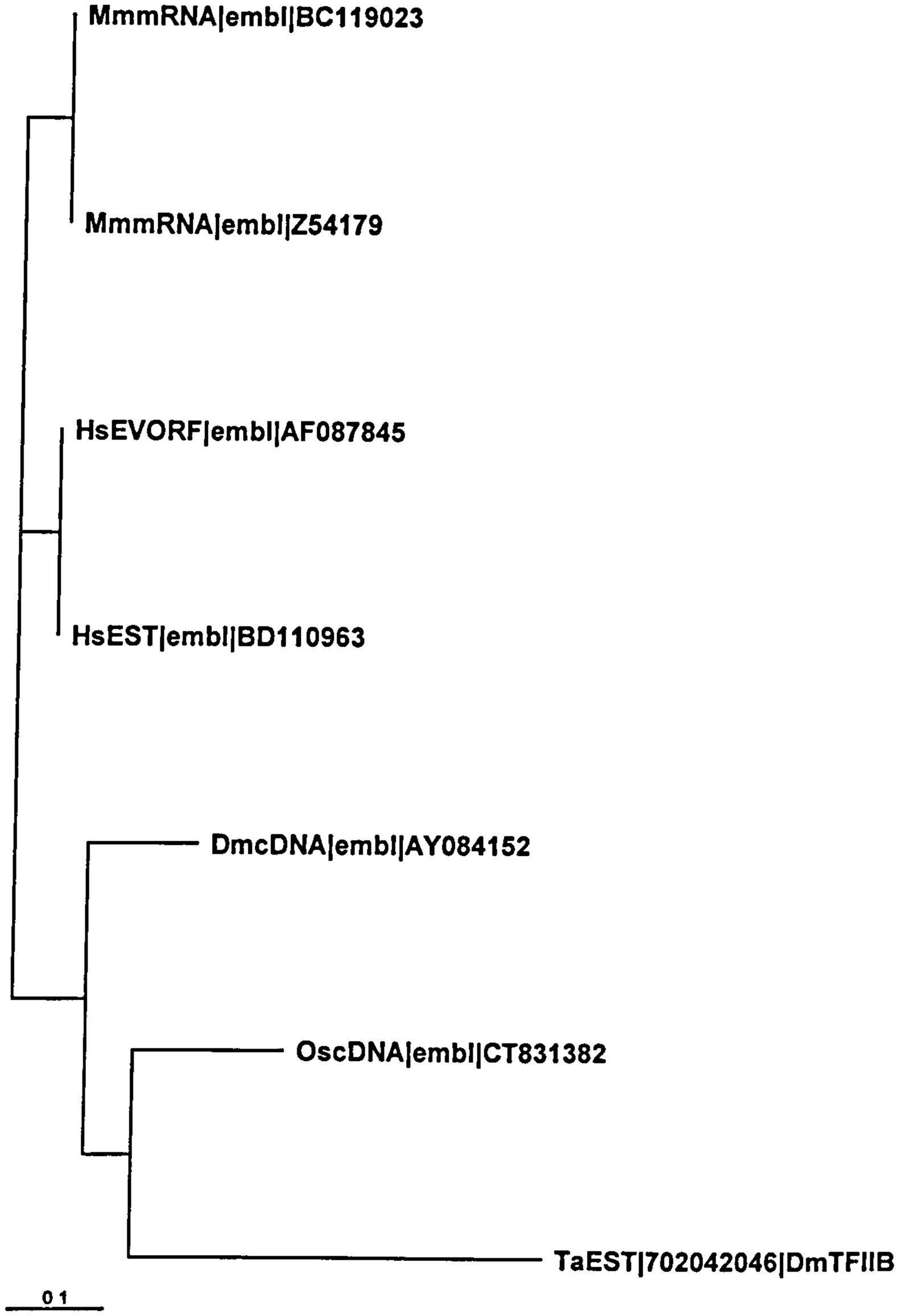
Phylogenetic Trees With No Inferred Functional Annotation



Phylogenetic tree for TaEST|701968565|NABP produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* NABP=Nucleic Acid Binding Protein; EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

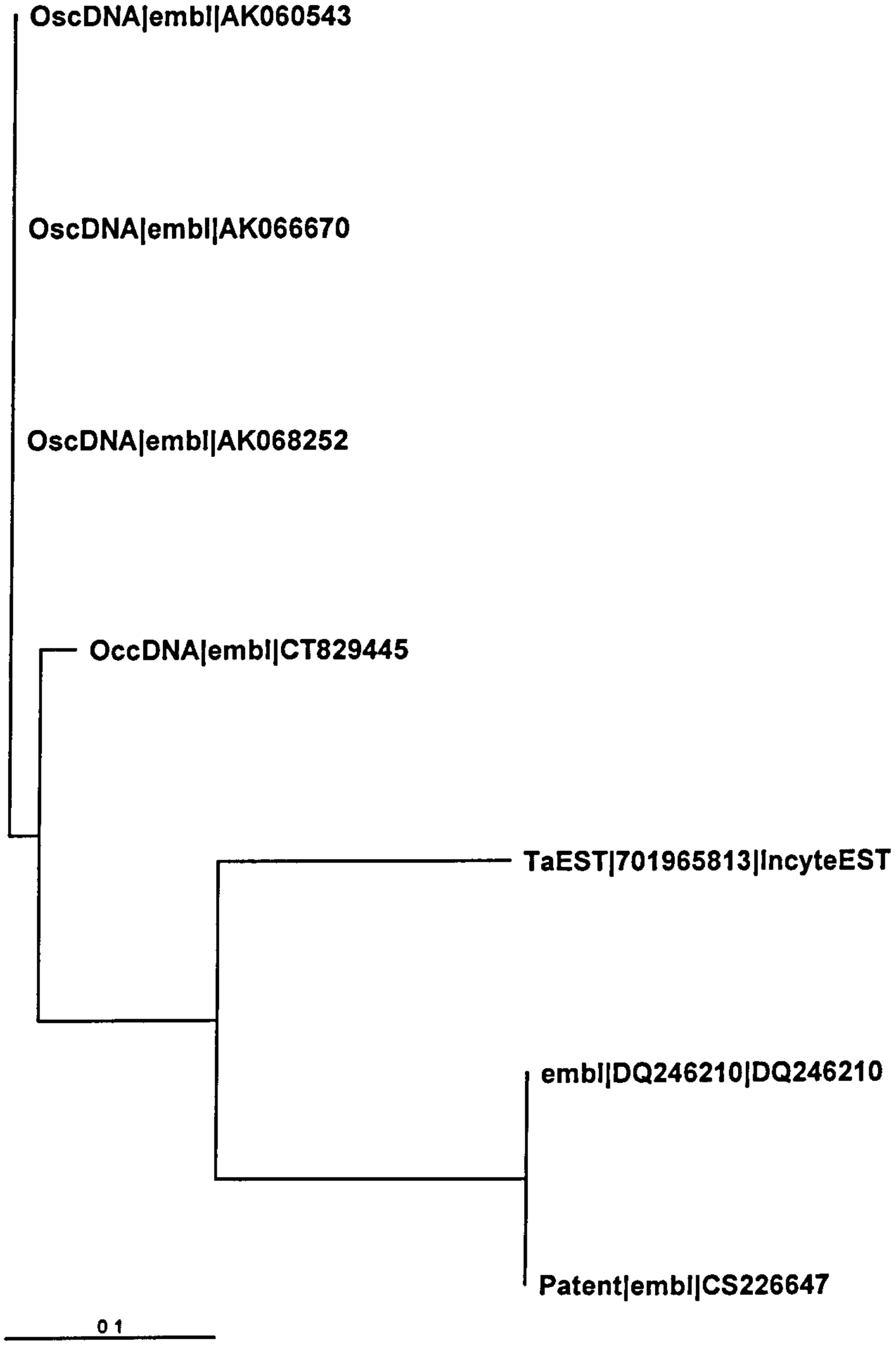


Phylogenetic tree for TaEST|701968565|NABP produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Zm=*Zea mays;* NABP=Nucleic Acid Binding Protein; EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.



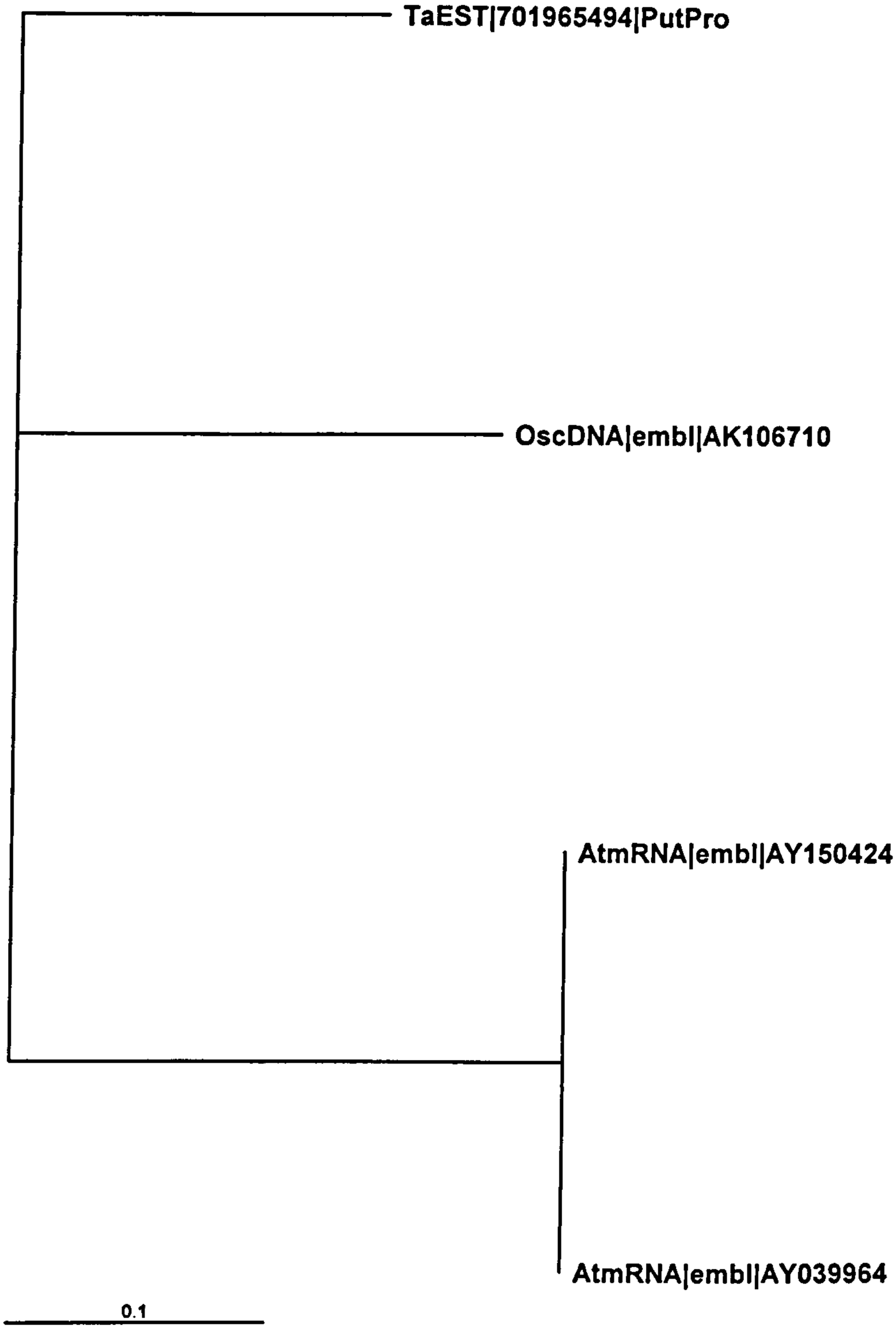
Phylogenetic tree for TaEST|702042046|DmTFIIB produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Mm=*Mus musculus;* Hs=*Homo Sapiens;* Dm=*Drosophila melanogaster;* TFIIB=Transcription Factor IIB; EVORF=Evoutionary conserved transcript; EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

Phylogenetic tree for TaEST|701968407|AHP1 produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* Hv=*Hordeum vulgare;* At=*Arabidopsis thaliana;* AHP1=Alkyl HydroPeroxide reductase; PutTF=Putative Transcription Factor; EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

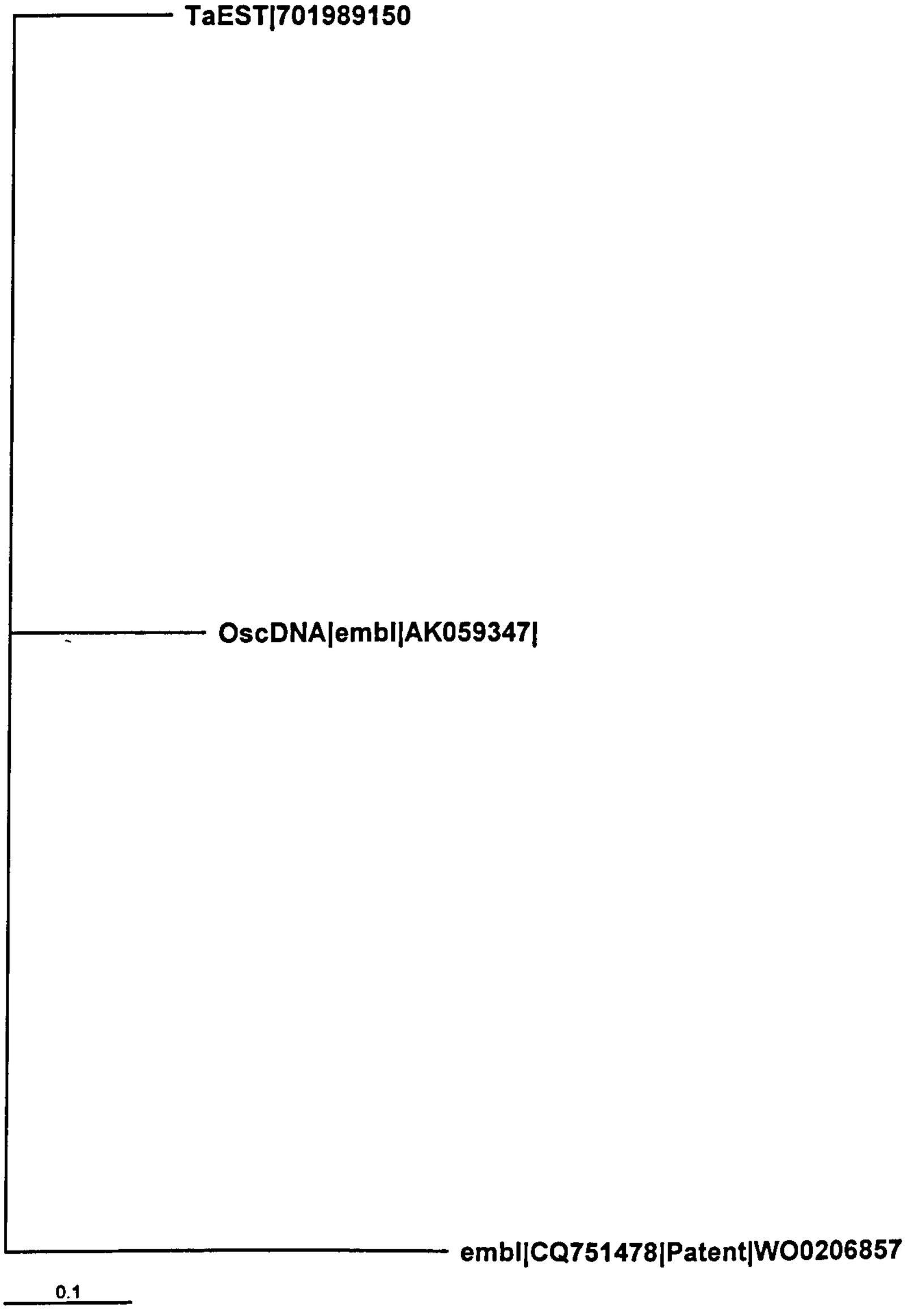


Phylogenetic tree for TaEST|701965813|IncyteEST produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* EST=Expressed sequence tag; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.

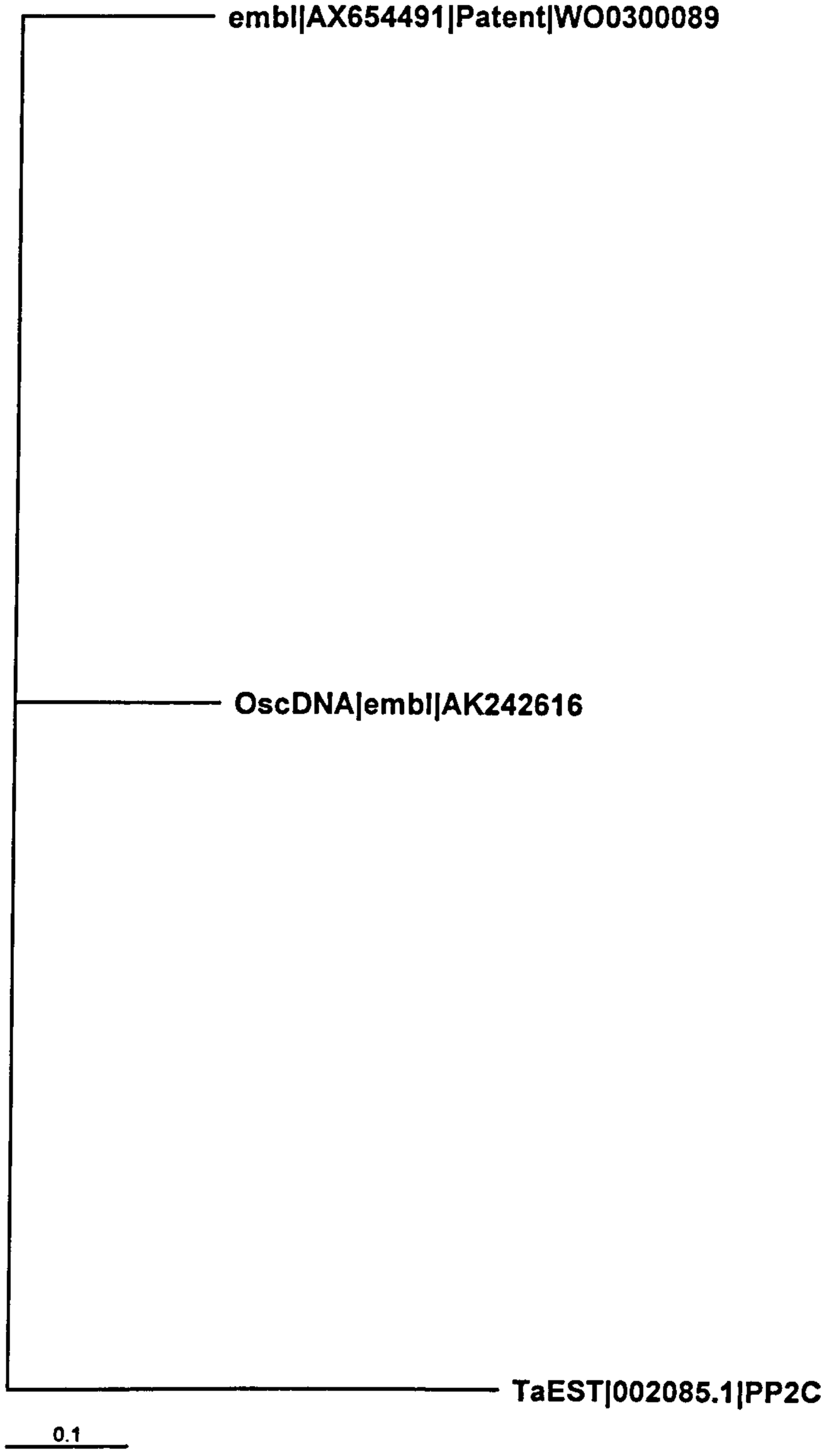
Phylogenetic tree for TaEST|020175.1|DsPTP produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* DsPTP= Dual Specificity Protein Phosphatase; EST=Expressed sequence tag; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.



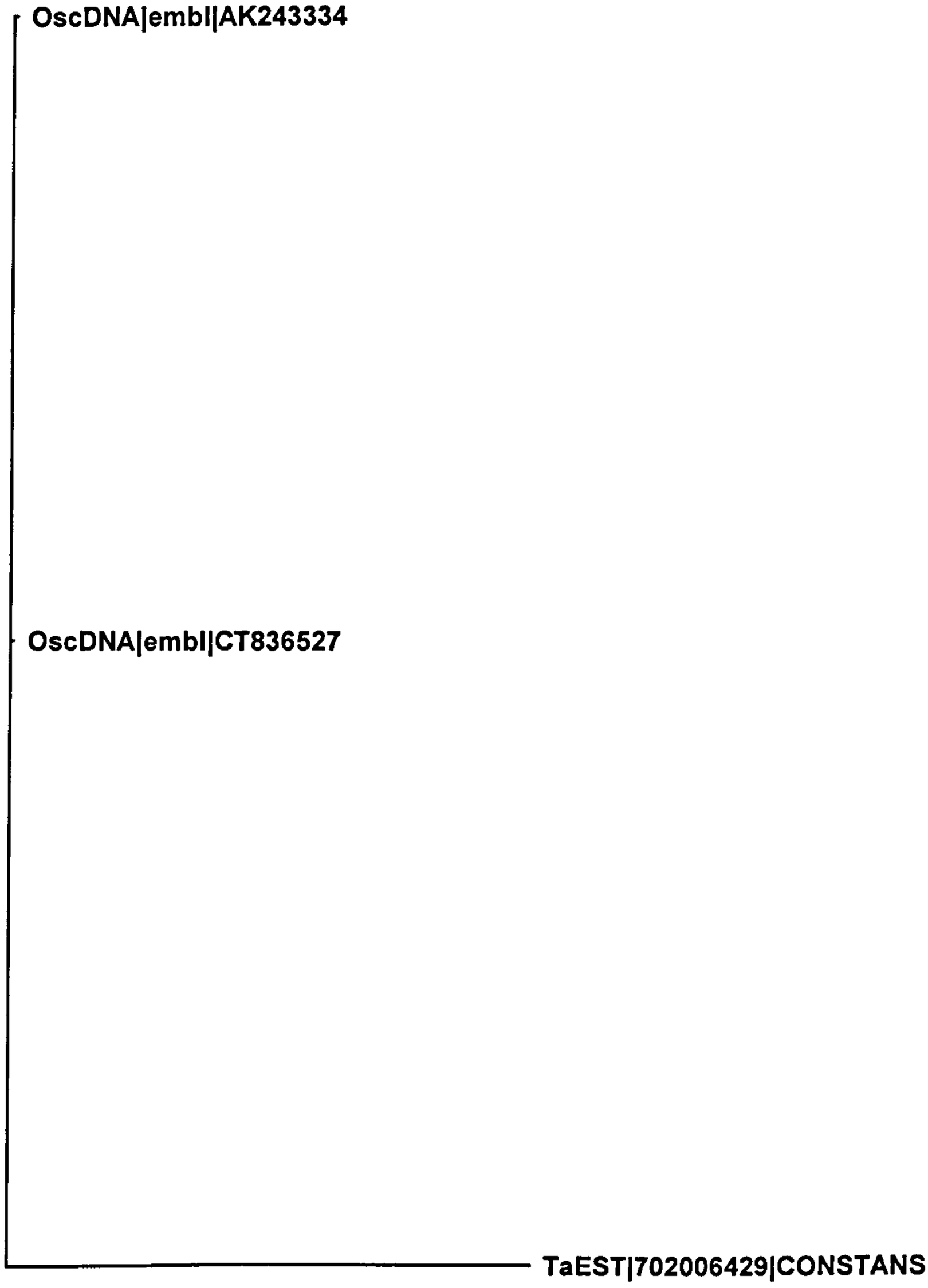
Phylogenetic tree for TaEST|701965494|PutPro produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a bootstrapped phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* At=*Arabidopsis thaliana;* PutPro=Putative Protein; EST=Expressed sequence tag; mRNA=mRNA of no known function; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.



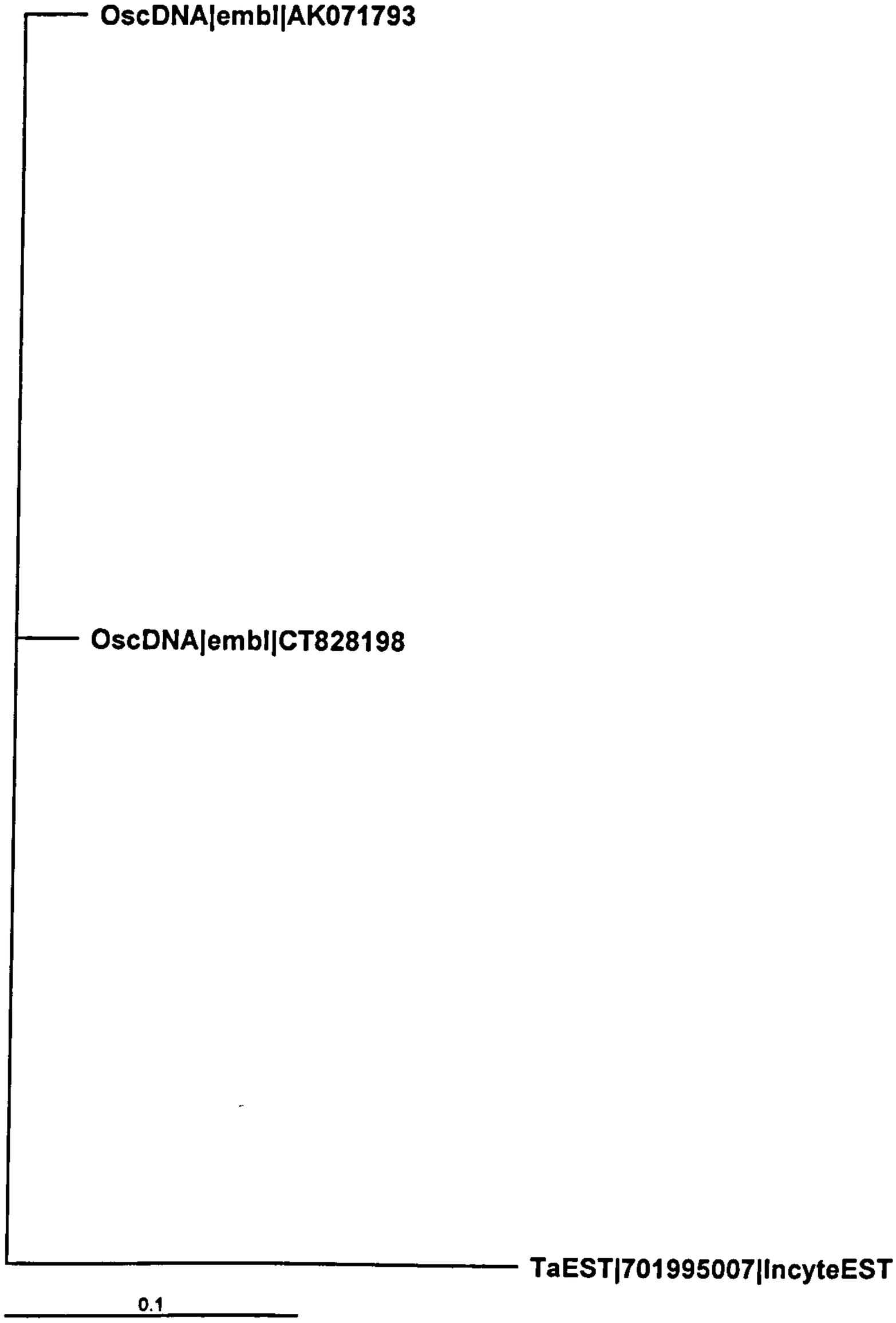
Phylogenetic tree for TaEST|701989150|Incyte EST produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* EST=Expressed sequence tag; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.



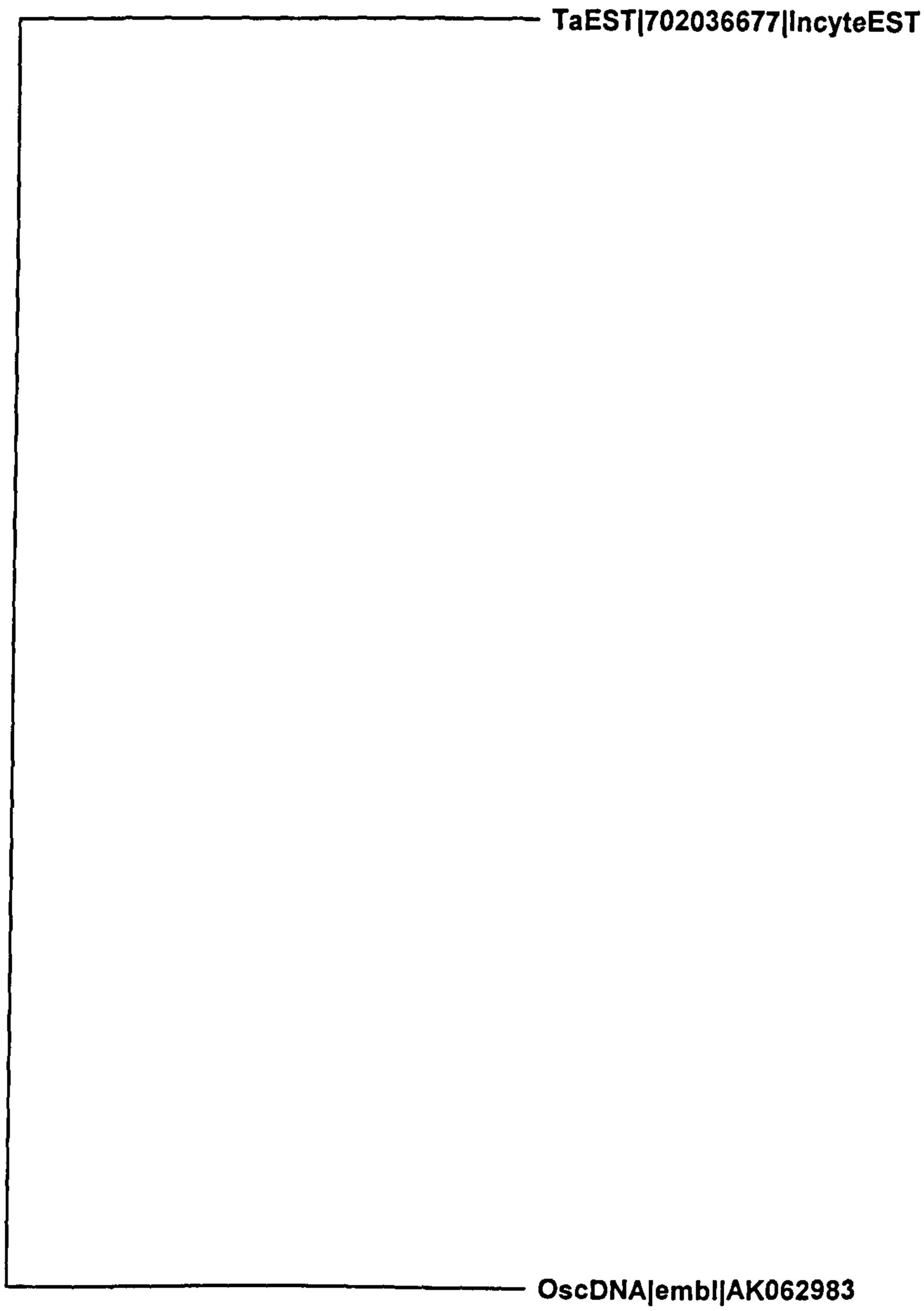
Phylogenetic tree for TaEST|002085.1|PP2C produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* PP2C= protein phosphatase 2C; EST=Expressed sequence tag; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.



Phylogenetic tree for TaEST|702006429|CONSTANS produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* CONSTANS=CONSTANS flowering protein; EST=Expressed sequence tag; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.



Phylogenetic tree for TaEST|701995007|IncyteEST produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* EST=Expressed sequence tag; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.



Phylogenetic tree for TaEST|702036677|IncyteEST produced by BLAST analysis in EMBL non-redundant database. Each sequence was aligned with Clustalx (Thompson *et al.*, 1997) to produce a phylogenetic neighbour-joining tree. Where Ta=*Triticum aestivum;* Os=*Oryza sativa;* EST=Expressed sequence tag; cDNA=cDNA of no known function. Scale bar indicates the number of nucleotide replacements per site.