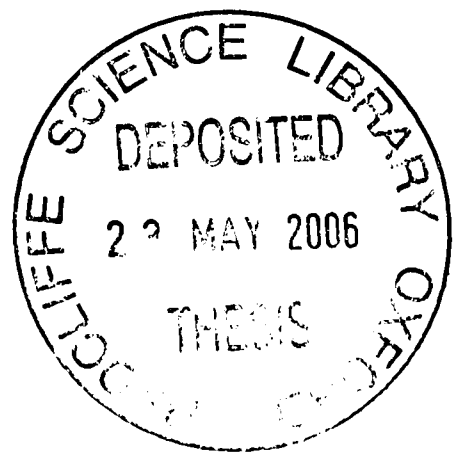
306119483Z

An analysis of interploidy crosses in maize

Paul Douglas Pennington,
MA (Cantab), MSc (Manchester)

Linacre College

University of Oxford
Trinity Term 2005
DPhil Thesis



An analysis of interploidy crosses in maize

Paul Douglas Pennington
Linacre College
Thesis submitted for the degree of Doctor of Philosophy, Trinity Term 2005

Reciprocal crosses between plants containing different genome numbers have been demonstrated to show alternate phenotypes in many species and are often lethal. The effects of interploidy crosses on kernel development in maize (*Zea mays*) have been reassessed using markers for two cell domains: the aleurone and the basal endosperm transfer layer (BETL); and by monitoring gene expression. The study confirms previous observations and offers novel insights into genomic interactions.

Endosperms from maternal excess crosses (tetraploid mother, diploid father) displayed reduced mitotic activity, and an accelerated rate of differentiation generally, including early starch accumulation, and earlier, atypical aleurone formation. The BETL failed to form normally, being replaced with aleurone-like cells, except for a few interspersed cells. Gene expression was altered, including those encoding cell domain markers and storage molecules.

Paternal excess endosperms (diploid mother, tetraploid father), showed an increase in cellular proliferation. Generally, differentiation was delayed: starch accumulation began late, and formation of the aleurone layer was delayed, and was morphologically highly irregular. BETL specification was also disturbed, with only a few cells showing characteristic wall modifications very late in development. Later development of the whole endosperm was characterized by ectopic gene expression, and zones of high cell proliferation. A greater proportion of the genome was mis-expressed in these crosses, reflecting the more extreme phenotype. Although expression levels of marker genes did not reflect the failure to develop a normal BETL, it may be that the cells in the BETL region had undergone an modified developmental program.

This work is considered in the context of evolutionary models which state that different parental roles lead to differential expression of genes from their respective genomes, and is compared with evidence from arabidopsis. The current study is considered generally supportive of such models.

Acknowledgements

This thesis depended on the help of a lot people, whom I should like to thank. Hugh Dickinson, my supervisor, has provided me with a lot of support from the start, and given me words of encouragement at all the right times, and helped put me in touch with all the right people. Jane Langdale also gave me enormous amounts of help, although I'm not sure she realizes just how much effect a few words of hers had on my work. I must also thank other members of the lab, including Jose Gutierrez-Marcos, who initially made suggestions about the design of the project, and Liliana Costa, who showed me a lot of the ropes at the start. I'd particularly like to thank Valérie Bourdon, who has been a soul-mate at the most difficult times, and given me real words of encouragement. I'd also like to thank Robert Grant-Downton, Alanna Keith Lucas, Peter Etchells, Jo Dowding, Janet Kenyon, Qing Zhang, and, of course, Sandie West for their enormous fun and company.

Practically, I have has support from many people, and their help must be acknowledged. From the department: Steve Rutherford for helping me at a hat's notice to print my microarrays, several times over; John Baker for photography; Ian Moore and Mark Fricker with help in confocal microscopy; Richard Buggs for help in flow cytometry; various people at Syngenta (Jealott's Hill) for help with greenhouse facilities; and all my DPhil classmates for camaraderie. From outside the department in Oxford: Gabrielle Reinkensmeier, from biochemistry for help with flow cytomtery; and most especially Nigel Saunders, and Lori Snyder from the William Dunn School of Pathology for life-saving microarray facilities, and the reassurance that I was not the only one to have seen the Aurora Borealis on a glass slide, when I was hoping for Manhattan.

My DPhil was supported by the BBSRC, although my fourth year would have been impossible without my own financing through employment I obtained through the University. For this I must express my enormous gratitude to Sue Bennett, Head of Student Administration, and Maureen McNaboe of the University Card Office, whose entertaining, generous nature helped keep me sane in the last year.

Oxford would not be Oxford, were it not for its colleges, and Linacre has been a lifeline at all stages of my DPhil, and as such I should thank all the staff, fellows and students, whom I have spent many a happy hour. In particular, my closest friends: Malti Gandhi, Eric Klotz, Joanne Pearson, Christiana Vonofakou, Stephen "6am" Whitelam and, of course, Bob Headington. And to Leonard, thank you for the music: very beautiful.

Most importantly, my family have been, as ever, the most important support for me during this thesis. Through their words of encouragement, by accepting my at times appalling moods, and with their generosity of financial help and holidays, they offered so much support. Thanks ever so much: mum, dad, and Julie – I couldn't have done it without you.

Contents

ACKNOWLEDGEMENTS I

CONTENTS II

LIST OF FIGURES AND TABLES V

UNITS, ABBREVIATIONS AND NOMENCLATURE VII

1 INTRODUCTION 1

1.1 MAIZE AS A MODEL ORGANISM IN DEVELOPMENTAL GENETICS 1
1.2 SEED DEVELOPMENT 3
1.2.1 Double Fertilization 3
1.2.2 Summary of male and female gamete formation 4
1.2.3 Fertilization 8
1.2.4 Development of endosperm 8
1.2.4.1 The basal endosperm transfer layer (BETL) 14
1.2.4.2 The embryo surrounding region (ESR) 18
1.2.4.3 The aleurone layer 19
1.2.4.4 The starchy central cells 21
1.2.5 Development of the embryo 22
1.3 CONTROL MECHANISMS OF SEED DEVELOPMENT 25
1.3.1 Parent-specific expression in the embryo sac 26
1.3.2 Dosage effects 27
1.3.3 Imprinting 29
1.3.4 Imprinting Mechanisms 31
1.3.5 Evolutionary origin of imprinting 32
1.3.6 The role of chromatin remodelling factors in epigenetics 33
1.4 INTERPLOIDY CROSSES 34
1.4.1 Interploidy crosses in maize result in seed abortion 35
1.4.2 Interploidy Crosses in arabidopsis 40
1.4.3 Interploidy crosses in other species 42
1.5 AIMS OF THIS WORK 44

2 METHODS AND MATERIALS 46

2.1 REAGENTS 46
2.1.1 Chemicals 46
2.2 GENERAL METHODS 46
2.2.1 Ethanol Precipitation of DNA 46
2.2.2 UV spectrophotometry 46
2.2.3 DEPC treatment of solutions 47
2.3 PLANT MATERIAL 47
2.3.1 Maize lines used 47
2.3.2 Plant Growth conditions 48
2.3.3 Endosperm isolation 48
2.3.4 Verification of genome number by flow cytometry 49
2.4 MICROSCOPY AND HISTOLOGY 49
2.4.1 Staining of samples to observe GUS activity 49
2.4.2 Fixation of material 50
2.4.3 Embedding and sectioning of samples 50
2.4.4 Staining of material in Periodic Schiff 's reagent for carbohydrate 51
2.4.5 Staining of material in Toluidine Blue 51
2.4.6 Microscopy 51
2.5 MICROARRAY TARGET PRODUCTION 52
2.5.1 Isolation of total RNA – Guanidium Isothiocyante Method 52
2.5.2 Isolation of total RNA – Trizol method 53
2.5.3 Isolation of RNA – Trizol and RNeasy clean-up method 54
2.5.4 RNA isolation – Hot Phenol Method 54
2.5.5 Denaturing Gel Electrophoresis 55
2.5.6 Agilent Bioanalyzer 55

2.5.7 RNA amplification 55
2.5.8 Fluorescent RNA labelling 55
2.6 CUSTOM MICROARRAY DESIGN 56
2.6.1 Selection of target sequences for array 56
2.6.2 Design of Primers 57
2.6.3 Polymerase Chain Reaction (PCR) 58
2.6.4 TBE gel 58
2.6.5 Isolation of DNA fragments from agarose gels 58
2.6.6 Ligation of DNA into pGEM-t easy vector 59
2.6.7 Transformation of competent E. coli 59
2.6.8 Colony PCR 59
2.6.9 Cell culture 60
2.6.10 Recovery of plasmid DNA from bacterial cultures 60
2.6.11 Sequencing of insert 60
2.6.12 Production of probe for microarray spotting 61
2.6.13 Printing of microarray slides 61
2.6.14 Propidium Iodide Staining of Microarray Slides 62
2.6.15 Microarray Prehybridisation 62
2.6.16 Hybridisation 63
2.6.17 Microarray Wash Protocol 63
2.7 OLIGOMER ARRAY PROTOCOLS 63
2.7.1 Array Processing 63
2.7.2 Dendrimer labelling protocol 64
2.8 SEMI QUANTITATIVE RT-PCRS 65
2.9 IT METHODS 66
2.9.1 Creation of a chromatin sequence database 66
2.9.2 Creation of a sample database 66
2.9.3 Primer_check: A program to check primers are in a given sequence 70
2.9.4 Listcompare: A program to count common genes between lists 70

3 PROPERTIES OF THE PLANT LINES USED IN THIS WORK 71

3.1 INTRODUCTION 71
3.1.1 Origin of plant lines used for different aspects of the work 71
3.1.1.1 Lines used in microarray and histological studies 71
3.1.1.2 Tetraploids used in crosses with reporter lines 73
3.1.1.3 Reporter lines; pBETL::GUS 73
3.1.1.4 Reporter lines; pVp1::GUS 74
3.2 RESULTS 74
3.2.1 Flow Cytometry confirms the ploidy of plants 74
3.2.2 Flowering time of maize lines 78
3.2.3 Seed mass variation between lines in balanced crosses 81
3.2.4 Interploidy crosses result in profoundly reduced seed masses 83
3.2.5 Interploidy kernels show low germination rates 84
3.3 DISCUSSION 88
3.3.1 Ploidy of plants 88
3.3.2 Seed mass effects between diploid plants demonstrate parental and hybrid effects 89
3.3.3 Seed mass following interploidy crosses 90

4 SPECIFICATION OF TWO CELL DOMAINS AND CELLULAR EFFECTS OF BALANCED AND INTERPLOIDY CROSSES 92

4.1 INTRODUCTION 92
4.1.1 The Vp1 Gene as a marker for the aleurone layer 92
4.1.2 The BETL1 gene as a marker for the basal endosperm transfer layer 92
4.1.3 Scope of this chapter 93
4.2 RESULTS 94
4.2.1 Activity of the pVp1::GUS reporter gene in balanced crosses 94
4.2.2 pVp1::GUS activity precedes aleurone morphology in balanced crosses 95
4.2.3 Activity of the pBETL-GUS reporter gene in balanced crosses 101
4.2.4 BETL cell morphology precedes pBETL1::GUS activity in balanced crosses 104
4.2.5 Activity of the pVp1::GUS reporter gene in maternal excess crosses. 105
4.2.6 Activity of the pBETL1::GUS reporter gene in maternal excess crosses 110
4.2.7 Activity of the pVp1::GUS reporter gene in paternal excess crosses. 112
4.2.8 Aleurone morphology is severely disrupted in paternal excess crosses 115

4.2.9 Activity of the pBETL1::GUS reporter gene in paternal excess crosses. *116*
4.2.10 Activity of an endogenous gene in paternal excess crosses *122*
4.2.11 Commencement of starch accumulation is altered in interploidy crosses *124*
4.2.12 Aleurone formation is distorted in both directions of interploidy cross *125*
4.2.13 Nuclear changes through development *127*
4.2.14 Embryo development in interploidy crosses *127*
4.3 DISCUSSION 132
4.3.1 Balanced crosses with reporter lines reveal markers for cell types but not cell identity *132*
4.3.2 Maternal excess interploidy effects on seed development *135*
4.3.3 Paternal excess interploidy effects on seed development *138*

5 EFFECT OF INTERPLOIDY CROSSES ON GENE EXPRESSION 141

5.1 INTRODUCTION 142
5.1.1 Microarray Production *142*
5.1.1.1 Target 143
5.1.1.2 Labelling Method 144
5.1.1.3 Signal amplification 144
5.1.2 Gene regulation by chromatin remodelling *145*
5.1.2.1 DNA methylation 146
5.1.2.2 Histone modifications 148
5.1.2.3 Polycomb Group Proteins 151
5.1.2.4 ATP-dependent remodelling enzymes 152
5.1.3 Summary of large scale experiments in maize *153*
5.2 RESULTS 154
5.2.1 Oligomer array design *154*
5.2.2 Experimental strategy *157*
5.2.2.1 Assessment of hybridisation quality 159
5.2.2.2 Dendrimer technology shows sequence specific biases 166
5.2.2.3 Estimation of the number of genes expressed in endosperms 170
5.2.2.4 Most genes are unaffected by interploidy crosses 174
5.2.2.5 Most highly expressed genes are similar between conditions 176
5.2.2.6 Genes of interest 176
5.2.2.7 Genes most significantly changed in different conditions 178
5.2.2.8 Analysis of functional categories of genes regulated between samples 179
5.2.2.9 Analysis of genes over-represented in different conditions 179
5.2.2.10 Correlations of expression patterns between samples 188
5.2.2.11 The number of genes regulated in the same channel between different comparisons 189
5.2.3 Development of a custom micro-array *191*
5.2.3.1 Levels of expression 192
5.2.3.2 Regulated Genes 193
5.2.4 Verification of results by RT-PCR *200*
5.3 DISCUSSION 203
5.3.1 Technological Insight *204*
5.3.2 Gene expression in early endosperm *204*
5.3.3 Changes in gene transcription following interploidy crosses *205*

6 GENERAL DISCUSSION 209

6.1 SUMMARY AND MAIN CONCLUSIONS 209
6.2 INSIGHTS INTO THE SPECIFICATION OF THE ALEURONE AND BETL 209
6.3 SUPPORT FOR THE CONFLICT HYPOTHESIS? 213
6.4 PATERNAL EXCESS CONFERS GENOMIC STRESS ON DEVELOPING ENDOSPERM 215
6.5 GENERAL DESCRIPTION OF INTERPLOIDY CROSSES 217
6.6 FUTURE PROSPECTS 217

7 REFERENCES 220

8 APPENDICES 232

8.1 CONTENTS OF ENCLOSED CD 232
8.2 PUBLICATION 232
8.3 TABLE OF SEQUENCE DATA AND PRIMERS USED TO PRODUCE CUSTOM MICROARRAY .. A1-16

List of Figures and Tables

Figures

Chapter 1

Figure 1-1 Ovule formation 7
Figure 1-2 Fertilization in maize. 7
Figure 1-3 Early mitoses in maize endosperm development 11
Figure 1-4 Cellularization of maize endosperm 11
Figure 1-5 Endosperm domains in maize 13
Figure 1-6 Embryogenesis in maize 24
Figure 1-7 Demonstration of interploidy crosses in maize 36
Figure 1-8 Images of kernels from balanced and interploidy kernels as reported by Cooper 36
Figure 1-9 Summary of Cooper's observations in maize interploidy crosses 37

Chapter 2

Figure 2-1 Screenshot from sample database 68
Figure 2-2 Table of relationships for kernel database 69

Chapter 3

Figure 3-1 Images of mature, flowering plants. 80
Figure 3-2 Time to flowering 80
Figure 3-3 Increase in seed mass following within ploidy pollinations 82
Figure 3-4 Increase in seed mass for interploidy crosses. 85
Figure 3-5 Comparison of seed mass between reciprocal interploidy crosses 86
Figure 3-6 Mature, dried kernels 86
Figure 3-7 Seed mass of mature, desiccated kernels 87
Figure 3-8 Germinated triploid kernels 87

Chapter 4

Figure 4-1 Logic for transgene parental origin control crosses 93
Figure 4-2 Frequency of kernels that stain for GUS 97
Figure 4-3 Progression of GUS activity in pVp1::GUS line x H99-2x crosses 98
Figure 4-4 Progression of GUS activity in H99-2x x pVp1::GUS line crosses 99
Figure 4-5 Comparison of aleurone-specific promoter activity and adoption of aleurone identity. 100
Figure 4-6 Progression of GUS activity in pBETL::GUS line x H99-2x crosses 102
Figure 4-7 Progression of GUS activity in H99-2x x pBETL::GUS line crosses 103
Figure 4-8 Initial appearance of transfer cell morphology in basal cells 104
Figure 4-9 Progression of GUS activity in HybT x pVp1::GUS line crosses. 107
Figure 4-10 Maternal excess Vp1 reporting pattern in a H99-4x maternal genotype. 108
Figure 4-11 Gaps of varying size within the ectopic basal 'aleurone' 108
Figure 4-12 PAS staining of non-GUS staining gaps in basal 'aleurone' 109
Figure 4-13 Progression of GUS activity in HybT x pBETL::GUS line crosses. 111
Figure 4-14 Progression of GUS activity in pVp1::GUS line x HybT crosses 114
Figure 4-15 Diffusion of GUS precipitate into underlying cells in paternal excess crosses 115
Figure 4-16 Vp1::GUS activity on ectopic surfaces 115
Figure 4-17 Progression of GUS activity in pBETL::GUS line x HybT crosses 118
Figure 4-18 Apically located GUS activity in a pBETL::GUS x H99-4x cross 119
Figure 4-19 Patches of GUS activity on base of paternal excess endosperm 119
Figure 4-20 Ectopic localization of GUS in pBETL::GUS x HybT plants 120
Figure 4-21 Kernel demonstrating patches within larger patches of GUS activity. 120
Figure 4-22 Variability between pBETL::GUS activity in paternal excess kernels 120
Figure 4-23 PAS staining of basally located cells 121
Figure 4-24 Ectopic localization of anthocyanin in paternal excess crosses 123
Figure 4-25 pVP1::GUS reporting profiles in comparison to aleurone identity 123
Figure 4-26 Dark field microscopy reveals pVP1::GUS production and starch accumulation 124
Figure 4-27 Toluidine blue staining of peripheral layer 126
Figure 4-28 Tumour like outgrowths of the surface of paternal excess endosperm 129
Figure 4-29 DAPI staining of centrally located nuclei 129
Figure 4-30 Areas of different mitotic activity in late paternal excess endosperms 130
Figure 4-31 Appearance of embryo kernels in paternal excess crosses 130
Figure 4-32 Appearance of embryos 131

Chapter 5

Figure 5-1 Mechanical isolation of 4 DAP endosperm 156
Figure 5-2 Design of oligomer microarray experiment 156
Figure 5-3 Details of RNA samples used in the oligo array experiment 158
Figure 5-4 Examples of poor quality spots eliminated from the array 162
Figure 5-5 Scan of one microarray (Hybridisation 1, slide A39) 163
Figure 5-6 Table of RI plots for oligomer array hybridizations 164
Figure 5-7 Frequencies of spot intensity values 165
Figure 5-8 Cross-hyb product of gene channel ratios. 169
Figure 5-9 Estimation of background thresholds using Alien controls 169
Figure 5-10 Histogram of ratio frequencies. 173
Figure 5-11 Histogram of fold ratio changes 175
Figure 5-12 Gene Ontology Groupings of regulated genes 185
Figure 5-13 Spearman's Rank Correlations between slides 188
Figure 5-14 Number of genes regulated more than 1σ in common between samples. 190
Figure 5-15 Number of genes regulated more than 2σ in common between samples. 190
Figure 5-16 Summary of percentage of genes in common between different samples 190
Figure 5-17 Visualization of RNA and aRNA on denaturing gel and Bioanalyzer 197
Figure 5-18 Genes most substantially regulated on custom array 198
Figure 5-19 RT-PCR analysis of gene activity through endosperm development 202

Chapter 6

Figure 6-1 Plan of interploidy series for future microarray investigation 219

Tables

Table 1 Summary of BETL expressed genes 17
Table 2 Summary of ESR expressed genes 19
Table 3 Maize lines in which interploidy crosses have been documented 36
Table 4 Summary of non-transgenic maize lines used in these studies 47
Table 5 Summary of transgenic maize lines used in these studies 48
Table 6 Sample of flow cytometry results 77
Table 7 Summary of histological observations following interploidy crosses 132
Table 8 Features of maize oligomer microarray 154
Table 9 Summary Statistics of microarray hybridisations 162
Table 10 Normalization statistics of microarray 172
Table 11 Frequency table fold-ratio changes between samples 175
Table 12 Table of regulation of genes of interest 177
Table 13 Most highly regulated genes in 2x × 4x crosses 181
Table 14 Most highly regulated genes in 4x × 2x crosses 182
Table 15 Most differentially expressed between 2x × 4x and 4x × 2x crosses 183
Table 16 Most up and down regulated genes in 2x × 2x and 4x × 4x crosses 184
Table 17 Summary of GO classes classed as significant from IGA package. 186
Table 18 Significantly different gene classes from BINGO analysis 187
Table 19 Genes represented on custom array 196
Table 20 Hybridisations carried out using custom array 196
Table 21 Tabulation of gene expression changes in custom microarray 199
Table 22 Summary of effects of interploidy crosses 217

Units, abbreviations and nomenclature

Units of Measure

Length

m	metre
cm	centimetre (10^{-2} m)
mm	millimetre (10^{-3} m)
µm	micrometre (10^{-6} m)
nm	nanometre (10^{-9} m)

Mass

g	gram
mg	milligram (10^{-3} g)
µg	microgram (10^{-6} g)

Volume

l	litre
ml	millilitre (10^{-3} l)
µg	microlitre (10^{-6} l)

Concentration

M	Molar (moles l^{-1})
mM	millimolar (10^{-3} M)
µM	micromolar, (10^{-6} M)

Time

s	seconds
min	minutes
h	hours

Other Units of Measure

°C	degrees Celsius
W	Watts
J	Joules
mJ	milliJoules

Other Abbreviations

aRNA	Amplified RNA
ABA	Abscissic acid
BETL	Basal endosperm transfer layer
BSA	Bovine serum albumin
CSE	Central starchy endosperm
DAP	Days after pollination
ddH2O	Double distilled water
DEPC	Di-ethyl pyrocarbonate
cDNA	Complementary DNA
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylene glycol bis(2-aminoethyl ether)-N,N,N'N'-tetraacetic acid
GO	Gene Ontology
GUS	ß-glucoronidase
mRNA	Messenger RNA
dNTP	Deoxyribonucleotidephosphates
LB	Luria Broth
NPT	Neomycin phosphotransferase
PCR	Polymerase Chain Reaction
PIPES	Piperazine-N,N'-bis(2-Ethanesulfonic Acid)
RNA	Ribonucleic acid
rpm	Revolutions per minutes
SDS	Sodium dodecyl sulphate
SSC	Saline–Sodium Citrate
Taq	Thermus aquaticus polymerase
TBE	Tris-Borate-EDTA
Tris-HCl	Tris (hydroxymethyl) aminomethane Hydrochloride
UTR	Untranslated region
UV	Ultraviolet

Nomenclature

Throughout this thesis, the common convention will be applied of writing the constitution of the female parent first when referring to endosperm balance and direction of cross. Hence a normal endosperm comprises three genomes of which two are female derived and one is male derived (i.e. 2♀:1♂). For crosses, the conventional notation 'x' indicating 'is crossed by' is used. Ploidy is denoted by the suffix, so '-2x' means 'diploid', and '4x' means 'tetraploid', where 'x' refers to the haploid set of genomes. Hence H99-2x x H99-4x means a H99 diploid female ear which is crossed with pollen from a H99 tetraploid male.

In references to interploidy crosses 'maternal excess' implies a tetraploid female ear crossed with diploid pollen, and 'paternal excess' refers to a diploid ear crossed with tetraploid pollen.
The font used for the main body text in this thesis is Serifa BT, point size 11.

1 Introduction

1.1 Maize as a model organism in developmental genetics

Maize is a monocotyledonous grass in the Poaceae family (Kellogg, 2001), and its relatives include the important crops rice, wheat, barley, and sorghum. In addition to being a staple of nutrition in humans and livestock (Gibbon and Larkins, 2005), maize has also played a significant role in biological research for many decades.

The historical origin of maize is unclear. It is widely held to be an ancient allotetraploid, although it is now a functional diploid, derived from hybridization between two uncertain progenitor species, at least 4.8 million years ago (Swigonova et al., 2004). One of the closest living ancestors to maize is teosinte, which is also a candidate for one of its progenitors (Mangelsdorf, 1974). Following tetraploidization maize has undergone extensive genomic changes, including rapid sequence expansion primarily due to the activity of retrotransposons, which comprise a large proportion of the genome (Meyers et al., 2001). Domestication of the species to improve traits such as seed size and quality (Weatherwax, 1923) has produced numerous maize lines displaying much phenotypic variation, and there are now about 3,500 identified variations in the plant (Coe, 2001) and considerable sequence variation between lines (Brunner et al., 2005; Fu and Dooner, 2002).

Maize could be considered technically more challenging than other model plants, such as *Arabidopsis thaliana* (arabidopsis), owing to its relatively long generation time (between three and four months) and its relatively large genome size of 2,500 Mb (Arumuganathan and Earle, 1991), spread over a haploid set of 10 chromosomes. The maize genome is highly complex, comparable to the human genome (Martienssen et al., 2004), and it is estimated to contain 59,000 genes (Messing et al., 2004) although these represent only around 10-15% of the genome (Bennetzen et al., 2001). Maize has,

however, a number of advantages that make it highly suited for certain areas of research. The extensive natural variation of maize facilitates the identification of new mutants, and assists in linkage experiments. Its diversity of transposable elements may be used to generate novel insertional mutants (Fernandes et al., 2004; Walbot, 2000). Some lines of maize also carry a supernumerary chromosome (B) alongside the normal haploid set (Auger and Birchler, 2002). This chromosome frequently undergoes non-disjunction in male meiosis (Birchler et al., 1990) which, in combination with chromosomal translocations, make it possible to generate seeds with one, two, or three copies of a chromosomal region.

For many of the reasons listed above maize has accumulated a distinguished track record in modern genetics, playing a central role in the discovery of phenomena such as imprinting (Kermicle, 1970), transposable elements (Federoff, 1983), paramutation (Chandler et al., 2000) and heterosis (Duvick, 2001; Moll et al., 1965).

Maize is a monoecious plant, in which the male and female inflorescences are physically separated; the pollen-bearing tassel appearing at the upper apex, and the female ear being subtended by leaves lower down on the plant. Hence, controlled pollinations are straightforward. The female inflorescence bears numerous fruits, each of which represents an independent fertilization event.

The seeds of maize are relatively large, probably largely due to the artificial selection imposed on it by man, rendering studies of seed development easier than in other species (Cheng and Pareddy, 1994) as the endosperm and embryo are easily isolated from about four and six days after pollination respectively. It is even possible to isolate the unfertilised egg and central cells using micromanipulation techniques (Kranz et al.,

1992), which can be fertilized *in vitro* to form embryos and endosperm (Kranz et al., 1992; Kranz et al., 1998).

Finally, research into this important crop species has been greatly advanced by the strong community of researchers, amongst whom resources, materials, and ideas are readily shared[1]. A combination of these shared resources, the availability microarrays supplied at cost, biological advantages such as large size of the endosperm, are major factors in the selection of maize for the work reported here.

1.2 Seed development

Understanding seed development is important for numerous reasons. Firstly, this product of reproduction represents a means by which the progeny of plants can remain dormant until appropriate growing conditions return, ensuring propagation of its genes to the next generation (Walbot and Evans, 2003). Secondly, seeds are important for human and animal nutrition, and a better understanding of seed development may enable us to improve nutritional quality (Wilson et al., 2004). Finally, a better understanding of reproductive processes may assist in engineering apomictic qualities into the plants, which would allow the stable propagation of desirable traits (Vinkenoog and Scott, 2001).

1.2.1 Double Fertilization

Central to the process of seed development and fundamental to our understanding of interploidy crosses is the fact that the angiosperms, the group of flowering plants which includes maize, undergo a process of double fertilization - first discovered in the late 19th century (Guignard, 1899; Nawaschin, 1898). Unlike in mammals, where a

[1] http://www.maizegdb.org/

single gamete from each parent contributes to the zygote, in angiosperms, the female gametophyte comprises two "germ" cells, the egg cell and the central cell, each of which must be fertilized by a separate male sperm cell. Whereas the egg cell contains a single haploid genome, and will develop into the embryo upon fertilization, the central cell contains two copies of the haploid genome, and will generate the endosperm.

The evolutionary origin of the endosperm remains unclear, with two main arguments proposed in explanation (Berger, 2003). First, it has been suggested that the endosperm might be a 'continuation' of female megagametophyte development, requiring a fertilization event to re-induce it (Baroux et al., 2002). In some plants, such as the gymnosperms, the female gametophyte performs a similar role to the endosperm in nourishing the growing embryo (Friedman and Williams, 2004). The second hypothesis suggests that the endosperm might have originated from a supernumerary, altruistic, 'twin' of the embryo, whose development promotes the development of the embryo, at the expense of its own progression into the succeeding generation (Friedman, 1992). Until recently it was thought that the occurrence of double fertilization, and the creation of twin embryos in the lower plant ephedra, supported the 'altruistic twin' hypothesis (Friedman, 1992). However, a reassignment of the phylogenetic position of this group (Bowe et al., 2000; Chaw et al., 2000), and the discovery of diploid and triploid endosperms, including in amborella, the basal-most angiosperm (Williams and Friedman, 2002), have required this interpretation to be re-examined, providing a reinforced support for a gametophytic origin of the endosperm.

1.2.2 Summary of male and female gamete formation

Unlike in animals, where the germline is specified early in development, plants only undergo reproductive development upon receiving appropriate internal or environmental cues (Walbot and Evans, 2003). As this process precedes the creation of

the germ cells which are involved in the fertilization that leads to seed development, it is instructive to understand the processes involved.

Megasporogenesis and megagametogenesis are key phases of female gametophyte development. These occur within the ovule which is a specialized structure comprising a nucellus, integuments and the funiculus (Gasser et al., 1998). The nucellus initiates growth as a protrusion from the placental wall of the ovary (Reiser and Fischer, 1993 and Figure 1-1A), and from within the nucellus a centrally-located archaesporial cell is formed, characterized by its large size, and large nucleus, which will undergo the processes of megasporogenesis. Two sheath–like appendages called the integuments emerge from near the base of the nucellar projection, extend upwards (Figure 1-1B), enveloping the nucellus except for a small gap at the top called the micropyle, which will define one pole of the embryo sac (Figure 1-1D). The opposite end of the ovule is attached to the placental wall by a thin stalk, called the funiculus, and the chalaza, where it attaches to the nucellus, defines the opposite pole of the egg sac: (Figure 1-1D). During the development of the ovule, the archaesporial cell develops into the megaspore mother cell and undergoes meiosis, to produce a linear tetrad of megaspores. Three of these cells abort, with a release of hydrolytic enzymes, and the breakdown of cellular components (Russell, 1979).

The next phase of embryo-sac development, megagametogenesis, involves three rounds of nuclear division without cell wall formation (Huang and Russell, 1992). The resulting eight-nuclear structure then undergoes a cellularization process, which establishes the two synergid cells and the egg cell at the micropylar end, and three chalazal-most antipodals - the other two 'polar' nuclei remaining in the central cell. Hence, the egg cell receives one haploid copy of the genome, whereas the central cell

carries two such nuclei. All three should be genetically identical, as they originate from the same meiotic event.

After this point the embryo sac matures, which involves a number of events including the modification of the synergids to attract the pollen tube (Higashiyama et al., 2001) and modification of the cell wall structure of the antipodal cells, which also undergo a polyploidization event (Diboll and Larson, 1966).

In the male sexual structure, the anther, the process of microsporogenesis generates the male gametes (Bedinger and Russell, 1993). The pollen mother cells are produced from archaesporial cells within the anther, and embark on a meiotic division to form a tetrad of four haploid cells or microspores, and these are encased in a callose wall (Knox and Heslop-Harrison, 1970). The surrounding cells of the tapetum release a β-1,3-glucanase that breaks down the callose to release the microspores into the anther locule (Steiglitz, 1977). During this time, the outer pollen wall, or exine, is deposited: a strong resistant material comprising the complex molecule sporopollenin. Additionally, a central vacuole forms, resulting in the displacement of the cytoplasm of the pollen grain to its outer margin. A post-meiotic mitosis then generates two nuclei: one vegetative nucleus organizes the pollen tube. The other is a precursor generative nucleus which undergoes a second mitotic division to form the two sperm cells, one of which will fertilize the central cell, and the other the egg cell of the same embryo sac, apparently at random (Birchler, 1993). As the pollen grain of maize matures, starch accumulates in the vacuole (Bedinger and Russell, 1993). Sperm cells have been shown to express a diverse range of mRNA species, and are rich in retrotransposons (Engel et al., 2003).

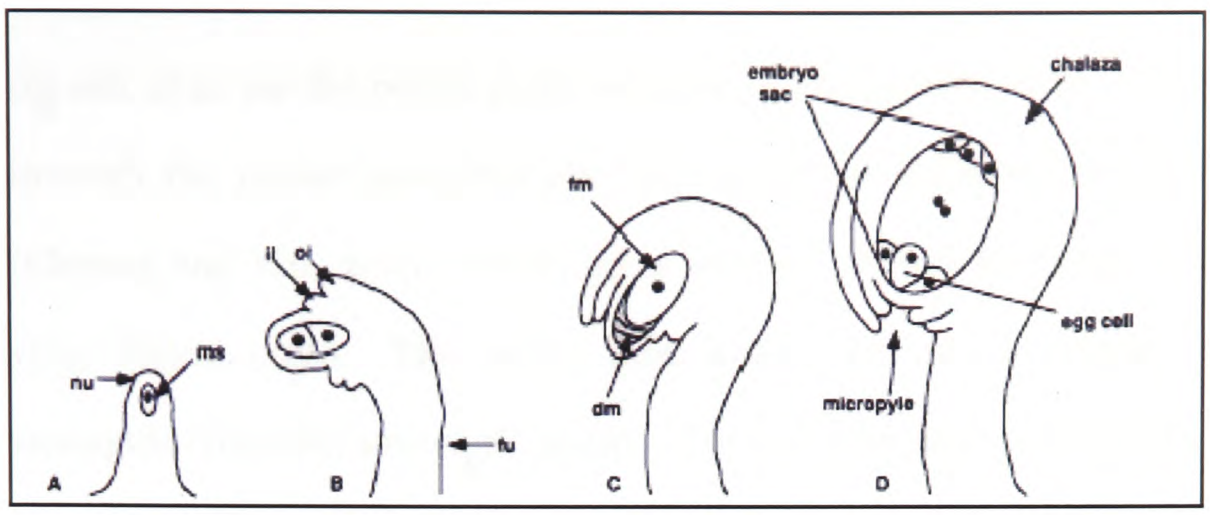


Figure 1-1 Ovule formation

Polygonum-type ovule development. **A**) Formation of a megasporocyte (ms) from within the nucellus (nu). **B**) Extension of the funiculus (fu), and growth of the outer and inner integuments (ii, oi), whilst the archaesporial cell undergoes the first meiotic division. **C**) After completion of meiosis, a linear tetrad is formed, one of the products of which becomes the functional megaspore (fm), whereas the others degenerate (dm). **D**) Ovule following megagametogenesis. Adapted from (Reiser and Fischer, 1993).

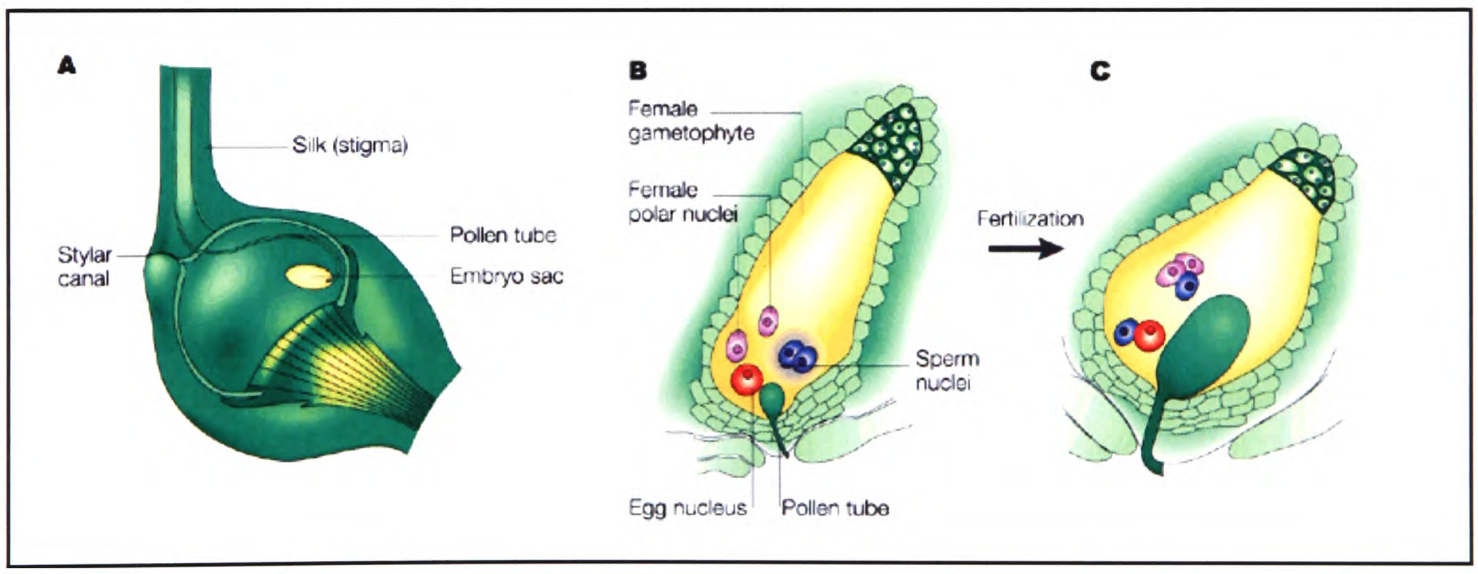


Figure 1-2 Fertilization in maize.

A) Female reproductive organs, in which the embryo sac is embedded within the ovary, the silk bears the stigmatic surface on which pollen grains germinate, to produce a germ-tube which grows towards the embryo sac. **B**) The embryo sac at the point of double fertilization as the pollen tube enters allowing release of the two male sperm cells. **C**) One sperm cell fertilizes the egg cell, and the other fuses with a polar nucleus in the central cell, karyogamy. Adapted from (Coe, 2001).

1.2.3 Fertilization

At maturity pollen is released through the process of dehiscence, when the anther elongates and bursts open to release pollen, which is transmitted by wind. On reaching the silk of an ear the pollen grain rehydrates, and growth of the pollen tube commences through the pollen pore into the silk and down its length towards the embryo sac (Cheung and Wu, 2001), which takes between 16 and 24 hours (Sheridan and Clark, 1994 Figure 1-2A). The pollen tube is attracted by the signals emitted from the synergids (Higashiyama et al., 2001). The tube then enters one of the synergids causing it to degenerate. The two sperm cells are then able to fertilize the egg cell and central cell, events which occur 'simultaneously' (Randolph, 1936).

Following the two fertilization events the two structures produced embark on different developmental pathways. The endosperm divides more rapidly than the zygote: the zygote does not divide until 40 hours after fertilization, by which time the endosperm already comprises between 8 and 32 free nuclei (Randolph, 1936).

1.2.4 Development of endosperm

Although different plant taxa display differing modes of endosperm development (Vijayaraghavan and Prabhakar, 1984), the nuclear type discussed here is the most common, occurring in both model plants maize and arabidopsis.

In outline, maize endosperm development can be split into several phases. The initial phase involves several rounds of nuclear division without cell division, to form a coenocyte: a structure in which many nuclei share a common cytoplasm (Olsen, 2001). A period of cellularization then follows, after which endosperm undergoes a phase of regular mitotic activity, when cell domains are elaborated. Then, starting from the

centre of the endosperm, cells cease mitosis and embark on a period of cell expansion. This activity gradually spreads towards the outer cells of the endosperm, until eventually only the outer cell layer, the aleurone, continues to be actively mitotic. The final phase of maturation involves programmed cell death of central endosperm, desiccation and metabolic changes in the aleurone in preparation for dormancy (McCarty and Carson, 1991), since this cell domain exclusively is reactivated during germination.

Following fertilization of the central cell, the sperm cell nucleus fuses with the homodiploid central cell nucleus. It is not clear exactly how long this takes, but in an *in vitro* system it was complete within three hours (Kranz et al., 1998). The first mitotic divisions begin within the next 2-4 hours (Kowles and Phillips), each in a predictable plane perpendicular to the central cell axis (Figure 1-3A and B). Significantly, these mitoses occur in the absence of cytokinesis, leading to a ring of eight nuclei at the base of the central cell (Figure 1-3C) (Olsen, 2004b). Each of these nuclei ultimately generate an 'eighth' sector of the mature endosperm, as determined by clonal analysis (McClintock, 1978). The following mitoses, which also lack cell wall formation, do not have predictable planes, but the spatial constraints imposed by the large central cell vacuole produces a multinucleate coenocyte of 256 to 512 nuclei in a single layer at the periphery of the coenocyte cell wall (Figure 1-3D) (Walbot, 1994). It is proposed that this community of nuclei in a shared cytoplasm might be a mechanism of specifying different ultimate cell fates on nuclei in different locations within the coenocyte (Becker et al., 1998). For example, the presence of morphogens, cell-fate determining factors whose effects change with concentrations (Gurdon et al., 1995; Wolpert, 1971), could form a gradient between different poles of the coenocyte. This is a mechanism known to determine cell fate in the drosophila embryonic blastoderm, which also has syncitial nuclei (St Johnston and Nüsslein-Vollhard, 1992).

In somatic cells, the products of mitosis are separated by the formation of interzonal phragmoplasts between the separating daughter nuclei. These are networks of microtubules which allow the deposition of polymers that constitute the cell plate. It is thought that this activity is suppressed in coenocytic endosperm nuclei, as in some species, non-functional interzonal phragmoplasts form (Brown et al., 1994b) and sometimes even rudimentary cell walls (Tian et al., 1998).

Coenocytial development is followed by a process of cellularization, in which networks of microtubules form between neighbouring nuclei of the coenocyte; these are referred to as radial microtubule systems (RMS) (Figure 1-4E) (Olsen, 2001). These microtubules radiate from nuclear envelopes to form so-called nucleocytoplasmic domains (NCDs) (Olsen, 2004b). Initially, microtubules from adjacent nuclei overlap, but shortly, distinct interzones develop (Brown et al., 1994a), and are termed cytoplasmic phragmoplasts in barley (Figure 1-4F). Cell wall material then starts to be deposited in the interzone of adjacent cytoplasmic phragmoplasts. These cell walls form initially at the outer edge near the boundary of the coenocyte, and progress centripetally to form tube-like structures surrounding each nucleus within the coenocyte (Figure 1-4B and G). At this point the nucleus within each alveolus becomes associated with the central cell wall via a microtubule network termed an adventitious phragmoplast (Figure 1-4H). These cells now undergo a periclinal nuclear division forming conventional interzonal phragmoplasts (Figure 1-4I). Normal cell plate deposition between daughter nuclei occurs creating a layer of cells at the periphery, and a new layer of alveoli towards the centre of the central cell (Figure 1-4J). A series of divisions forming an outer cell layer and new inner alveolar layer proceeds four to five times until eventually the central cell is fully cellularized, this process taking between two to three days (at around 4 DAP) (Olsen, 2001).

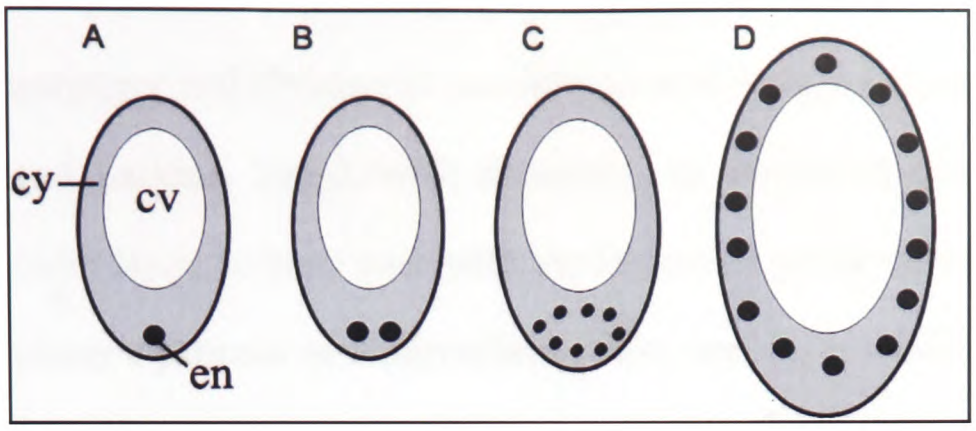


Figure 1-3 Early mitoses in maize endosperm development.

A) The fused triploid nucleus (en), lies towards the base of the central cell, which comprises a large vacuole (cv), restricting the cytoplasm (cy) to a thin layer at the periphery. **B**) An initial mitosis occurs in a predictable plane, perpendicular to the long axis of the central cell in the absence of the formation of a functional interzonal phragmoplast, nor cell wall formation. **C**) Two more rounds of mitosis without cell wall formation form a ring of nuclei at the base of the endosperm. **D**) Further mitoses are no longer in predictable planes, but the spatial constraints imposed by the central vacuole lead to a single layer of nuclei, in a common cytoplasm. This structure is referred to as the coenocyte. Taken from (Olsen, 2004b).

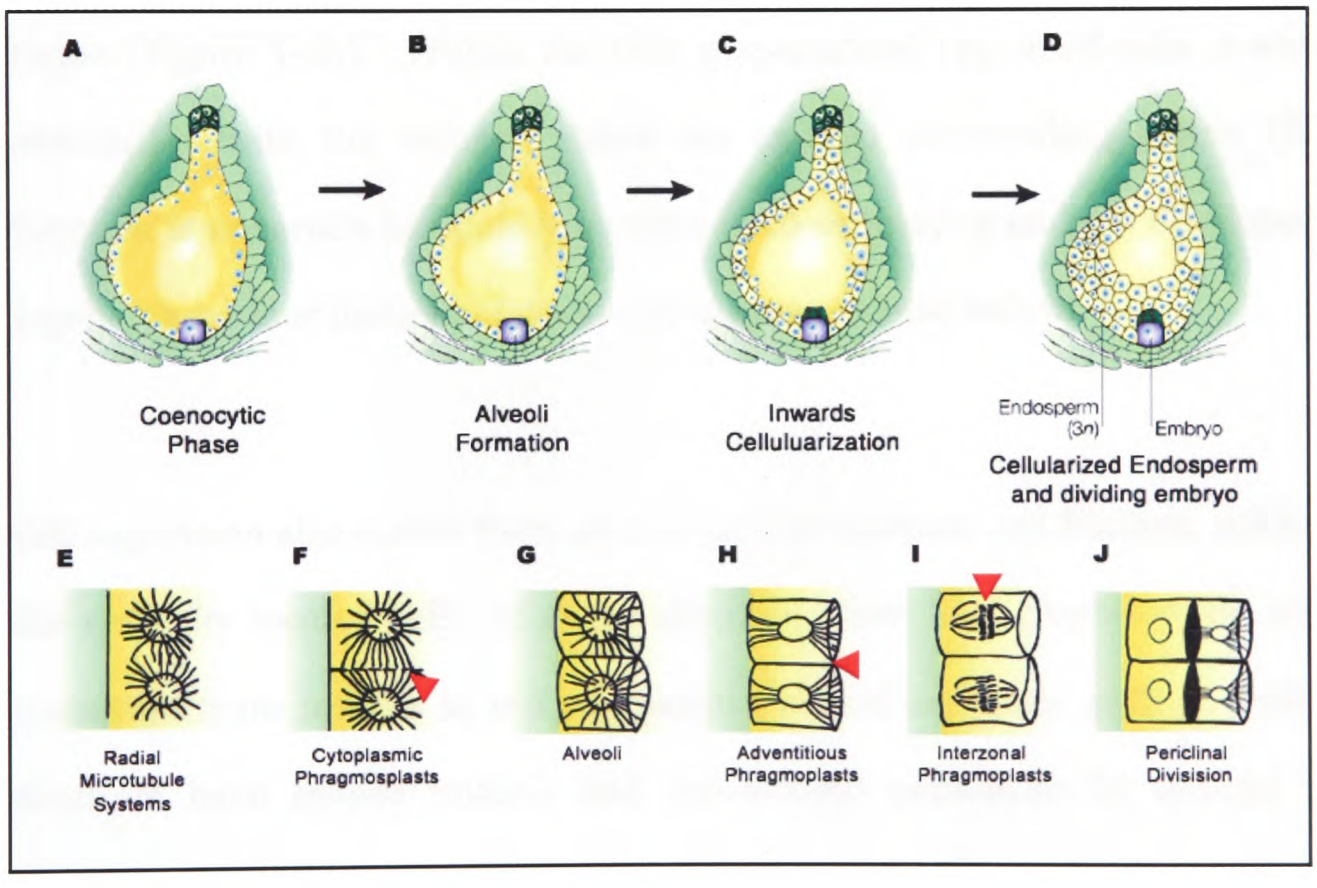


Figure 1-4 Cellularization of maize endosperm

Panels **A**-**D** depict gross changes in the cellularizing endosperm, whereas **E**-**H** depict details of cellular events. (**A**,**E**) Following the initial cell divisions in maize, which occur in the absence of cell division, a syncitium is formed, in which the nuclei form a layer around the periphery of the future endosperm. During this phase RMSs form between adjacent nuclei. (**B**,**F**-**G**) Cellularization commences with the deposition of cell wall material via cytoplamsic phragmoplasts (**F** - red arrow) from the edge of the coenocyte inwards forming alveolar structures of tubes of cell walls separating individual nuclei. (**C**, **H**-**J**) Anchored to the coenocyte cell walls via adventitious phragmoplasts (**H**- red arrow), the alveolar nuclei undergo a periclinal cell division involving interzonal phragmoplasts (**I** – red arrow) to form a peripheral layer of fully cellularized nuclei and a new inner layer of alveoli. (**D**) After several rounds of centrifugal cellularization the endosperm is fully cellular. (Adapted from (Coe, 2001)).

The post-cellularization phase is characterized by expansion of the endosperm through anticlinal cell division of centrally located cells, creating files of descendant cells (Lopes and Larkins, 1993), with simultaneous predominantly periclinal cell division of the outer layer, to keep pace with the increased surface area (Randolph, 1936). During this phase a process of differentiation also occurs, in which the four main cell domains are defined: the basal transfer layer (BETL), the embryo surrounding region (ESR), the aleurone, and the central starchy endosperm (CSE). The basal endosperm transfer layer occurs in the peripheral-most layer, adjacent to the site of contact with the maternal conducting tissue (Figure 1-5c), and is held to have a role in nutrient uptake. The aleurone layer is continuous with the BETL, forming an "epidermal" layer around the remainder of the endosperm, and is implicated in seed germination processes (Figure 1-5a). The CSE forms the bulk of the seed, and is the centrally-located storage tissue (Figure 1-5b). Within the CSE a specialized region of cells is found in close association with the embryo, called the embryo surrounding region (ESR), whose function is uncertain but may have roles in co-ordinating embryo development (Figure 1-5e). These four tissues are discussed in greater detail below.

Cell expansion also occurs from around 12 DAP (Kowles and Phillips, 1988), starting in the centrally located cells, at about the time when they cease mitotic activity. This transition from mitosis to cell expansion radiates outwards until all cells except the aleurone have ceased mitosis and commenced expansion by around 20-25 DAP (Kowles and Phillips, 1988). During this phase of cell growth the endosperm begins acquisition and storage of nutrient reserves, predominantly in the form of carbohydrate (starches) and storage proteins (such as zeins) (reviewed in (Soave and Salamini, 1984). Furthermore, the CSE cells undergo a process of endoreduplication, in which DNA is replicated, but without the separation of sister chromatids, to cause a massive increase in the DNA content of these cells (reviewed in (Kowles and Phillips, 1988)).

At later stages in development (after 30 DAP), undergoes profound metabolic changes. This phase is characterized by programmed cell death of centrally located cells (Young et al., 1997), and the accumulation of the hormone abscissic acid (ABA), which represses embryo germination and promotes desiccation tolerance on the embryo (Lopes and Larkins, 1993).

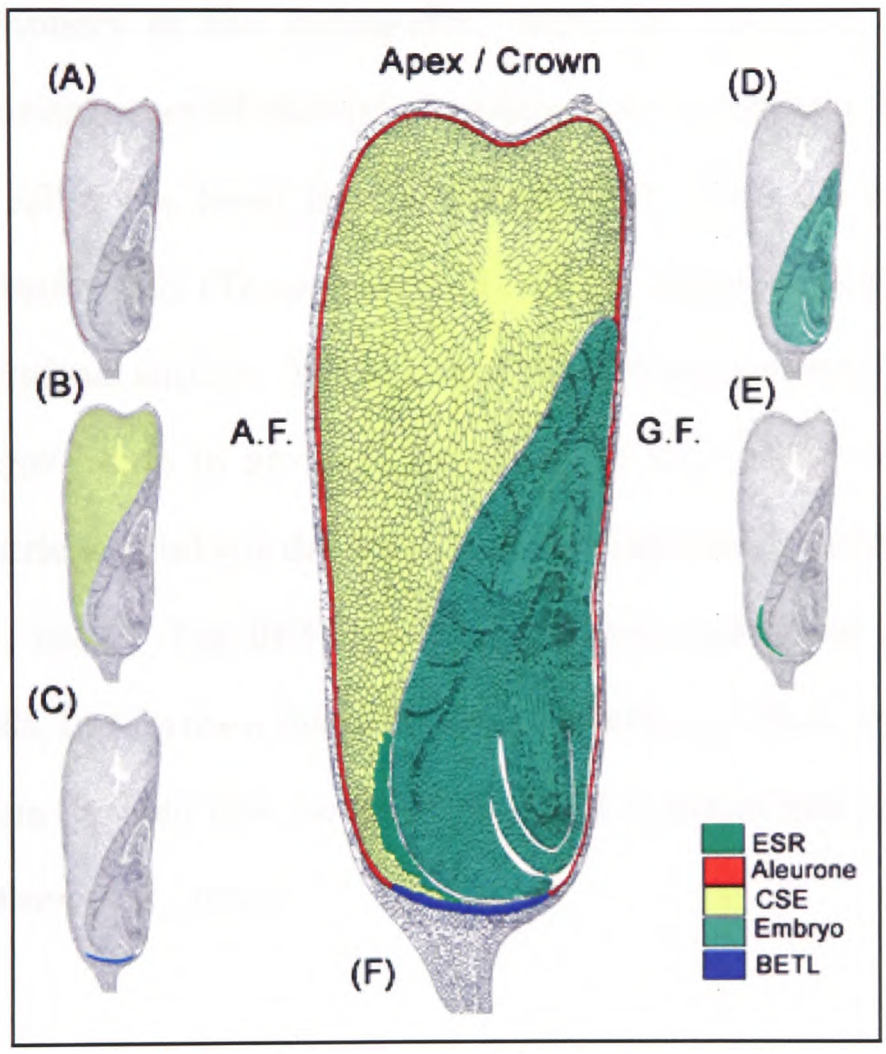


Figure 1-5 Endosperm domains in maize

Cartoon demonstrating the location of different endosperm highlighted in different colours (red = aleurone (**A**), yellow = central starchy endosperm (**B**), blue = basal endosperm transfer layer (**C**), green = embryo (**D**), light green = embryo surrounding region (**E**) A.F. = Adgerminal face, and G.F. = germinal face of the kernel. Adapted from (Kiesselbach, 1949a).

1.2.4.1 The basal endosperm transfer layer (BETL)

The seed is a metabolite sink, and obtains significant amounts of nutritional resources from the maternal parent. The endosperm performs the role of acquiring and storing reserves of carbohydrate and protein, it has however no plasmodesmatal connections to the maternal plant (although many plasmodesmatal connections exist between internal cell walls) (Davis and Smith, 1990; Soave and Salamini, 1984). The uptake of these reserves into the endosperm occurs through highly specialized cells located at the basal periphery of the endosperm, above the maternal pedicel, or chalazal end, where nutrients carried through the phloem are unloaded (Thompson et al., 2001). This zone is called the basal transfer layer (BETL), and the cells of this region are typical of transfer cells (Thompson et al., 2001), displaying highly invaginated cell walls on the periclinal surface. The enlarged cell surface increases the area over which transport can occur by up to 20-fold (Thompson et al., 2001). Also indicative of their activity in nutrient uptake is the massive accumulation of mitochondria in these cells (Charlton et al., 1995). The BETL occupies the same outermost endosperm layer as the aleurone cells, but forms a distinct tissue, the cells of which are morphologically quite different from the aleurone, however, the BETL is sometimes referred to as a 'modified aleurone' (Olsen et al., 1992).

The BETL is first cell type to become histologically distinct (Becraft, 2001) and is apparent from 6 DAP. As the kernel grows the BETL domain broadens and extends up to three cell layers deep (Schel et al., 1984a), the mechanism of which has been proposed to be by a lineage-dependent clonal mode of inheritance (Costa et al., 2003).

A number of genes which are expressed specifically in the BETL domain have been identified and are summarized in Table 1. BETL genes were identified as being specifically expressed in the basal portion of the endosperm (Hueros et al., 1999b;

Hueros et al., 1995), and are expressed during mid-term development (between 8-20 DAP). They lack significant similarity to other known sequences, although all have cysteine-rich regions, and *BETL1* and *BETL3* show some weak similarity to defensins. BETL2/BAP2 (basal layer type antifungal protein) has been demonstrated to have antifungal properties, and shows some sequence homology with *BAP1* and *BAP3* which are also specifically expressed in the BETL layer (Serna et al., 2001). The BETL2/BAP2 protein is initially synthesized as a propeptide, located in the endomembranes of the BETL, but its processed form is found in the placento-chalazal region of the pedicel.

Another gene has been shown to be expressed specifically to this area is *Maternally Expressed Gene 1* (*MEG1*) (Gutierrez-Marcos et al., 2004), which is also produced as a propeptide, and is cleaved and gylcosylated. MEG1 is localized to the cell walls, and like the BETL genes, has limited similarity to other known sequences and is also cysteine-rich. Interestingly *MEG1* expression shows parent of origin expression, only being expressed from the maternal allele from 10 DAP.

One of the most interesting genes to be identified to date is *MYB-Related Protein* (*ZmMRP*) This sequence is also expressed in the basal portion of the coenocytic endosperm prior to its expression in the BETL during later endosperm development (Gomez et al., 2002). *ZmMRP* encodes a MYB-domain transcription factor which has been shown to transactivate the promoters of some of the other BETL-specific genes discussed above (Gomez et al., 2002; Gutierrez-Marcos et al., 2004), leading to the proposal that the ZmMRP protein may be involved in regulating differentiation of the BETL. Should this be so, it may well be that the transfer layer is already specified before cellularization, which has already proposed for barley, where the *END1* transcript is also found in the chalazal region prior to cellularization (Doan et al., 1996).

Two related genes *ZmEBE1* and *-2* (*Zea mays embryo sac/basal endosperm transfer layer/embryo surrounding region*) are interesting as they show expression in the BETL and ESR, in addition to being expressed in the central cell prior to fertilization (Magnard et al., 2003). These genes have only weak similarity with other genes in sequence databases, and also show a cysteine-rich region, and a potential anti-microbial role has been ascribed to them.

A cell wall invertase gene (*INCW2*), has been shown to be expressed specifically in the BETL region (Cheng et al., 1996). The biochemical function of hydrolysing sucrose into fructose and glucose is thought to be important in maintaining a sucrose gradient to maintain the flux on metabolites into the pedicel. Mutants in this gene show defective development of the pedicel, although a normal BETL layer forms (Miller and Chourey, 1992). Similarly, a seed mutant which results in small kernels, *reduced grain filling* (*rgf1*), results in a decrease in the amount of BETL1 and BETL2 proteins, even though the transfer layer appears essentially normal (Maitz et al., 2000), suggesting they are not essential for BETL cell function.

Interestingly, restricted development of the BETL domain occurs in the maize seed development mutant, *globby1-1*, which also has a 'miniature kernel' phenotype (Costa et al., 2003). This gene is thought to have a role in cytokinesis and cell division throughout endosperm development.

Normal development of the transfer layer can be distorted in at least one direction of interploidy cross, (paternal excess), with cells lacking the normal extensive cell wall invaginations (Charlton et al., 1995). However, it has also been suggested that following similar paternal excess crosses the *BETL2* transcript is found in the BETL in punctuate domains along the base of the endosperm, and cells located within these

domains displayed typical transfer-like morphology (Hueros et al., 1999b). An earlier, but more detailed survey using light microscopy showed that both interploidy directions resulted in delayed or reduced transfer layer development (Cooper, 1951). This is described in more detail in section 1.4.1.

Gene Name	Timing of Expression in Endosperm (DAP)	Location of expression	Predicted Size (kD)	Sequence Similarities / Features	Proposed Function	Reference
BETL1	8-22	BETL – cell wall	7	Defensin (weak) Cysteine-rich	Defence, structural, mechanical	(Cai et al., 2002; Hueros et al., 1995)
BETL2 / BAP2	6-22	BETL – membrane / microtubules (propeptide) Placento-chalazal region (cleaved)	8(propeptide) 4.5 (cleaved)	Cysteine-rich	Antifungal	(Cai et al., 2002; Hueros et al., 1999b; Serna et al., 2001)
BETL3	8-16	BETL		Defensin (weak) Cysteine-rich	Defence	(Hueros et al., 1999b)
BETL4	11-14	BETL, not cell wall	8	Cysteine-rich		(Cai et al., 2002; Hueros et al., 1999b)
BAP1a	10-24	BETL		Cysteine-rich		(Serna et al., 2001)
BAP3a	10-18	BETL				
ZmMRP	3-17	Basal coenocyte, BETL – nuclear	27.5	MYB-related transcription factor	Domain specification Activation of other BETL genes	(Gomez et al., 2002)
ZmEBE1	3-12	BETL, ESR	33.4	DUF23 family (arabidopsis)		(Magnard et al., 2003)
ZmEBE2	1-15	Central cell, BETL, ESR	31.3			
MEG1	4-20	BETL - cell wall ingrowths	9.7propeptide 6.7 cleaved	Cysteine-rich	Structural, signalling, defence	(Gutierrez-Marcos et al., 2004)
INCW2		BETL –cell wall ingrowths	72	Cell wall invertase	Sugar hydrolysis/ sucrose partitioning	(Cheng et al., 1996)

Table 1 Summary of BETL expressed genes.

1.2.4.2 The embryo surrounding region (ESR)

A specialized region of cells is located adjacent to the cavity in which the embryo is located and is termed the embryo surrounding region (ESR). It is characterized by cells of a compact size and dense cytoplasm, and containing high levels of rough endoplasmic reticulum (Becraft, 2001; Bonello et al., 2000; Schel et al., 1984b). As the embryo develops and grows, this region, which does not expand, remains localized at the base of the embryo and around the suspensor. The ESR is thought to be the first cell domain to form, and it appears to already be cellular whilst the majority of the endosperm remains coenocytic. The function of this region is not clear, although obvious roles have been proposed: in embryo nutrition; as a transfer zone between the embryo and endosperm; and in protection of the developing embryo (Olsen, 2001).

A number of genes specifically expressed in the ESR are detailed in Table 2. A family of related genes, *ZmESR1-3* (Embryo Surrounding Region) (Bonello et al., 2000; Opsahl-Ferstad et al., 1997, Bonello, 2002) are expressed in the ESR from 4 to 28 DAP. The function of these proteins is unclear, as there are no similar sequences in the databases, although a conserved critical 15 amino acid motif ([DS][RE][KL]R[TP]VPSGP[DN][HP][MIK]H[HN]) of apparent functional importance is also found in the *Clavata3* gene (Bonello et al., 2002). Additionally, the small size of the proteins, their generally hydrophilic nature, and the fact they are secreted and cell wall associated suggests they may be involved in signalling - possibly involved in embryo-endosperm communication. Instructively, spontaneous embryoless kernels do not show expression of the *ZmESR1*, so formation of the domain may depend on signalling from the embryo (Opsahl-Ferstad et al., 1997).

Two related genes *ZmAE1* and *ZmAE2* (androgenic embryo), are also located specifically in the ESR, although, unusually, they were originally isolated as transcripts

which were expressed in androgenic embryos derived from pollen grains (Magnard et al., 2000). These two genes lack sequence similarity with other ESR genes, and also other genes in the databases, however, their common location and timing of expression suggests they may be involved in a similar process as the other ESR transcripts.

Another gene, *ZmInvIn* has also been shown to be specific to the ESR, and is expressed for a limited time in early development (Bate et al., 2004). This gene is an inhibitor of invertases, and shown to be associated with an apoplastic invertase, so could be involved in compartmentalizing invertase activity to match the different needs of the embryo and endosperm.

Gene Name	Timing of Expression in Endosperm (DAP)	Location of expression	Predicted Size (kD)	Sequence Similarities / Features	Proposed Function	Reference
ESR1,2,3	4-28	ESR – secreted, cell wall	~14	Clavata3, Hydrophilic	Signalling, pathogen defence	(Bonello et al., 2000; Bonello et al., 2002; Opsahl-Ferstad et al., 1997)
ZmAE1 ZmAE3	3-12 4-12	ESR and androgenic embryos	10-20	Remote, EGF-like cysteine domain.	Signalling	(Magnard et al., 2000)
ZmInvIn	4-7	ESR – cell wall	17.7	Invertase inhibitor	Resource partitioning	(Bate et al., 2004)

Table 2 Summary of ESR expressed genes

1.2.4.3 The aleurone layer

The aleurone is the outermost layer of the endosperm and is one cell thick in maize, forming a periphery overlying the CSE cells. In some cereals the aleurone comprises more layers of cells (three in barley, one to several in rice) (Olsen, 2004b). Furthermore, maize mutants have been described which also have multiple layers, suggesting that control of aleurone depth has a genetic basis (Lid et al., 2004; Shen et al., 2003; Wolf and al, 1972). The aleurone consists around 250,000 cells (Walbot, 1994), characterized by their regular, cuboidal shape. Aleurone cells are

cytoplasmically dense, being rich in endoplasmic reticulum, mitochondria, and aleurone grains which are small vacuoles with inclusion bodies containing digestive enzyme precursors (Olsen, 2004b). Following a late-stage developmental program that confers desiccation tolerance on the aleurone cells, this cell domain is the only one to survive after maturation (McCarty and Carson, 1991), and is re-activated during germination, in which it plays a critical role. Imbibition of water by aleurone cells activates the digestive pro-enzymes (Becraft, 2001), which are released into the central endosperm to release nutrients the embryo will utilize in development.

The aleurone is first histologically distinct between 10 and 15 DAP (Becraft and Asuncion-Crabb, 2000), although its specification begins sooner, and is characterized by cell-type specific transcripts such as *Vpp1* and *Ltp2* from 8 DAP (Wisniewski and Rogowsky, 2004).

It was originally thought that aleurone fate was determined by lineage, largely because cells located in this layer divide predominantly anticlinally after 21 DAP (Randolph, 1936). However, genetic analyses demonstrated that aleurone and cells of the CSE can share a common lineage (Becraft and Asuncion-Crabb, 2000). This critical study also revealed that neither aleurone nor CSE cell identity is a terminal cell fate. Ds-induced chromosome breakage in heterozygotes for the *dek1* mutation revealed that aleurone identity could be lost (and replaced with the mutant starchy-like morphology), up until the last cell division. Additionally, Becraft revealed transposition of the Mu element out of a *dek1:Mu* insertion allele restored the gene function and regenerated an aleurone identity. Hence, aleurone identity depends on a continuous signalling involving the DEK1 gene, a 21- loop transmembrane protein with a cytosolic region similar to a calpain-like Cys proteinase domain (Lid et al., 2002).

Another protein which appears to be important in the specification of the aleurone cell fate is the product of the *Crinkly4* gene (*Cr4*), mutants of which lack an aleurone (Becraft et al., 1996). The protein is a cell-autonomously acting receptor tyrosine kinase with homology to tumour necrosis factor receptors (TNFRs) of mammals acting in both aleurone and epidermal cells (Becraft et al., 2001). TNFRs are activated by binding a homotrimeric TNF-α cytokine complex and it is hypothesized that Cr4 is also likely to be activated by a currently unidentified ligand which might originate from the maternal cell wall, to promote the aleurone developmental pathway (Olsen et al., 1998).

Intriguingly, aleurone identity has been shown to occur in 'external' faces of the endosperm, even in aberrant development, where cell surface invaginations, and internalized cisternae have created 'ectopic' edges leading to a 'surface rule' hypothesis for its determination (Costa et al., 2003; Olsen, 2004a).

1.2.4.4 The starchy central cells

Starchy cells form the central mass of the endosperm and it is thought that they function primarily in the storage of nutrients in the form of prolamine storage proteins, including zeins and globulins (Song et al., 2001; Washida et al., 2004), and starch (Lopes and Larkins, 1993). The low oxygen concentration characteristic of the CSE is important in partioning of reserves through the endosperm (Rolletschek et al., 2005). In spite of the fact that this is the most numerous cell type, very little is known about its mechanism of specification (Olsen, 2004b), although the observation of starchy endosperm cells replacing aleurone cells in the mutants *dek1* and *crinkly4*, suggests that they may be a default cell type in the endosperm (Becraft and Asuncion-Crabb, 2000; Becraft et al., 1996). Cells of the CSE are descendants of the internally located nuclei produced following cellularization, and their nuclear division occurs apparently

at random planes, which explains why the cells in this region do not display files of cells radiating from the centre of the endosperm.

One distinctive characteristic of the CSE, is that these cells undergo endoreduplication (DNA replication without mitosis), increasing the DNA content of the cells to up to 96C starting from around 12 DAP (Kowles and Phillips, 1988). It is unclear exactly why these cells should undergo endoreduplication, but it could be to increase the transcriptional capacity for the genes which are expressed in these cells (Larkins et al., 2001), although recent works suggests endosperms with decreased endoreduplication show normal levels of gene expression of storage molecules (Leiva-Neto et al., 2004). An alternative proposal for the function of endoreduplication is as a storage mechanism of nucleotides (Lopes and Larkins, 1993). It is suggested that endoreduplication occurs through the loss of activity of M-phase cyclin-dependent kinases (CDK), and oscillations of S-phase CDK activity (Larkins et al., 2001).

Later stages of endosperm development are characterized by programmed cell death, and nuclear disintegration, which might be a mechanism of making nutrient resources more readily available to the embryo when it germinates (Young and Gallie, 2000).

1.2.5 Development of the embryo

Fertilization of the central cell and egg cell occur simultaneously, but the development of the endosperm proceeds much more rapidly than that of the embryo. The first zygotic division occurs after 10-12 hours after fertilization, in comparison with between three and five hours for the primary endosperm cell (Sheridan and Clark, 1994). Embryogenesis then proceeds in three distinct phases: an initial formation of polarity; establishment of an axis and the shoot and root meristems; followed by a maturation

phase in which the primordial leaves are formed prior to the dehydration which precedes dormancy (Figure 1-6).

The first cell division is predictable in the maize embryo, occurring perpendicular to the long axis of the megagametophyte (Figure 1-6-A). It produces two distinct cells: at the top a small cytoplasmically dense cell which will form the majority of the embryo (Figure 1-6-A, coloured orange) and a larger, vacuolated cell at the bottom which will contribute to the suspensor (Figure 1-6-A, coloured yellow) (Randolph, 1936). Although in arabidopsis many of the initial cell divisions contributing to the embryo are predictable (Bommert and Werr, 2001), in maize subsequent divisions are not. Following several cell divisions an initial structure starts to form, and by 100 hours after pollination around 12-24 cells at the apex form a hemisphere (Figure 1-6-B-C). This structure enlarges and increases in cell number to form a club-shaped structure (Figure 1-6-D), which expands, producing the transition phase embryo (Figure 1-6-E).

The second major phase of embryo growth starts with the initial definition of an apical-basal polarity. A subset of cytoplasmically dense cells become defined within the anterior face of the embryo (Figure 1-6-E, highlighted in green). Shortly afterwards a similar region develops on the opposite side of the embryo (Randolph, 1936). These two regions expand and join to form a region of higher meristematic activity which defines the future shoot- root meristem axis, which is slightly acute to the embryonic apical basal axis, around 11-13 DAP. During this time the proembryo elongates and enlarges to form an axis and, concurrently, the coleoptilar ring and scutellum form to generate the so-called coleoptilar stage embryo (Figure 1-6-F). By 16 DAP the first leaf primordium also appears as a rim around the lower edge of the SAM (Figure 1-6-G).

The final phase of embryogenesis involves the development of further leaf primordia (up to six), alternating between the lower and upper side of the SAM (Figure 1-6-I and J). Also during this phase the embryo undergoes metabolic changes that enable it to experience the dehydration phase which occurs prior to dormancy (Sheridan and Clark, 1994).

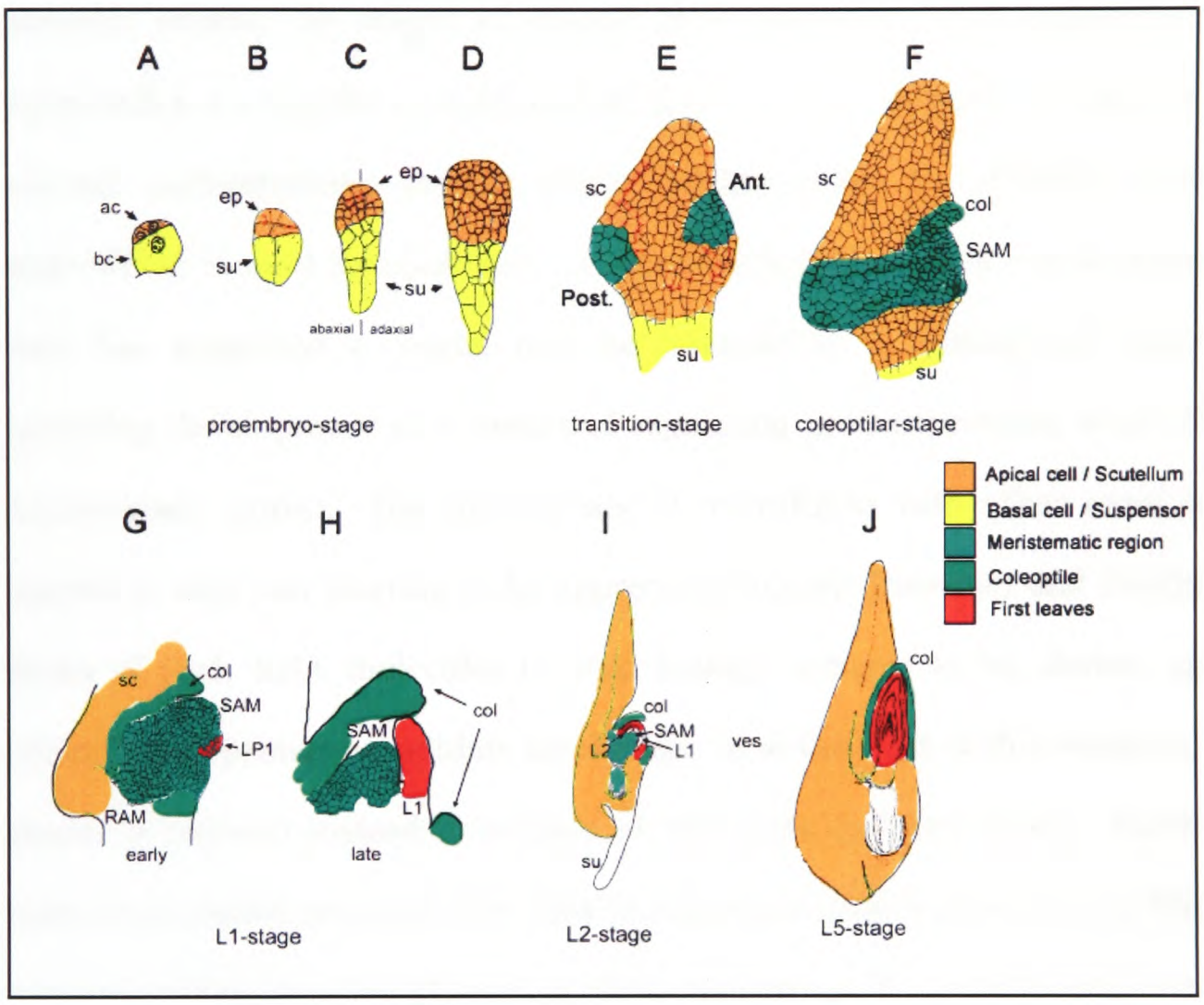


Figure 1-6 Embryogenesis in maize

Details the different phases of embryo development from the initial cell division (**a**), through to formation of the proembryo stage and development of the suspensor (**b**-**d**), the first appearance of polarity at the transition stage (**e**), and the origination of the shoot apical meristems at the coleoptilar-stage (**f**), and the root meristem by the L1 stage (**g**). ac: apical daughter cells; bc: basal daughter cell; ep: embryo proper; su: suspensor; sc: scutellum; col: coleoptile; SAM: shoot apical meristem, RAM: root apical meristem, LP1: first leaf primordium; L1 + 2: first and second leaf. Different stages are not drawn to scale. Ant = anterior face, Post = posterior face. Adapted from (Bommert and Werr, 2001).

1.3 Control mechanisms of seed development

Molecular control of seed development by inherited factors is a complex, multi-faceted and still relatively poorly understood area of research. Regulation is likely to occur at a series of genetic and epigenetic levels, and an awareness of these phenomena is critical to the interpretation of the effects that follow interploidy crosses, since any of these processes may be disturbed. This section will give a brief summary of some of the main modes of control that are currently known, including normal genetic inheritance, parental effects, the effects of dosage and imprinting, and epigenetic phenomena. Epigenetics is a rapidly expanding field and too broad a theme to describe a complete current understanding in this thesis. Additionally, the possible involvement of microRNAs in seed development , is acknowledged here, as a rapidly growing body of data has suggested a crucial role for microRNAs in widespread areas of biology including development as a means of regulating gene expression levels (Vaughn and Martienssen, 2005). The convergence of microRNAs with other areas of epigenetic control is only just starting to be appreciated (Grant-Downton and Dickinson, 2005). Roles of such RNA molecules in seed biology remains to be shown, and were not considered a primary candidate mechanism of at the start of this research, as such the reader is referred instead to reviews on this topic as listed above. Furthermore, the even more recent proposal that RNA-based intermediates may act as a form of genetic memory acting in a non-Mendelian fashion (Lolle et al., 2005) serve to remind us of how naïve our understanding of inheritance and transcriptional control may be. For this reason, it seems prudent to acknowledge, that there may yet remain other genetic and epigenetic modes of inheritance pertinent to seed biology of which we have, as yet, no knowledge.

1.3.1 Parent-specific expression in the embryo sac

It is widely reported that the maternal genotype is critical in determining the phenotype of progeny seed in maize. Crosses between large seed varieties and small seed varieties resemble the maternal parent, whereas the F1 seeds are intermediate between the two (Birchler, 1980). However, a role is played by the paternal genome, since paternally derived gene copies can rescue the phenotypes of mutant maternal copies (Costa et al., 2003), and the differential effects of interploidy crosses necessarily imply a paternal involvement (Scott et al., 1998). However, the maternal genome must be the most important contributor to seed development, especially in the early stages of development. It has been reported that in early seed development in arabidopsis gene transcription occurs exclusively from the maternal genome at least for the first three to four days (Vielle-Calzada et al., 2000), although this issue remains contentious (Vielle-Calzada et al., 2001; Weijers et al., 2001). A similar general delay of paternal genome activity is also reported for maize, for a period of three to four days (Grimanelli et al., 2005), although another study suggested that paternal expression of a transgene occurs almost immediately after nuclear fusion (Scholten et al., 2002).

The nature of the plant life-cycle, in which there is a sporophytic and gametophytic phase, means that the maternal parent can influence the seed development through sporophytic acting genes which determine the development of the embryo sac and embryo, and also from the haploid genome set present in the female gametophyte. Additionally, cytoplasmic inheritance is determined by the maternal rather than the paternal parent. Mutations have been identified acting at each of these stages (Chaudhury et al., 1997; Colombo et al., 1997; Grossniklaus et al., 1998; Hulskamp et al., 1995; Luo et al., 1999).

Several genes which have parent of origin effects and are critical for seed development in arabidopsis were identified through separate screens for seed development without fertilization, and for seed abortion (Chaudhury et al., 1997; Grossniklaus et al., 1998; Ohad et al., 1996). Three mutants displayed not only autonomous seed development without fertilization, but also perturbed and over-proliferative development when ovules were fertilized. Such over-proliferative development was characterized by large ectopic chalazal cysts. All three genes were maternal effect mutants, which could not be rescued by a functional paternal copy of the gene. The three genes identified were homologues of drosophila Polycomb group genes which control memory of transcriptional states through gene repression (Francis and Kingston, 2001). *MEDEA* (*MEA*) is a homologue of the drosophila Enhancer of Zeste (*E(Z)*) gene (Grossniklaus et al., 1998; Kiyosue et al., 1999). FIE encodes a WD40 protein similar to drosophila Extra sex combs (*ESC*) (Ohad et al., 1999). *FIS2* is a zinc-finger containing protein similar to drosophila Supressor of zeste (12) (*Su(z)12*) (Luo et al., 1999). FIE and MEDEA proteins have been demonstrated to interact *in vitro* (Luo et al., 2000; Spillane et al., 2000; Yadegari et al., 2000), as would be predicted by their homologues in drosophila working in a functional repressive complex (Tie et al., 2003).

1.3.2 Dosage effects

Dosage refers to the concept of a gene whose level of expression (as measured by mRNA or protein) is dependent on the physical number of copies of that gene present. Dosage-dependent effects of genes are thought operate in the endosperm, which normally contains three copies of each gene in the endosperm, although after the initiation of endoreduplication this number increases many-fold (Birchler, 1993). Clearly, dosage dependent expression of genes from the diploid embryo and the triploid endosperm could account at least in part for the very different developmental fates these two organisms follow, in spite of their identical genetic make-up.

It was once held that most genes show a direct correlation between the number of gene copies present in a cell and the level of expression (Birchler, 1993), more recent work suggests a much more complex picture. On varying the number of copies of chromosomal arms contained in the nucleus, different effects are seen between those which are physically located on the varied fragment (direct effects), and other genes located genes elsewhere in the genome. Genes located on the varied arm often displayed dosage compensation, the maintenance of a constant level of total expression for a gene in spite of a change in copy number (Birchler, 1979; Birchler and Newton, 1981). Most surprisingly was the effect on regions of the genome that were not on the altered chromosomal arm which showed a number of different effects: although predominantly inverse correlations between expression levels and the number of the varied chromosomal fragments were observed, direct and static levels were also seen for some genes in an unpredictable manner (Birchler and Newton, 1981; Guo and Birchler, 1994). However, on varying the total number of genome sets, the majority of genes assayed showed a level of expression proportional to the genome number (Guo et al., 1996). These effects appeared to suggest that stoichiometric interactions between regulatory genes were critical in maintaining normal development, and that if their balance was distorted in an aneuploid condition they caused altered expression elsewhere in the genome. Several such genes were identified, and involved a relatively limited subset of regulatory genes including transcriptional regulators and chromatin remodelling proteins (Birchler et al., 2001, Birchler, 2005). A review of the potential biological implications of dosage effects on evolution is provided in (Birchler et al., 2005).

1.3.3 Imprinting

Genomic imprinting is a phenomenon in which the level of expression of a certain allele is dependent upon whether it originates from the mother or from the father, in contravention to the Mendelian rule of equivalence of parental genomes. Such parent-dependent expression occurs between widely phylogenetically distributed range of organisms, including the angiosperms maize and arabidopsis; and the mammals humans and mouse. Although most research into imprinting and evolutionary theories for its existence originates from research in these species, the distribution of imprinting is thought to be more widespread than this (Alleman and Doctor, 2000).

Parent of origin phenomena have been known in certain species for a long time, although these involved the selective switching off of an entire chromosome, or genome complement, often from the father, such as in the X-inactivation of marsupials (Sharman, 1971), or the turning off of the paternal genome in rodent extra-embryonic tissue (Tagaki and Sasaki, 1975). The first example of imprinting of a single gene, however, was in maize, and involved a gene involved in the pigmentation of the aleurone layer of the seed (Kermicle, 1970). A certain allele, R-r:std, produces a complete purple aleurone when inherited through the mother, but patchy incomplete pigmentation if inherited through the father. Dosage effects could be ruled out by the fact that the addition of an additional paternal copy via a B chromosome could not rescue the phenotype (Kermicle, 1978).

The most important discovery relating to the importance of imprinting in development was in mice, where parthenogenetic embryos were created containing either two maternal or two paternal genome copies (Surani et al., 1984). Embryos from both crosses aborted, but those created from maternal genomes possessed reduced extra-

embryonic tissue, whereas embryos created from the paternal genome also aborted, but with an overgrown placenta.

In contrast to mammals, imprinting does not seem to be critical for embryonic development in plants, as embryos can be recovered from various tissues, including haploid plants derived from gametes (Sarkar and Coe, 1971). Furthermore the existence of apomixis, in which seed development occurs without fertilization of the embryo confirms that imprinting is not essential for embryonic growth (Grossniklaus et al., 2001). However, imprinting effects have been shown to be important for seed development, but all imprinted loci identified to date are active in the endosperm. The R-locus gene operates in the aleurone, the peripheral cell layer of the endosperm (Kermicle, 1970), and other imprinted genes include the zein family of genes, important for protein storage in the endosperm (Lund et al., 1995a), and a regulator of accumulation of zein proteins (Chaudhuri and Messing, 1994). These imprinted genes only display allele specific imprinting, so only certain genotypes display these effects.

Imprinted loci in maize were deduced by the work of Lin, who showed loss of chromosomal fragments of the long arm of chromosome 10 (10L) from the female parent of the seed resulted in reduced grain sizes which could not be compensated for by paternally inherited copies of the same fragment, which he inferred was due to maternally expressed imprinted genes (Lin, 1982). Additionally, Lin's work on *indeterminate gamete* (*ig*) mutants, which produce central cells with a range numbers of nuclei from one to seven (Lin, 1984), to alter parental genome ratio, showed whole genome parental effects. Two classes of seed contained four genomes, those with three maternal and one paternal genome which were smaller than normal seeds, but still viable, and those containing two paternal and two maternal genomes which were abortive (Lin, 1984). These studies revealed that a deviation from the two maternal to

one paternal ratio resulted in seed abortion and failure of endosperm development. The net effect of all imprinted genes can be observed by altering the balance between maternal and paternal genomes which contribute to the endosperm, as achieved by crossing plants of different ploidy (interploidy crosses), as discussed in section 1.4.

The *MEDEA* gene, discussed above under parental effects has been demonstrated to show locus imprinting (Kinoshita et al., 1999), being expressed from the maternally inherited copy. The activity of the gene is regulated by another maternal effect mutant, *DEMETER* (*DME*), which is a DNA glycosylase, which is expressed in the early ovule, but not in the male germline, and is proposed to activate *MEA* by a base-excision mechanism, possibly through the removal of 5-methylcytosine (Choi et al., 2002).

Maize has two homologues of *FIE*, called *ZmFie 1* and *2*, which have different expression patterns, suggesting non-redundant roles (Danilevskaya et al., 2003; Springer et al., 2002). *ZmFIE1* is regulated by imprinting, as determined by restriction analysis of alleles from reciprocal crosses of different inbred lines. ZmFIE2 shows expression patterns most similar to those in arabidopsis, as it is expressed pre-fertilization, repressed in early kernel development, and then re-expressed by around 10 DAP.

1.3.4 Imprinting Mechanisms

The occurrence of imprinting poses a number of important questions. For example, what 'marks' are conferred on imprinted genes to reveal parental ancestry in their progeny, and how and when are these marks applied to the genes involved. Unfortunately the mechanisms of imprinting in plants remain poorly understood, and

most of our knowledge is derived from mammalian systems (Kaneko-Ishino et al., 2003; Murphy et al., 2003).

It is clear that epigenetic mechanisms are fundamental to the silencing of imprinted genes, but since the field of epigenetics itself remains quite poorly understood, which of these diverse mechanisms serve to generate parent-dependent expression is still unclear. The most well characterized modification associated with imprinted genes is DNA methylation status. In maize the *zeins* and *R1* genes show hypomethylation (Alleman and Doctor, 2000; Lund et al., 1995b). In arabidopsis homozygotes for *decrease dna methylation* (*ddm1*) which display reduced DNA methylation can rescue mutants of the *mea* and *fis2* genes presumably through activation of the normally silenced paternal functional copy (Luo et al., 1999; Vielle-Calzada et al., 1999).

The finding of a DNA glycosylase which is required to specifically activate maternal MEA transcription through demethylation suggests that in plants active demethylation is required to activate the expressed allele of an imprinted gene (Choi et al., 2002), whereas in mammals it seems that DNA methylation is necessary to silence the imprinted allele (Brannan and Bartolomei, 1999).

1.3.5 Evolutionary origin of imprinting

The existence of imprinted genes exposes the organism to a potential problem: for this locus it will be a functional haploid, which exposes the plants to the potential lethality of mutants which would be masked in the diploid condition (Wilkins and Haig, 2003). This might be actually a mechanism of eliminating deleterious mutations from the gene-pool (Walbot and Evans, 2003), especially if most imprinted genes function during fertilization and early development, in which case they are eliminated before

resources are wasted in creating a new individual. However, a more widely held explanation for the existence of imprinting is that it is the consequence of differential interest of the two parents over the allocation of resources from the mother to the progeny, the so-called 'parental conflict hypothesis' (Haig and Westoby, 1989; Haig and Westoby, 1991). In an out-crossing population, a father will not necessarily be related to the other kernels growing on the female, hence it has an interest in maximizing the amount of resources it can obtain from the mother, so the theory predicts that there will be a tendency for paternally expressed nutrient transporters and growth factors. In contrast the mother is equally related (50%) to all offspring she bears, so has an interest in equitably distributing these resources. The model predicts the evolution of female expressed, paternally silenced antagonists of paternally expressed growth promoters. The model is general for imprinted genes in plants and mammals, where there is a connection between the mother and her offspring, in which nutrients are passed: the placenta in mammals, and the endosperm in plants. Although some studies suggest directions of imprinting as predicted for growth enhancers (Dechiara et al., 1990), there are also contradictory observations (Hurst and McVean, 1997). However, the observations of interploidy crosses in arabidopsis are considered supportive of the theory as a paternal excess results in larger seeds and maternal excess results in smaller seeds (Scott et al., 1998). However, the failure of interploidy crosses in maize and the reported failure to form a nutrient transfer layer is apparently contradictory to the theory (Charlton et al., 1995; Gehring et al., 2004).

1.3.6 The role of chromatin remodelling factors in epigenetics

Central to the mechanisms of gene regulation reported to occur in the seed outlined in this chapter is epigenetic regulation, that is changes to the transcriptional state of a gene which is not due to the underlying DNA sequence, but to the modifications of the DNA, the proteins associated with it, or their modifications. These mechanisms alter

the conformation of the chromatin which determines transcriptional levels of the gene. Such phenomena will be described in greater detail in Chapter 5.1.2.

1.4 Interploidy Crosses

Perhaps the clearest demonstration of the different contributions to seed development provided by the male and female parents is provided by reciprocal crosses between plants of different ploidy, typically between diploids and tetraploids. These so-called interploidy crosses still offer the molecular biologist one of the most accessible ways of distorting the parental contribution to the seed, thereby allowing investigation of the different roles of inherited factors from either parent (most significantly the genomes, but also cytoplasmic constituents etc) have on the seed. Although an absolute understanding of these differential parental roles would be gained through the creation of endosperms and embryos comprising exclusively genomes from either parent, similar to in mice (Surani et al., 1984), this is not currently technically feasible in plants. However, interploidy crosses are straight-forward, requiring only the controlled application of diploid pollen onto a tetraploid ear and vice versa. Commonly, interploidy crosses lead to profound effects on the phenotype of the seed, which often abort. The two directions of cross frequently show alternate phenotypes, despite the genetic equivalence of their genomes.

Double fertilization in angiosperms means there are two main consequences of interploidy crosses that must be considered. The total number of genomes that ordinarily contribute to the embryo is altered (from two to three in both crosses), and the number of genomes that contribute to the endosperm nuclei is changed: from three in normal crosses to four in interploidy crosses where the male is the tetraploid parent (paternal excess), and to five when the mother is the tetraploid parent (maternal excess) (Figure 1-7).

Adding to the complexity, and of significance to the study of imprinting and other parental effects, is the altered balance of parental genomes. Normal endosperms comprise two maternally and one paternally derived genomes (2m:1p), and normal embryos comprise one genome from each parent (1m:1p). In interploidy crosses with paternal excess (2x × 4x), the endosperm has a 2m:2p genomic constitution, and the embryo becomes 1m:2p. For maternal excess crosses, the resultant endosperm is 4m:1p, and the embryo is 2m:1p. Another possible consequence of interploidy crosses is that the amount of cytoplasm contributed by the female parent is changed, as tetraploids are reported to have larger ovules than diploids in some species (von Wangenheim and Peterson, 2004); and the paternal cytoplasmic contribution is considered negligible. This might affect both mitochondrial and plastid inheritance, and also RNAs and proteins already synthesized and stored in the embryo sac (von Wangenheim and Peterson, 2004).

1.4.1 Interploidy crosses in maize result in seed abortion

Interploidy crosses in maize have been reported for several genetic backgrounds (Figure 1-8 and Table 3). The first study was by DC Cooper (Cooper, 1951); in a detailed histological survey from between two to 24 days after pollination (DAP) he observed the differences in gross morphology between the two interploidy crosses: maternal excess (4x × 2x) and paternal excess (2x × 4x), and two same ploidy controls: diploid (2x × 2x), and tetraploid (4x × 4x). These data are summarized in Figure 1-9 and examples of observations he made presented in Figure 1-8.

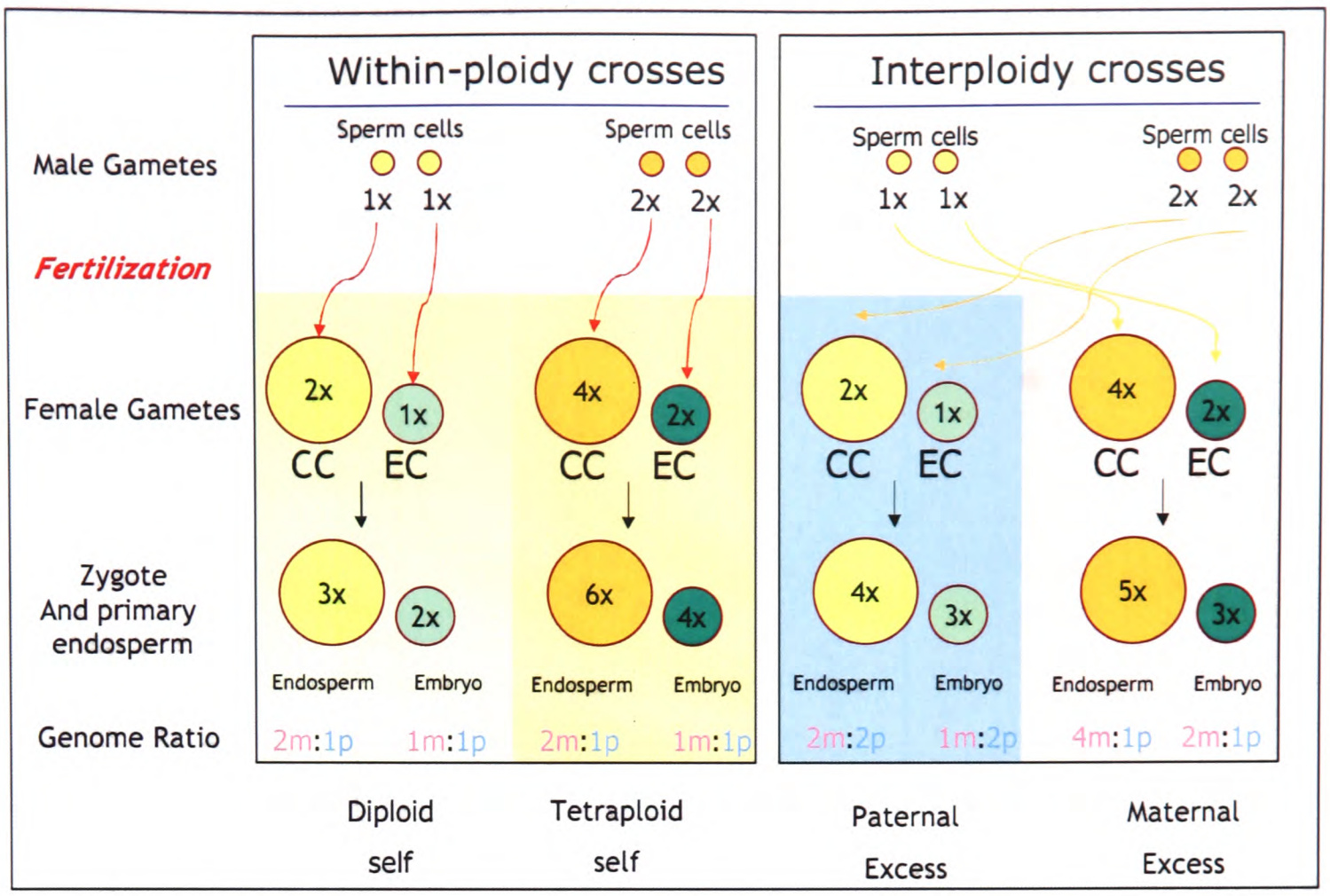


Figure 1-7 Demonstration of interploidy crosses in maize

The cartoon demonstrates the genomic consequences of interploidy crosses in maize. The left panel details the situation in normal genomically balanced crosses for diploids and tetraploids. The central cell (CC), has either two or four genome copies respectively, and the egg cell (EC) has one or two. In the diploid cross the resulting endosperm contains two maternally derived and one paternally derived genome, whereas the embryo contains one of each. In the tetraploid this is doubled. Interploidy crosses, in which pollen from diploids and tetraploids are crossed onto the ear of the other ploidy result in imbalanced chromosomes containing a paternal excess (2x x 4x), or a maternal excess (4x x 2x) of genomes.

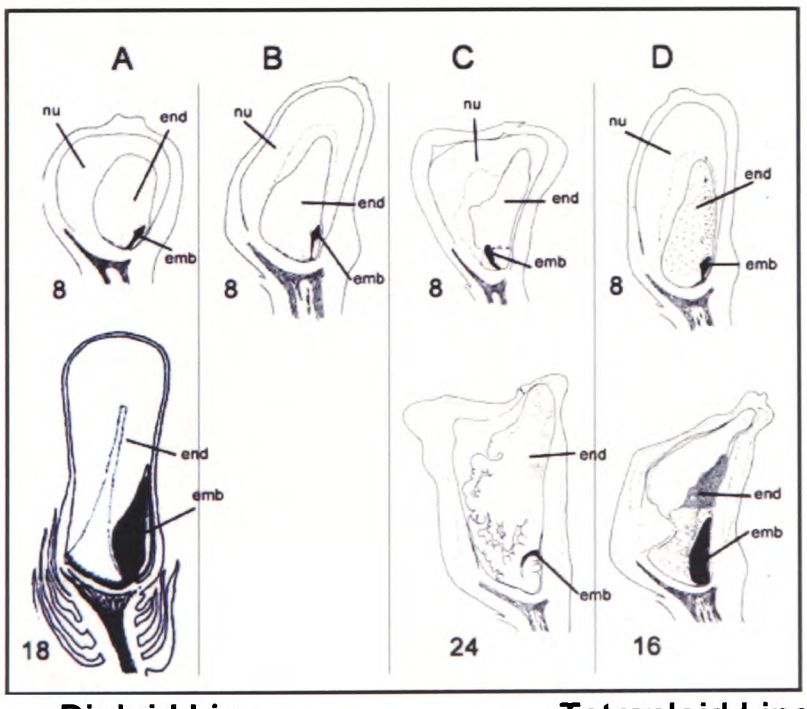


Figure 1-8 Images of kernels from balanced and interploidy kernels as reported by Cooper

Line drawings of sectioned material observed by DC Cooper in his studies of maize seed development (Brink and Cooper, 1947; Cooper, 1951). Samples represent two time-points (noted in DAP at the bottom left of each panel) for kernels resulting from the following crosses: **A**) diploid-selfed, **B**) tetraploid-selfed **C**) paternal excess and **D**) maternal excess. nu = nucellus, end = endosperm, emb = embryo.

Diploid Line	Tetraploid Line	Reference
Unknown	Unknown	(Cooper, 1951)
W23	N107B (W23 autotetraploid)	(Charlton et al., 1995)
CML135, CML346, CML395, W23	4x-B (synthetic population derived from N107B, N107C, N104B)	(Leblanc et al., 2002)

Table 1 Maize lines in which interploidy crosses have been documented

Figure 1-9 Summary of Cooper's observations in maize interploidy crosses

Observations of developmental consequences in embryo and endosperm from balanced diploid, and interploidy crosses collated from Cooper et al, 1951.

Cross \ DAP	1	2	4	6	8	10	12	16	20	24
2n(X) Endosperm	Fertilization. 2-nucleate endosperm	16-32 nuclei		Basal epidermal cells near suspensor become absorbing tissue	Growth correlated with transformation of entire basal layer into absorbing tissue		All nucellus absorbed by endosperm First storage starch		Transformation of centrally located cells into conducting tissue	
		Cellularization	Active mitoses	Active mitoses only in peripheral layers, layer covering embryo face, and conducting tissue at base.			Active mitosis continues until 20 DAP		Further expansion is by cell growth	
2n(X) Embryo		2-celled pro-embryo	Average number of cells ~102							
2n x 4n Endosperm		16-32 nuclei					Nucellus has been reduced to a few crushed layers Basal epidermis fails to form absorbing tissue, and disintegrates	Central portion of endosperm that normally forms conducting tissue breaks down. This region is instead invaded by meristematic tissue in central region Peripheral layers of abgerminal face become irregularly meristematic, forming outgrowths		
		Cellularization – complete by 4 DAP	Highly active mitoses adjacent to embryo (up to 5 fold higher than 2n(x))	Active mitosis between germinal face and embryo – results in deeply embedded embryo.	Sub-epidermal cambium forms on adgerminal face.		Basal layer formed between embryo and face of kernel becomes absorbing tissue		Storage of starch commences	
2n x 4n Embryo		2-celled pro-embryo	Average number of cells ~59							
				Rate of growth slows down, or misshapen embryos form						
				Evidence of embryo disintegration						
4n x 2n Endosperm		Early cellularization	Initiation of sup-epidermal cambium-like zone at lateral and apical periphery	Very slow mitosis in cambial zone, ceases after 12 DAP Little or no mitotic activity in the central portion of endosperm Basal epidermis remains in close association with placenta. Cells near the suspensor become elongate, but rest fail to differentiate.				Well-defined aleurone is present		
				First evidence of starch accumulation	Indiscriminate storage of starch except for in cambial zone			Entire endosperm packed with starch, except epidermis		
4n x 2n Embryo			Average number of cells ~76							
				Growth of embryo at similar rate to that of 4n(X) embryo until 16 DAP						

The ultimate consequence of these crosses was seed failure, and in only a very small fraction of the kernels was it possible to recover a viable embryo. In both within-ploidy and interploidy crosses, fertilization was shown to have occurred within the first 24 hours. In the case of paternal excess crosses, the rate of embryo growth was only similar to the selfed diploid for the first two days, when both comprise around 16-32 cells, but by four DAP had decreased considerably. Subsequently embryo growth decreased until it stopped completely, usually before the formation of a radicle axis. In the maternal excess crosses the embryo grew more quickly than in the tetraploid self, and by four DAP the embryos comprised an average of 76 cells, compared with 44 in the selfed-tetraploid. Afterwards, the growth of the embryos from the different crosses are not very different until around 16 DAP, after which the cells of the embryo from maternal excess crosses appear to become starved of nutrients and cells start to vacuolate by 20 DAP.

However, the effect of these interploidy crosses on endosperms was even more extreme (Cooper, 1951). Paternal excess endosperms cellularized prematurely before the normal time of 4 DAP. Mitosis in these endosperms was also more widespread, occurring throughout the endosperm rather than being confined to the peripheral regions as in normal kernels. This was most marked in the region where the embryo faces the seed coat, which comprises just one cell layer in a normal endosperm, but in the paternal excess endosperm was composed of up to 20 layers. Although mitosis was much more widespread, it was poorly regulated, and produced ectopic outgrowths on the abgerminal face. Differentiation is reported to be disturbed, as a normal aleurone was not formed, and an "absorbing tissue" (the BETL) did not form at the usual time, although a similar tissue was found at a much later stage located between the embryo and kernel face.

In the case of maternal excess interploidy crosses, endosperm cellularization was shown to occur prematurely, and mitosis was slow in the cells of these endosperms. Growth was generally slowed, and the endosperm never fully occupied the entire region of the nucellus. Differentiation did occur, as a well-defined aleurone layer was observed by 16 DAP. Starch deposition occurred prematurely, and only a reduced absorbing tissue was observed (Cooper, 1951).

In more recent studies a clearer insight into the nature of endosperm failure following paternal excess crosses was provided by transmission electron microscopy (Charlton et al., 1995). This work suggested that the cells located in the region normally occupied by the BETL had either completely failed to develop, or had undergone an aberrant process of differentiation. Cells in this region normally display distinctive transfer cell morphology with invaginated cell walls, and possess many mitochondria. Instead, the cells in these paternal excess endosperms showed large electron opaque aggregates of unknown composition and very few mitochondria.

Intriguingly, another study into the expression pattern of a transfer layer-specific marker, BETL2, showed that the gene was expressed in the general location of the endosperm, although not in the expected smooth layer seen in balanced endosperms, but in irregular patches along the base (Hueros et al., 1999b). However, the authors reported that cells expressing this transcript did display the characteristic cell wall invaginations of BETL cells at a histological level.

It has also been shown that interploidy endosperms demonstrate a difference in the normal progression of mitosis in maize (Leblanc et al., 2002). The series of mitotic

events occurring in normal endosperms progress through several distinct transitions. Initially, nuclei divide without cytokinesis to form a large multi-nucleate coenocyte (Olsen, 2001). This is followed by a phase of cellularization, then a possible mitotic hiatus before a resumption of mitosis coupled with cell division and cell expansion. During this process, cells within the endosperm embark upon yet another transition: a phase of endoreduplication, in which the genomes are copied without daughter strand separation to form cells with a very high C value of up to 192C (Kowles and Phillips, 1988). In the work of Leblanc *et al* two main phase transitions were identified in the nuclei of endosperm using flow cytomtery. The first was at around 10 DAP, which represented the initiation of endoreduplication in 2x × 2x crosses (in 4x × 4x crosses this occurred up to two days later). During endoreduplication, a gradual increase in nuclear DNA content occurred until around 20-22 DAP, depending on the line involved, followed by a decline in DNA content. In contrast, in paternal excess interploidy crosses there was no distinct transition to a phase of rapidly increasing DNA content within the time points observed (8-22 DAP), and first signs of endoreduplication did not occur until 20 DAP. Conversely, in maternal excess endosperms the mean DNA content had started to increase by 8 DAP, peaked prematurely at 14 DAP, before decreasing considerably by 18 DAP.

Studies on the chromosome numbers of diploid, triploid, and tetraploid maize excluded the possibility that seed failure in interploidy crosses was due to incomplete transmission of genome sets through the microspores of tetraploid plants (Punyasingh, 1947). This confirmed that in maize interploidy effects genuinely result from the imbalance between complete maternal and paternal genomes.

1.4.2 Interploidy Crosses in arabidopsis

The other species in which the effect of interploidy crosses has most recently been reported is the model plant arabidopsis (Scott et al., 1998). Superficially there are

significant differences between such crosses and those in maize. For example, the seeds resulting from interploidy crosses in arabidopsis are mostly viable. However, there are also some instructive parallels. The two directions of interploidy cross in arabidopsis show broadly complementary phenotypes, with the final seed mass from maternal excess crosses being less than those of balanced crosses, and those from paternal excess crosses being greater. Whereas embryo and endosperm differentiation is retarded in maternal excess crosses, and results in fewer cells (a characteristic of maize maternal excess endosperm), paternal excess crosses display an accelerated rate of mitosis and have more nuclei, also something that happens in maize.

Additional crosses were performed between diploids and hexaploid lines of arabidopsis (Scott et al., 1998). This greater deviation from the normal endosperm dosage and balance does result in seed abortion, but it is preceded by what appears to be an exaggerated version of the series of events observed in the respective interploidy crosses between diploids and tetraploids. The authors proposed that these results might reflect the action of genomically imprinted genes expressed in the endosperm, as shown in maize (Lin, 1982). The difference in the 'severity' of interploidy crosses between maize and arabidopsis was suggested to reflect a partial breakdown of the imprinting function in arabidopsis, which is a natural inbreeder. For this reason it may not be subject to the same selective pressures on parental contributions to progeny which have been postulated to be responsible for the evolution of imprinting (Haig and Westoby, 1989).

Further work on arabidopsis has shed light on the mechanism by which parent-specific effects on the endosperm may be manifested in the genome (Adams et al., 2000). These experiments utilized diploid lines carrying an anti-sense construct to the *MET1* gene, the major methylase in arabidopsis, which show a down-regulation of the gene and

reduced levels of global methylation. These MET1-a/s lines were crossed with diploid wild type lines, and resulted in seed which partially phenocopied the effects of interploidy crosses. When hypomethylated 2x seed parents were crosses with a wild-type diploid pollen parent, a phenotype similar to paternal excess crosses was observed. In the reciprocal cross between a wild-type seed parent and a hypomethylated pollen parent, the resultant seed were similar to female excess crosses. The authors inferred from these results a role for methylation in the genome which confers differential effects on the two parents, possibly at loci that are genomically imprinted. They speculate that the reduced methylation in a MET1-a/s line 'paternalizes' the female genome, by removing female-specific methylation marks, and vice versa.

1.4.3 Interploidy crosses in other species

The outcome of interploidy crosses has also been reported in a much wider range of species including barley (Balkanjieva, 1983), primulas (Woodell and Valentine, 1961), and brassica (Nishiyama and Inomata, 1966), and is reviewed in (Haig and Westoby, 1991). The effects of interploidy crosses varies between species; however, it is often seen that interploidy crosses result in either inviable seeds, which frequently lack an embryo, or complete seed failure. Commonly, reciprocal crosses have different and alternate phenotypes but both often fail. However, it has been reported in one species, currant tomatoes (*Lycopersicon pimpinellifolium*) (Cooper and Brink, 1945) that the aberrant developmental processes preceding seed failure are essentially the same in both directions of interploidy cross (Cooper and Brink, 1945). Most frequently, 4x × 2x crosses (maternal excess) yield more viable seed than the reciprocal.

The work of Nishiyama in brassica is interesting as the observations parallel those in maize (Nishiyama and Inomata, 1966). Crosses in both species resulted in seed failure, showing very low seed viability, albeit slightly higher in the maternal excess crosses.

The paternal excess crosses displayed more cells than the control cross at early stages, and the maternal excess displayed fewer cells in both species. Early degeneration of the endosperm is seen in both species, whereas the maternal excess endosperms display a premature halt to increase in cell number. An early increase in starch uptake was observed in 4x × 2x crosses, as in maize, and development of the embryo never progressed past globular or heart stage embryos.

Interestingly, the 2:1 maternal:paternal endosperm genome balance is not a prerequisite necessary for the polarized outcome of interploidy crosses (Birchler, 1993). In oenothera, the endosperm of which has a 1m:1p genome ratio, interploidy crosses also display alternate phenotypes, the 'maternal excess' crosses (2m:1p endosperm) show accelerated differentiation, whereas the 'paternal excess' endosperm (1m:2p) show decelerated differentiation (von Wangenheim, 1962).

An intriguing insight into how the hybridization barrier of interploidy crosses might be overcome is provided by the observations that interploidy crosses between different genotypes in citrus display a range of recovery of germinable seed, suggesting that some element of genetic control underlies the process (Esen and Soost, 1973).

The apparent similarity of consequences following interploidy crosses across diverse taxa suggests that the effects are due to common control mechanisms amongst these species. This leads to the exciting prospect that discoveries in one species may elucidate analogous mechanisms in others.

1.5 Aims of this work

The aim of this project is to investigate the effects of interploidy crosses on maize seed development by extension of previous work (Charlton et al., 1995; Cooper, 1951; Leblanc et al., 2002) using a range of modern techniques. Interploidy crosses provide a relatively straightforward way in which to uncouple maternal and paternal contribution to the seed, and should provide an insight into the effects exerted by these genomes over seed development.

Crosses involving reporter line plants for two cell domains of the endosperm will be used to supply a developmental profile for these domains. This will first be carried out in balanced diploid crosses, from the time of first appearance of the reporter marker until the start of seed maturation. Histological work will establish whether expression of these markers correlates with the adoption of the cell identity which they report. This work will provide a benchmark for comparison with data emerging from experiments in which these reporter lines are used in interploidy crosses, which will reveal how cell domain specification is altered by an when parental genome ratio in the endosperm is disturbed.

Microarrays will be used to investigate how the patterns of gene expression following interploidy crosses compare with balanced controls. Two types of array will be used: one to assess a large fraction of the genome, with the aim of assessing the global genomic response to imbalanced parental ratios; and another on a limited subset of genes of interest. In addition to information on interploidy crosses, these arrays should provide gene expression profiles from early endosperm at a time-point immediately following cellularization. RT-PCRs will be used to verify some of the results from these arrays, and also track changes in gene expression through a period of development.

In short a combination of two powerful complementary techniques of cytology and profiling gene activity should probe in great detail the nature of interploidy crosses, and thereby provide one of the clearest insights into the different roles conferred on the seed by the two parents.

2 Methods and Materials

2.1 Reagents

2.1.1 Chemicals

Unless otherwise specified all chemicals were obtained from Sigma-Aldrich Co. Ltd (Poole, Dorest, UK), enzymes were purchased from Invitrogen (Paisley, UK), New England Biolabs (Hitchin, Hertfordshire, UK), or Promega (Southampton, Hampshire, UK), oligonucleotides were supplied lyophilised by MWG-Biotech (Milton Keynes, UK), or Sigmas-Genosys (Cambridgeshire, UK). Sequences for all Oligonucleotides produced are detailed on the accompanying CD. Water was purified to a purity of 18.2 Ω cm, and is referred to as ddH_2O.

2.2 General methods

2.2.1 Ethanol Precipitation of DNA

DNA was precipitated from PCR components by adding 2 volumes of ice cold 100% ethanol, followed by vortexing, and freezing at −20°C overnight (see Section 2.6.12 for rationale). DNA was recovered by centrifugation (13,500 x g for 10 minutes at 4°C), the pellet washed in 70% ethanol at 4°C, and centrifuged (13,500 x g, 5 minutes at 4°C). The pellet was air dried and resuspended in an appropriate amount of distilled or DEPC-treated water. DNA quality and quantity was assessed by measuring the OD260 and OD280 using UV spectrophotometry (Section 2.2.2 below), and by running on a TBE gel.

2.2.2 UV spectrophotometry

DNA and RNA was quantified using a UV spectrophotometer. A volume of the sample (from 1-10 µl) was diluted in a final volume of 140 µl of either ddH_2O (for DNA) or DEPC-treated distilled water (for RNA). The machine was referenced using 140 µl of the solvent in which the nucleic acid was suspended (DNA: ddH_2O, RNA: DEPC-treated

ddH_2O (see section 2.2.3)). The absorbance measures were taken at OD260 and OD280. An OD260 of 1.0 was considered to represent a concentration of 50 µg/ml for DNA and 40 µg/ml for RNA (pathlength = 1cm). The purity of the sample was judged by the ratio of the OD260:OD280, which should have been between 1.6-1.8 for DNA and between 1.8-2.0 for RNA.

Other samples were measured on a Nanodrop spectrophotometer (Nanodrop, Wilmington, DE), in which samples were diluted to an appropriate concentration, and 1 µl pipetted onto the sensor after having been calibrated to the buffer. A full spectrum was obtained, and again quantity and quality determined from the OD260/280 ratios, in context of the general shape of the trace. Samples were also analyzed in a GeneQuant DNA calculator (Pharmacia Biotech), for which samples were diluted appropriately, and 7 µl of the solution used for analysis.

2.2.3 DEPC treatment of solutions

Solutions intended for use with RNA were treated with the addition of 0.1% DEPC overnight at room temperature to inactivate RNAases, before autoclaving to breakdown the DEPC molecules.

2.3 Plant Material

2.3.1 Maize lines used

The following non-transgenic plant lines were used in these studies, and are described more fully in sub-chapter 3.1.1:

Plant abbreviations used in this text	Genetic Background	Ploidy	Source
H99-2x	H99	Diploid	Maize Stock Centre
H99-4x	H99	Tetraploid	Dr James Birchler, University of Missouri
HybT	½ OH43-4x, ½ MO17-4x	Tetraploid	Dr J Gutierrez-Marcos, University of Oxford

Table 4 Summary of non-transgenic maize lines used in these studies

The following transgenic lines were used in these studies:

Plant Names used in this text	Genetic Background	Insert Carried	Ploidy	Source
pBETL::GUS line	Derived from Hi2, and introgressed into W23 by four generations	pBETL::GUS	Diploid	Dr Richard Thompson (INRA-Dijon, France)
pVp1::GUS line	Derived from Hi2, and introgressed into W23 by four generations	pVP1::GUS	Diploid	Dr P. Perez (Biogemma, France)

Table 5 Summary of transgenic maize lines used in these studies

2.3.2 Plant Growth conditions

Plants were grown between the years of 2001 and 2004, in greenhouses in Oxford (Oxfordshire) and Jealott's Hill (Berkshire). In addition to natural light, the plants were exposed to an artificial light regime from sodium bulbs in a 16 hour daylight cycle. Plants were watered regularly, and grown in a 1 part compost : 1 part peat mix in pots appropriate to the plant size, but ultimately in 7.5 litre pots.

2.3.3 Endosperm isolation

4 DAP endosperm and embryos were isolated for RNA analysis, and immediately placed in 2ml tubes, with screw-cap tops, and then transferred to liquid nitrogen before long term storage at −80°C. Endosperms were isolated intact by making an incision into the ad- or ab-axial face of the endosperm using a fine scalpel blade (Size 11, fitment number 3, Swann-Morton, #0303). The nucellar tissue was prised open using the scalpel and tweezers, to reveal the endosperm, which occupied only around 1/4 of the nucellus at this stage. The endosperm was released by pushing up from the base of the kernel. As the embryo was very small, and seldom visible it was decided not to attempt to separate it from the endosperm.

2.3.4 Verification of genome number by flow cytometry

In order to confirm the ploidy of the different lines under investigation around 4 cm^2 of leaf material was obtained from each plant and finely chopped in 2 ml of buffer LB01 (Dolezcaronel et al., 1989) and 50 µg/ml propidium iodide and 50 µg/ml RNAse using a double edged razor blade in a petri dish. Tomato leaves were used as standards. The sample was then passed through a 42 µm sieve, and then analysed in a FACSScan machine (BD Biosciences, San Jose, CA), peaks measured by the CellQuest program (BD Biosciences). Four replicate measures were taken for each sample, including controls.

2.4 Microscopy and Histology

For the histological work detailed in chapter 4, numerous samples were observed, and statements detailing observations or images taken, only where typical of the vast majority of kernels. For most directions of cross and time points kernels from two independent crosses (separate ears pollinated with distinct pollen samples) were observed, and from these between 30 and 60 kernels were stained and observed.

2.4.1 Staining of samples to observe GUS activity

Seeds resulting from crosses with transgenic plants were analysed for GUS reporter activity by initially pre-cutting the samples by hand using a fine scalpel. The samples were then rinsed in 50mM sodium phosphate buffer (pH 7.0) for a minimum of ten minutes, which was replaced once to remove all the starch. The buffer was replaced with staining solution (0.05 M sodium phosphate buffer (pH 7.0), 10 mM EDTA (pH 7.0), 0.5 mM potassium ferricyanide, 0.5mM potassium ferrocyanide, 1 mM X-GlucA substrate, 0.1% triton), vacuum infiltrated for 10 minutes, and placed at 37°C. The staining reaction was left for 1 - 12 hours, and stopped by replacing the staining solution with 1 X PIPES (50mM PIPES, 5 mM EGTA, 1mM $MgSO_4$) solution for 1 hour.

2.4.2 Fixation of material

Following staining samples were fixed overnight in an cold solution of either 1 x PIPES, 4% formaldehyde, 1% gluteraldehyde solution; FAA (2% formaldehyde, 5% acetic acid, 45% ethanol), or 3:1 ethanol: acetic acid, after an initial period of 10 minutes of vacuum infiltration. Samples were then dehydrated in an increasing ethanol series, from 30% to 100% ethanol (starting at 50% for samples from FAA fix, and 70% for those from 3:1 fix). Samples were left in each stage of the series for 30 minutes.

2.4.3 Embedding and sectioning of samples

Samples were fully dehydrated with 3 changes of 100% ethanol at 4°C for 4 hours, then another change of 100% ethanol at room temperature for 2 hours. Ethanol was gradually replaced with Histoclear (National Diagnostics, Atlanta, GA) in a series of 3:1, then 1:1 Histoclear:ethanol, and left for 1 hour each time, followed by three changes with 100% Histoclear, each for an hour. This was then replaced with 1:1 Histoclear:wax (Paraplast plus, Sigma, #P3683) and left overnight at room temperature. The following day the samples were then transferred to 45°C for 1 hour, then to a wax oven set at 60°C. The Histoclear:wax was mixture was gradually replaced with molten wax over the course of the day. At the end of that day, a complete wax change was performed with molten wax. Two further wax changes were carried on each of the succeeding two days. Finally, individual kernels were embedded in individual embedding moulds, left to cool, and stored at 4°C until processed.

Wax blocks were trimmed close to the samples, ensuring edges were parallel, and sectioned at a thickness of 5-12 μm on a rotary microtome. Samples were floated on warm ddH2O on a cover slip on a hot plate until samples were fully stretched. Water was removed by blotting, and samples left in a 37°C incubator overnight to allow

covalent attachment of the sample to Superfrost slides (#406016902, BDH-VWR, Poole).

2.4.4 Staining of material in Periodic Schiff 's reagent for carbohydrate

Paraffin wax was removed from slides through a Histoclear series, and then rehydrated via ethanol. Slides were incubated in 0.5% periodic acid for 10 minutes, washed in running tap water for 10 minutes, then Schiff's reagent (3952016, Sigma) was applied to slides and covered with a coverslip. After 10 minutes slides were washed in running tap water for a further 10 minutes, then placed in a 2% sodium bisulphite solution for 2 minutes. Samples were dehydrated, and taken to 100% ethanol, then Histoclear before embedding in DPX mounting medium.

2.4.5 Staining of material in Toluidine Blue

Wax-embedded sections were de-waxed, then rehydrated before being dipped in a 0.05% solution of Toluidine Blue O in 50 mM citrate buffer. The samples were washed in three changes of water for two minutes each, before dehydration through 95% ethanol and two brief changes of 100% ethanol. Samples were then soaked in two changes of Histoclear for five minutes each before embedding in DPX mounting medium.

2.4.6 Microscopy

Microscopy was performed using a Zeiss AxioPhot microscope, and images taken with a Nikon Coolpix 990 (3.34 Megapixel) digital camera, using bright field, differential interference (DIC3-5), or dark field optics. Fluorescence microscopy used a 50 W mercury lamp and the following filter set: 365 nm excitation, 395 nm dichroic and 420 nm long-pass emission.

2.5 Microarray Target Production

RNA used in both kinds of microarray employed (purchased oligo arrays or custom-made cDNA arrays) represented a pooling of individual kernels (usually between 25 and 50) from the same cross (sharing the same maternal and paternal parent), but there was no pooling between ears.

2.5.1 Isolation of total RNA – Guanidium Isothiocyante Method

Isolated samples which had been stored in a freezer at –80°C were cooled in liquid nitrogen, and then crushed in the tube using a chilled micropestle. 700 ml of extraction buffer (100mM Tris-HCl pH 8.6, 2% sarkosyl, 4M guanidine thiocyanate, 25mM EDTA, 25 mM EGTA, 100mM β-mercaptoethanol) and 300ml phenol was then added, and any endosperm that were not already homogenized were clearly visible, so were further ground until no longer visible. Samples were then shaken vigorously for 10 seconds, before the addition of 350 µl of 24:1 chloroform: isoamyl alcohol, and thoroughly shaken for a further 10 seconds before centrifugation at 14000 x g at 4°C for 10 minutes. The upper aqueous phase was then transferred to a new tube, and an equal volume of 25:24:1 phenol:chloroform:isoamyl alcohol was added, before shaking and centrifugation at 14,000 g at 4°C for 5 minutes. The upper aqueous phase was then transferred to a new tube, and nucleic acid was precipitated by adding 3.5 µl glacial acetic acid, and 0.6 volumes of isopropanol. After shaking, the sample was placed at –80°C for 30 minutes, then centrifuged at 14,000 x g for 10 minutes. The pellet was rinsed in 95% ethanol, and centrifuged at 14,000 x g for 10 minutes at 4°C. After removing the supernatant the pellet was air dried, and resuspended in 400 µl of a 0.1% DEPC- treated solution. The sample was placed at 37°C for 10 minutes and on ice for 15 minutes before centrifugation at 14,000 x g at 4°C for 10 minutes. The supernatant was transferred to a new tube, to which 100 µl 8M lithium chloride was added, and left overnight at 4°C. The sample was then centrifuged at 14,000 x g for 10 minutes at 4°C, and the supernatant reserved for DNA precipitation (by keeping overnight at –20°C in 2 volumes of 100% ethanol, and then centrifugation at 14,000 x

g, and washing of the pellet in 70% ethanol), the pellet was resuspended in 400 µl DEPC-treated distilled water, to which 40 µl 3M sodium acetate and 1ml 95% ethanol was added. The sample was stored at –80°C for 10 minutes, before centrifugation at 14,000 x g for 10 minutes at 4°C. The sample was then air dried, before being resuspended in DEPC treated distilled water which was kept at 60°C for 10 minutes and on ice for 30 minutes before a final centrifugation at 14,000 x g at 4°C, after which the solution was transferred to a new tube, before quantification by UV spectrophotometry (section 2.2.2), visualization on a denaturing formaldehyde gel (section 2.5.5) and storage at –80°C.

2.5.2 Isolation of total RNA – Trizol method

Frozen samples of 10-30 mg were pre-chilled in liquid nitrogen. 1ml of Trizol reagent (#18068-015, Invitrogen) was added and the samples crushed with a micropestle until no longer visible. Samples were then shaken for 15 minutes, then left at room temperature for 5 minutes. 200 µl of chloroform were then added, before shaking for a further 20 seconds. After leaving at room temperature for 5 minutes the samples were centrifuged for 20 minutes at 12,000 rpm at 4°C. The upper aqueous phase was removed and transferred to a new RNAase-free tube, and either processed as follows or as in section 2.5.3. 250 µl isopropanol was added to the supernatant and incubated for 10 minutes at room temperature. The sample was centrifuged at 12,000 x g for 10 minutes, the supernatant removed and the resultant pellet was washed in 500 µl 80% ethanol, vortexed, and centrifuged at 7,500 x g for 5 minutes at 4°C. The pellet was dried using a vacuum centrifuge, before being dissolved in typically 20 µl of RNAase-free water. To ensure full resuspension, the pellet was dissolved by incubating at 55°C for 5 minutes.

2.5.3 Isolation of RNA – Trizol and RNeasy clean-up method

Supernatant from the first centrifugation produced using method 2.5.2 was transferred to a fresh RNAase-free tube, to which one volume of 80% ethanol was added, and mixed immediately. This solution was then flushed through an RNeasy column in 700 µl volumes by spinning at 10,000 x g for 15 seconds at room temperature. After discarding flow through, the sample was washed with 700 µl RW1 buffer provided with the columns, centrifuging for 15 seconds at 8,000 x g. Two additional washes with 500 µl buffer RPE were performed for 15 seconds (or two minutes the final time) at 10,000 x g. After transferring the column to a clean RNA free tube, typically 25 µl RNAase-free water was added to the column matrix, and left for 5 minutes at room temperature. The tube was spun for 1 minute at 10,000 x g, reapplied to the column, and spun once more for a minute. 1 µl of SUPERasin RNase inhibitor (Ambion, #2696) was added, and the sample was stored at –80°C.

2.5.4 RNA isolation – Hot Phenol Method

This method was effective for starchy samples, such as the RT-PCR time series extractions. Up to 100mg of tissue was ground in a pestle with liquid nitrogen, to which was added 250 µl phenol and 250 µl extraction buffer (0.1M lithium chloride, 0.1M Tris-HCl, 10 mM EDTA,1% SDS; pH 8.0 and filter sterilized), which had been pre-heated to 80°C. Samples were transferred to an 1.5 ml tube, and vortexed for 30 seconds. 250 µl chloroform:IAA was added and vortexed a further 30 seconds. Samples were centrifuged for 5 minutes at 13,000 x g at room temperature. The aqueous phase was transferred to a new tube, to which an equal volume of 4M LiCL was added. The tube was vortexed, and left at 4°C overnight. Tubes were then centrifuged for 10 minutes at 13,500 x g, the supernatant was removed, and resuspended in 250 µlDPEC-ddH_2O. 0.1 volumes of 3M sodium acetate (pH5.2) and 2 volumes cold 100% ethanol were added, before vortexing and incubating for three hours at -20°C. Samples were centrifuged for 10 minutes at 13,550 x g, the supernatant removed, and the pellet

washed in 70% ethanol. The pellet was then dried and resuspended in 50 µl DEPC-ddH_2O.

2.5.5 Denaturing Gel Electrophoresis

RNA integrity was verified by running on a denaturing formaldehyde gel. 7 µl of the sample (typically 0.5 – 1 µg) was mixed with a 23 µl of RNA Sample buffer (300 µl 10 x gel buffer, 1500 µl formamide, 14 µl ethidium bromide, 486 µl bromide), heated to denature, before cooling on ice and running through a 1.5% agarose denaturing gel (17ml formaldehyde, 10 ml 10 X gel buffer) in 1 x RNA running buffer. 10 X gel buffer: 200 mM MOPS, pH 8.0, 50 mM NaOAc, 10 mM EDTA, pH 7), 10 X running buffer (200 mM MOPS, pH 7.0, 50 mM NaOAc, 10 mM EDTA).

2.5.6 Agilent Bioanalyzer

RNA integrity for the dendrimer microarray labelling protocol was deemed essential, so RNA samples for this application were analysed using an Agilent 2100 bioanalyzer. Typically 1 µl of my RNA sample was diluted in 10 µl total volume of RNAase free water and 1 µl of this was then processed according to the manufacturer's instructions.

2.5.7 RNA amplification

RNA was amplified from 1 µg total RNA using the MessageAmp II kit (Ambion, #1751) following the manufacturer's instructions, using a 14 hour in vitro transcription incubation time. aRNA was quantified using a GeneQuant spectrophotometer.

2.5.8 Fluorescent RNA labelling

Either 30-50 µg of total RNA, or 3 µg of aRNA was heated to 70°C for 10 minutes in the presence of 1 µg oligo dT primers, or 3 µg of random hexamers, respectively. After a pulse spin and chilling on ice the RNA was reverse transcribed in the presence of a

fluorescent label in a final volume of 40 µl which containing 1X 1st strand buffer, 4 µl 10 mmoles DTT, 1 unit RNase inhibitor, 3 nmoles Cy3-dCTP (PA53021, Amersham) or Cy5-dCTP (PA55021, Amersham), and 2.5 mmoles of dATP, GTP, and dTTP, and 1 mmole dCTP. 400 units Superscript II (18064-014, Invitrogen) were then added before incubation at 42°C for 2 hours. The reverse transcription was terminated by the addition of 5 µl EDTA (0.5M, pH 8.0), and 10 µl sodium hydroxide (1N), followed by incubation at 65°C for 30 minutes, and neutralization with 25 µl 1M Tris pH 7.5. The sample was purified and the label concentrated using a Microcon YM-30 filter concentrator (Millipore, #4209). The labelled cDNA samples were pooled, and 380 µl of 1X TE was added, before passing through the filter at 14,000 x g for 8 minutes. The sample was twice cleaned by the addition of 450 µl 1 X TE and centrifugation at 14,000 x g for 8 minutes. On the final centrifugation step, 4 µl of 'liquid block' (Amersham, RPN3601) was also added, and the centrifugation time increased by a few minutes until the column contained around 30 µl. The Microcon was then inverted in a new 1.5 ml tube, and centrifuged at 14,000 x g for a few minutes until all the sample was recovered. The sample was then dried in a heated vacuum concentrator, and resuspended in a volume of 34 µl of hybridisation buffer (0.1% SDS, 25% formamide, and 5 X SSC). The probe was denatured by heating to 95°C for 5 minutes before applying to the printed area of a prehybridised slide, covering and incubation at 42°C for 16 hours.

2.6 Custom Microarray Design

2.6.1 Selection of target sequences for array

Sequences were mostly based on those catalogued in the ChromDB database[2] on the 20 November 2001. Sequences were initially downloaded as a text file. Data was transformed to extract data for *Zea mays* sequences only, and produce a double

[2] www.chromdb.org

stranded sequence that could be used to assist primer design, by implementation of a Perl program, Primer_check, detailed in 2.9.3, and on the enclosed CD.

These sequences represented the consensus of the contigs for each chromatin protein present in the database at that time. These sequences were not in the sense direction, so had to be translated in each of the six possible reading frames to identify the correct location, using the DNA-Protein Translate tool at the ExPASy proteomics server at the Swiss Institute of Bioinformatics[3]. The frame in which the longest contiguous region lacking stop codons was assumed to be the correct frame of the protein, and this sequence was then verified by using the BLAST algorithm against a protein database to ensure high scoring sequences were obtained (Altschul et al., 1997). The proximal-most stop codon to the coding region was then identified. An EST from each gene contig was selected (where possible a Unigene representative, which could be supplied by a collaborator, Dr V Chandler, University of Arizona). Where possible ESTs spanning the 3' UTR were chosen as this region is usually the most variable in a gene, and should minimize sequence redundancy between duplicate genes (Fernandes et al., 2002). When no ESTs meeting these criteria were available, coding sequences were considered. For all genes, forwards and reverse primers were designed between 150 and 300 bp apart. Intervening sequences were PCR amplified either from plasmid where ESTs were available, or from genomic DNA. In the latter case sequence identity was verified by cloning into the pGEMt-easy vector, and sequencing. Full sequence data is provided in Appendix 1.

2.6.2 Design of Primers

Primers were designed using a number of criteria to maximise efficiency of amplification. Primers were selected such as to amplify fragments that were around

150-300 base pairs in length. The primers themselves were around 17-25 nucleotides long, and had a melting temperature of 60-70°C. Melting temperature was estimated, and amplified product length determined using the PERL script Primer_check, detailed in 2.9.3, and on the accompanying CD, and tabulated in Appendix 1.

2.6.3 Polymerase Chain Reaction (PCR)

Sequences were amplified by polymerase chain reaction (PCR), in volumes of 50 or 100 µl containing approximately 1 µg of template DNA, 0.2 µM of forward and reverse primer, and PCR buffer (20 mM Tris-HCl(pH 8.3), 0.75 mM magnesium chloride, 25 mM KCl, 0.05% Tween 20, 100 µg/ml BSA, 100µM mixed dNTPs, 2% DMSO, 2 units Taq DNA polymerase. Reaction conditions varied for different primer pairs, but generally the reaction conditions were as follows: pre-denaturing step at 94°C for 2 minutes, followed by 30-35 cycles of three temperatures: denaturing at 94°C for 30 seconds, annealing at 53-60°C for 20 seconds, an extension at 72°C for 20-40 seconds; and a final extension at 72°C was carried out for 5 minutes. PCR reactions were carried out in a PTC100-thermal cycler (MJ research).

2.6.4 TBE gel

5µl of the reaction product was loaded in 1X Tris-Borate-EDTA (TBE) running buffer on a 1.5% TBE gel containing 30µM ethidium bromide, run at 80-140 volts for 30 minutes, and visualized on a UV transilluminator.

2.6.5 Isolation of DNA fragments from agarose gels

Following amplification of products that were of expected length, and unique bands, 30µl of the PCR product was run out on a gel, and the band excised with a razor blade,

[3] http://us.expasy.org/tools/dna.html

following visualization on a UV transilluminator. The sample was then purified using a QIAquick Gel Extraction Kit (Qiagen #28704), according to the manufacturer's instructions.

2.6.6 Ligation of DNA into pGEM-t easy vector

DNA was ligated into pGEM-t easy vector system 1 (Promega #A1360) using a 3:1 molar ratio of insert DNA to vector plasmid according to the instructions of the supplier.

2.6.7 Transformation of competent *E. coli*

Competent *E. coli* cells (strain DH5α) was transformed with pGEMt easy vector, containing amplified sequences, to provide a stable source of insert plasmid allowing the generation of sufficient amounts of template plasmid. 50 µl of cells were thawed on ice, and then mixed with 5 µl of plasmid by gentle flicking. Tubes were then left in ice for 20 minutes before being exposed to a heat-shock by placing in a water bath at 42°C for 50 seconds, before being returned to ice for 2 minutes. 950 µl of liquid LB was added to the cells, before being placed at 37°C with shaking (around 150 rpm) 1 hour. Cells were then pelletted by centrifugation at 1000 x g for 8 minutes, and then resuspended in 100 µl of the culture medium. Up to 50 µl of these cells were then plated out 10 cm^2 petri dishes containing 25 ml solid LB containing 2mM IPTG, 0.004% X-Gal and 100 µg/ml ampicillin, to allow blue/white selection of cells containing an insert in the multiple cloning site.

2.6.8 Colony PCR

Up to 10 replicas of white colonies from overnight incubations of pGEM-t easy transformed cells were plated on new LB plates containing just the selective antibiotic.

New colonies were grown overnight at 37°C, before being used to test the insert was of the expected size, by colony PCR. Components for the reaction, and reaction cycles were as described in section 2.6.3 although, the template was generated by lysing cells stabbed from the colony using a pipette tip, which was placed in PCR reaction wells containing ice cold water, and were left to shake for 10 minutes. Amplification primers used were either M13 forward and reverse, or else T7 and SP6 pairs. Samples were run out on 1% TBE gels and visualized. Expected product lengths were now slightly longer due to the presence of the multiple cloning site within the area of the universal cloning site primers. Colonies containing products of expected length were selected for overnight growth in liquid culture.

2.6.9 Cell culture

Colonies were picked with sterile toothpicks, and then dropped in 5ml of LB, containing 100 μg/ml ampicillin, and placed in a 37°C incubator overnight with shaking.

2.6.10 Recovery of plasmid DNA from bacterial cultures

Cells suspended in LB from overnight cultures were pelleted by centrifugation (12,000 x g for 5 minutes). The pellet was then lysed and DNA purified using a Qiagen QIAprep miniprep kit (#27104), following the manufacturer's instructions.

2.6.11 Sequencing of insert

Plasmids isolated through minipreps were sequenced using the BigDye Terminator v3.1 kit (Applied Biosystems, #4337455), using 4 μl of purified plasmid, 4 μl BigDye reagent, and 3.2 pM of T7 primer. The reaction was performed in a thermocycler according to the manufacturer's instructions, and the labelled samples purified using a

Qiagen DyeEx 2.0 spin kit (#63206), following the supplied instructions. Samples were sequenced on a Applied Biosystems 3730xl DNA Analyzer.

2.6.12 Production of probe for microarray spotting

Following optimization of reaction conditions for all sequences from which it was possible to amplify a fragment of expected length, targets were amplified in 100 µl volumes. These were cleaned up initially using a 96-well PCR clean-up plates (Machery-Nagel, #743100.10) following the supplier's instructions. The results provided from microarrays spotted from these plates were disappointing, and led to several poor hybridisations. Plates were stained with propidium iodide, to identify spots which had low amounts of DNA bound to them. These spots were re-amplified, cleaned up using QIAquick PCR clean-up columns, and a new batch of slides were printed including these re-amplified spots. Similarly disappointing results were obtained from these slides. Following numerous attempts to optimise hybridisation conditions, it was decided to again re-amplify all spots that were deemed too low in DNA concentration from the latest printing using propidium iodide staining. The original spots, this time up to 96 spots were re-amplified, using 2 or 3 100 µl PCR reactions (Section 2.6.3). It was also identified that the method of DNA clean-up that gave the greatest DNA yields, and highest purity as identified by OD260:OD280 ratio of between 1.8 and 2.0, was ethanol precipitation in the absence of salt (see 2.2.1).

2.6.13 Printing of microarray slides

Amplified and purified probe sequences were arranged in 96-well polycarbonate plates (Costar, #6511) following resuspension in one of two printing buffers. In the initial print onto Corning UltraGAPS slides (#40016), probe was spotted in a solution of 50% DMSO and 3X SSC in a final volume of 30 µl. In a later print on Telechem ArrayIt Superamine slides (SMM-BC), the DMSO, SSC solution resulted in very irregular

shaped spots, or variable slides. A 3M betaine solution which produced homogenous round spots was therefore used, using a 30 µl volume. Slides were spotted using an Affymetrix pin-and-ring 417 arrayer (MWG Biotech).

Slides were stored following printing in a dark airtight box with desiccant, and DNA immobilized as slides were needed. DNA was immobilized on the slides as follows: rehydration over a beaker of 90°C distilled water, snap drying of the slide, spot side up, on a hot plate (~90°C); slides were then left at room temperature before repeating the process another two times. Slides were UV crosslinked using a UVP CL-1000 UV crosslinker at a power of 200 mJ/cm^2, then baked in an oven at 80°C. Slides were not washed before use, as it was found that pre-hybridization yielded a much more efficient method of removal of unbound DNA.

2.6.14 Propidium Iodide Staining of Microarray Slides

DNA bound to slides could be visualized by applying 30 µl of a propidium iodide solution (30 µg/ul), covering with a coverslip and leaving in the dark for 15 minutes. Slides were washed by dipping in water, and then dried by centrifugation in a 50ml plastic tube (Corning) (2,000g x g, 2 minutes). Slides were visualized by scanning in an Affymetrix 427 microarray reader, using a 532nm filter.

2.6.15 Microarray Prehybridisation

Spotted slides were prehybridized for 2 hours in a rotary shaker at 42°C in a solution of 0.1% SDS, 25% formamide, 5X SSC, and 10 mg/ml BSA. Slides were then washed in double distilled water by dipping for 5-10 seconds before being dried by centrifugation at 2000 g in a 50 ml tube.

2.6.16 Hybridisation

After applying a labelled probe in a hybridisation solution (section 2.7), a glass coverslip (15 x 50 mm) was carefully placed on top, avoiding the formation of air bubbles. The slide was then placed in a preheated oven at 42°C and left to hybridise for 14 hours.

2.6.17 Microarray Wash Protocol

After overnight hybridisation, slides were washed by dipping in a solution of 2X SSC, 0.1 % SDS until the coverslip was displaced, before being left in the same wash solution for an additional 5 minutes in incubator at 42°C with shaking. The slide was transferred to a solution of 0.1X SSC, 0.1% SDS at room temperature, and left shaking for 10 minutes. The slide was transferred to a new solution of 0.1X SSC, replaced three times after incubation for a minute without shaking. The slide was then dipped for around 5-10 seconds in double distilled water, before being dried by centrifugation at 2000 g in a 50 ml plastic tube.

2.7 Oligomer Array Protocols

Community-produced arrays containing up to 58,000 oligomers separated between two phyical glass slides were purchased from the Maize Oligonucleotide Project as detailed in Sub-Chapter -5.2.1.

2.7.1 Array Processing

Supplied slides had not been processed following printing. Slides were scanned initially, and the close layout of some spots on one slide (B39) suggested it would be preferable not to rehydrate spots, for risk of them merging. Hence, slides were baked at 80°C for 40 minutes on a heating block, before being cross-linked twice at 300 mJoules in a Stratalinker (Wang et al., 2003b).

Prehybridization was carried out at 65°C for 20 minutes in a plastic coplin jar containing 3.5X SSC, 0.1% SDS and 10% BSA. Slides were then washed by dipping in MilliQ water for 1 minutes, and in isopropanol for 1 minute. The slide was dried with an air brush, and then scanned at a resolution of 50 μm to record quality of the array.

2.7.2 Dendrimer labelling protocol

The 3DNA 900 labelling kit (W500180, Genisphere) was used to label RNA, in a modification of the manufacturer's protocol. RNA was first quantified and assessed for quality using a nanodrop spectrophotometer, and Agilent bioanalyzer respectively. 2 μg of total RNA was denatured at 80°C for 5 minutes in a thermocycler with 1 μl of either of the primers specific to each dendrimer capture sequence (Cy3 or Cy5). After cooling on ice for 2 minutes, 1st strand reverse transcriptase buffer was added to a final concentration of 1X, 20mM DTT, 0.5 μl superasin, 250 μmoles dNTP, and 100 units SuperscriptIII (Invitrogen, #18080). Samples were incubated at 42°C for 2 hours in a thermocycler. The reaction was stopped by the addition of 1 μl of a 1M NaOH /100 mM EDTA solution, followed by an incubation of 10 minutes at 65°C, and neutralized with 1.2 μl 2M Tris-HCl (pH 7.5). The two cDNA solutions for each slide were then mixed and made up to 85 μl with nuclease-free water, and then mixed with an equal volume of the supplied 2X hybridisation solution (vial 5, non-formamide based). The sample was preheated to 80°C for 10 minutes before cooling to the hybridisation temperature of 55°C. The slides were placed on a SlideBooster (SB400, Advalytix), on top of 17 μl of coupling solution, a LifterSlips (Erie scientific, #22x60I-2-4861) coverslip was placed on top of the slide, and the hybridization solution was applied to the array. The slide was left for 16 hours on a SlideBooster in the presence of 700 μl humidifying solution in each of the two wells, and at a temperature of 55°C (power 27; pulse 3/7).

The slide was washed initially in a solution of 2X SSC, 0.2% SDS pre-warmed to 55°C by dipping to allow the lifterslip to fall off, and then for a further 10 minutes on an orbital shaker at 55°C. A second wash was performed in 2X SSC at room temperature for a further 10 minutes, with shaking, and then for another 10 minutes in 0.2X SSC with shaking at room temperature. The slide was then dried using an airbrush.

A second hybridisation was then performed to allow the fluorescent capture reagent to bind to its complementary primer sequence. For each array 1 µl of both Cy3 and Cy5 capture reagents were made up to a final volume of 170 µl using the 2X hybridisation solution (vial 6) and nuclease-free water, and then preheated to 80°C for 10 minutes. The solution was applied to the array which had been set up in the slidebooster as for the first hybridisation. The slide was left for four hours in the SlideBooster using the same settings as before being washed, using the same conditions as after the first hybridisation, except that these steps were carried out in the dark.

After drying the slides were scanned at a resolution of 5 µm on a GenePix 4000B scanner (Axon instruments, Union City, CA), with a power adjusted to get a balanced signal from both channels, and a minimal number of saturated spots.

2.8 Semi Quantitative RT-PCRs

RNA was isolated using the hot phenol method (2.5.4), and up to 2 µg RNA was DNAase treated in a volume of 10 µl using 1.0 µl DNase I (Invitrogen, #18068-015), and 1.0 µl 10X buffer then incubated at room temperature for 15 minutes. 1 µl of 25 mM EDTA was added, before heating to 65°C for 10 minutes. cDNA was quantified, and ran out on a 1.5% TBE gel to verify integrity (2.6.4). 1.0 µg of DNase-treated RNA was used in a reverse transcription reaction in a final volume of 20 µl. Initially the

RNA, 500 ng oligo dT, and 1µl dNTP (10mM) was heated to 65°C for 5 minutes then cooled on ice. To this 4 µl 5X 1st strand buffer, 1.0 µlDTT, 1 µl RNAse inhibitor (Ambion, #2696) and 1 µl Supercsript III (Invitrogen, #18080) was added. The reaction cocktail was incubated at 50°C for 60 minutes in a thermocycler, the reaction was stopped by heating at 70 °C for 15 minutes. Samples were diluted up to 1/10 (adjusted for initial starting amount of RNA) 0.5µl of which was used in a 50 ul PCR reaction which also contained 20 nmoles of primer, and 45 µl of 1.1X PCR mix (ABGene, AB-0575/LD/A). A PCR cycle was developed to allow fractions of the reaction to be reserved at several intervals through the whole cycle. After the initial 20 cycles (94°C-20 s, 55°C-20s, 72°C-30s), a step with a long pause is repeated for a further 20 cycles (94°C-20 s, 55°C-20s, 72°C-30s, 20°C-5 minutes). This cycle facilitated the removal of 5 ul of sample at every fifth cycle. 5 µl samples were collected in a 96 well plate, and then separated simultaneously in a 1.5% TBE gel.

2.9 IT Methods

2.9.1 Creation of a chromatin sequence database

To catalogue the process of obtaining PCR-amplified fragments for spotting on a microarray, and to deal with the different sources of each fragment, a database was created in Microsoft Access 2002 (Richmond, VA) to record details of amplification template, primer details, history of amplifications and sequence data. This database is included in the enclosed CD.

2.9.2 Creation of a sample database

This work involved the collection of many samples that were to be processed in different ways, stored both for short- and long-term periods, and used for the generation of large number of images. In order to collate this data in a manner that allows logical, structured retrieval a database was developed using MS Access (Figure

2-1 and Figure 2-2). Although the database was initially developed for material harvested in Winter 2001, in which image storage was the primary concern, the harvesting of Winter 2003 and Summer 2004 was more complex, and therefore a much more comprehensive version was produced, and is discussed here. The database is relational comprising distinct tables. Each plant used in these crosses is detailed in a table of plant details [Plant Details] as both a female and male parent. This table details the date planted, the genotype of the plant and whether it proved positive or not for the presence of the GUS transgene.

For each time point a cross was harvested (some crosses studied at multiple time points), data was put in another table, [Cross], this detailed the pollen and ear parents, the date of harvest and the number of days after pollination, and also stored details of the ear: an image, a text description, and approximate fertilization success rate, details of the number of kernels stained for GUS, and the number of those that were positive. Material that was fixed was stored in a separate table, [Fixation Cassette], detailing the fixative, and the storage location. The [Seed Mass] table stored details of kernel weight measurements. Images taken using the microscope were stored in the [Microscope Images] table , detailing the microscope setup and sample staining details. A separate table of images was produced for those images taken without the aid of a microscope [HandCut Images], and further tables were also created to detail material harvested and stored for down-stream applications ([Starch], [PCD], and [RNA]). The data was linked via different primary keys, as displayed in the relationship table presented in Figure 2-2, whilst Figure 2-1 displays a screenshot from the database. The database was then able to provide information about biological samples by use of queries in a logical and effective fashion.

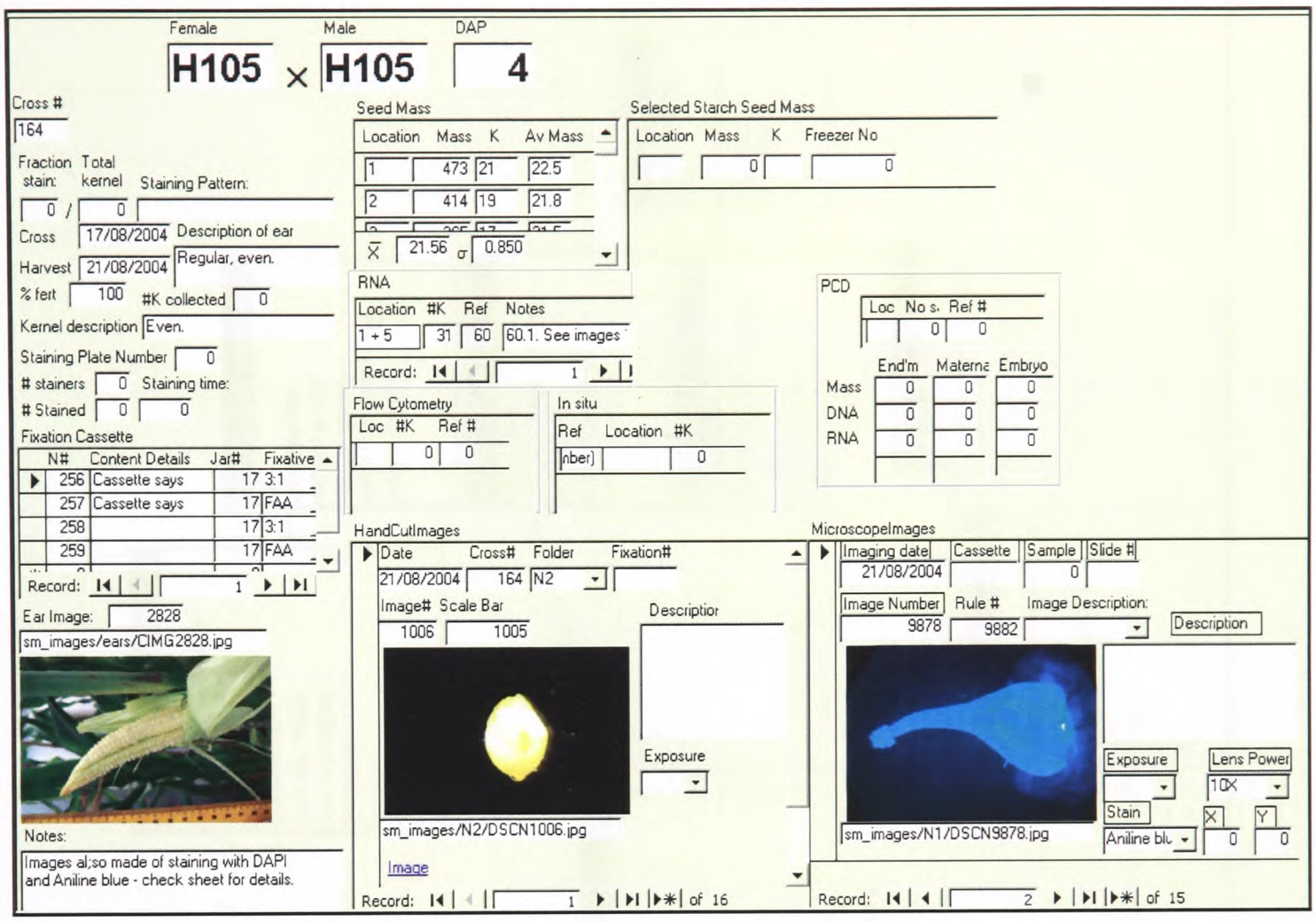


Figure 2-1 Screenshot from sample database

A database was created in MS Access of samples harvested, detailing mass of kernels, downstream applications (GUS staining, fixation, storage for analysis etc), alongside a record of photographs taken for kernels from each cross. This screen represents the main interface: a form called 'Crosses', which comprises several separate sub-forms, as detailed in Section 2.9.2.

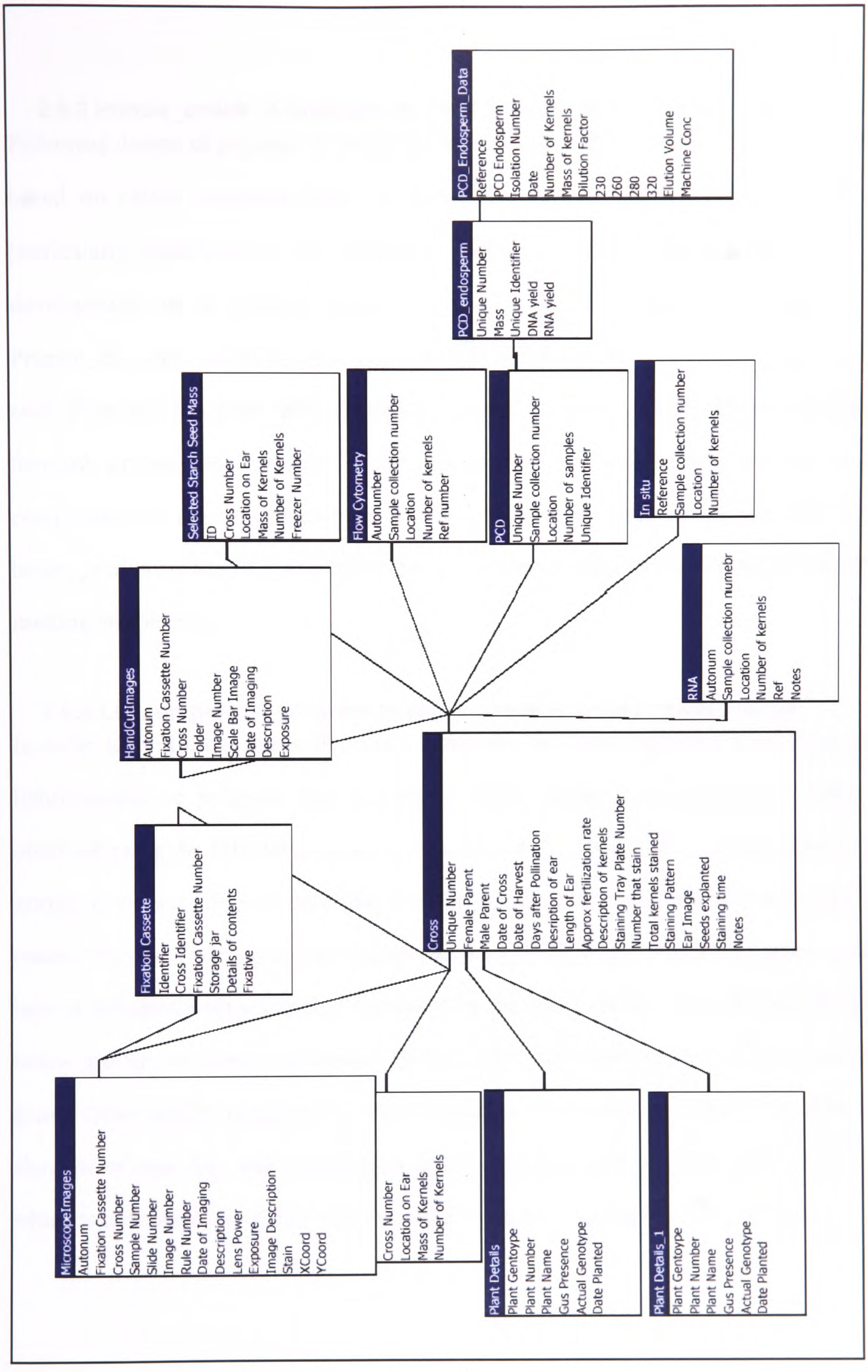


Figure 2-2 Table of relationships for kernel database

Summary of tables, including all field names for the Sample database, also detailing nature of relations between these tables.

2.9.3 Primer_check: A program to check primers are in a given sequence

Following design of primers, it was desirable to check that a product could be amplified based on cDNA sequence data, to ensure no mistakes had been made, this was particularly important for the design of almost a hundred primer pairs during the development of a custom array. To do this, a PERL script was written, Primer_check.pl, which accepts in a list of forward and reverse primer sequences for each gene and the gene DNA sequence in FASTA format. The program verified the forward primer was present in the sequence, and also looked for the reverse complement of the reverse primer. It also reported the distance between the two, in bases, predicting the PCR amplification product size, and also providing an estimate of melting temperature.

2.9.4 Listcompare: A program to count common genes between lists

In order to find those genes that were either up- or down-regulated in more than one hybridisation, a program was written in PERL called Listcompare.pl. Data was obtained using the MIDAS program component of the TIGR TM4 package (Saeed et al., 2003), to obtain slices of data that was greater than a certain amount from the local mean. This data was processed to generate a tab-limited file, which contained multiple lists of columns with a heading representing the hyb number, and whether the genes below are up- or down regulated. Each header has underneath it a list of all those genes differentially regulated in that condition. The program then compares each element of each list, and counts how many common elements there are in each list, which is outputted as a table, and also a list of those common elements is included.

3 Properties of the plant lines used in this work

3.1 Introduction

The aim of the work was to understand how development of seed resulting from interploidy crosses differs from seed resulting from balanced crosses. Such analyses in maize could, in theory, be confounded by differences between genetic lines used, which are extensive and well documented, (see chapter 1.1). The current chapter details the rationale for the choice of the plant lines used, describes some of the biology of these plants as assessed during the experiments, and sets out the justification for the validity of the use of the plants in this thesis.

3.1.1 Origin of plant lines used for different aspects of the work

Ideally, isogenic inbred lines should be used to eliminate the potential effects of different genetic backgrounds. However, this is not always possible or practical. The interploidy crosses investigated in this thesis (Chapters 4 and 5) make use of transgenic reporter lines that mark specific cell domains in the endosperm and allow an analysis of how these domains (sub-chapter 1.2.4) are affected in the two reciprocal crosses. Owing to the need for plant lines which can be readily transformed, these plants have a hybrid origin, a background which can be transformed with comparative ease (Armstrong et al., 1991). Because tetraploid varieties of these transgenic lines were unavailable, it was necessary to use a different line for the tetraploids. Further, due to the low fertility observed in the tetraploid line used in primary investigations, an alternative tetraploid line was also used in later studies.

3.1.1.1 Lines used in microarray and histological studies

For the microarray experiments (chapter 5), it was viewed as essential to cross plants which were genetically identical. The sensitivity of the system, and the sampling of

such a wide proportion of the transcribed genome means that any differences in the transcriptome of different lines could be observed, and confound true interpoloidy effects. Commonly used tetraploid lines have a hybrid origin so an alternative was required (Charlton et al., 1995; Leblanc et al., 2002). A novel technique of obtaining tetraploid plants by using nitrous oxide gas to induce chromosome doubling in diploid progenitors has been developed (Kato, 1998; Kato, 2002) and a number of such tetraploid lines derived from inbred progenitor lines were supplied by Dr James Birchler (University of Missouri). These plants were grown and compared for suitability for the experiments. Of these, the H99 tetraploids had the greatest germination success, and seemed to be the most vigorous, so were selected for further experiments.

H99 is an important line, since it is a dent corn variety, and is genetically similar to commercially grown varieties. Therefore, any potential advances indicated by this work should be more readily transferred to lines of agronomic importance, than would data emerging from other lines more commonly used in research.

H99 was also used in preliminary histological studies detailed in chapter 4. However, during these studies, it was realised that the tetraploid was not particularly fertile. Crosses often failed to take when used as a male parent, and many plants never bore ears, precluding full investigations as a female parent. Furthermore, the H99-4x plant was vulnerable to infection (possibly viral, Dr S. Gurr, personal communication), which often restricted its development. It has been reported that maize tetraploids derived from direct chromosome doubling are much less vigorous than those of a hybrid origin (Punyasingh, 1947). For this reason, following an initial analysis of interploidy effects on cell biology in winter 2001, in which only a limited number of timepoints were

observed, another tetraploid line, HybT detailed below, was employed for the broader time-scale analysis.

3.1.1.2 Tetraploids used in crosses with reporter lines

Crosses were made between the different inbred tetraploid lines supplied from Missouri and these yielded one plant which was particularly healthy and vigorous, displaying signs of hybrid vigour. This plant was obtained through crossing a tetraploid OH43 female with a tetraploid W23 male.

Like H99, both of these inbred lines are dent varieties of maize, and OH43 and H99 show a high degree of genetic similarity as assessed by random amplified polymorphic DNA (RAPD), simple sequence repeats (SSR), and amplified fragment length polymorphisms (AFLP) (Pejic et al., 1998). W23 is the plant line into which the transgenic lines used in these studies had been introgressed.

While this new tetraploid variety had the disadvantages of being neither isogenic, nor exactly the same genetic background as the plants with which it was crossed, these were outweighed by the advantages of it being healthy, and highly fertile. Additionally, the fact it shared genetic similarity with both the transgenic reporters, and the other tetraploid line (H99) used in this thesis, further added to the validity of its use.

3.1.1.3 Reporter lines; pBETL::GUS

This line contains a transgenic construct consisting of the promoter for an endosperm cell domain reporter gene, (*BETL1a*), the β-glucoronidase (GUS) reporter gene (Hueros et al., 1999a), and the selectable marker neomycin phosphotransferase (NPT). The line was generated in diploid lines (HiII, (Armstrong et al., 1991)), and then introgressed into a W23 diploid background through five to six generations.

The promoter carried by the pBETL::GUS reporter line consists of 985bp fragment, shown to confer transfer cell specific expression onto endosperm (Hueros et al., 1999a). Only a broad pattern of activity has previously been published, and for this reason it was considered critical to perform a detailed developmental time course for the present work.

3.1.1.4 Reporter lines; pVp1::GUS

The genetic construct and genetic origin of the other reporter line used in these studies were essentially same as for the pBETL::GUS line, except for the promoter region. The line carrying the pVP1::GUS reporter contains a construct of approximately 1kb of promoter region for the *Vivaparous1* (*Vp1*) gene (McCarty et al., 1991), in addition to the GUS transgene, and an NPT selectable marker. The pVp1::GUS line was supplied by Dr P. Perez of Biogemma (France). The plant was also derived from HiII plants, followed by several generations of introgression into the W23 diploid background. The pattern of reporting of this transgenic line has been described previously (Costa et al., 2003), although a detailed developmental series had yet to be carried out.

3.2 Results

3.2.1 Flow Cytometry confirms the ploidy of plants

In order to confirm that the plants used in these experiments were genuine tetraploids, leaf samples were isolated, and flow cytometry, using the FL2-A channel units which measure propidium fluroescence, as a measure of nuclear DNA content, which should in theory be exactly double that of the diploids (Galbraith, 1983).

Several comparisons were made to estimate the ploidy of plants. Initially, diploid leaf material was analysed simultaneously with an exogenous control, tomato (*Solanum lycopersici*), to ensure an expected ratio was observed between the peaks of non-dividing cells (interphase nuclei). Maize, which has a 1C value of 2.73 pg, would be expected to have a peak 2.65 times that of tomato (1C=1.03), as detailed at the Plant DNA C-value database[4]. However, known interlab variability (Dolezcaronel et al., 1998), and variation between maize lines (Laurie and Bennett, 1985), means a range of values could be observed between 2.4 and 2.88. Hence, the observed value of 2.76, is well within the expected range (Table 6A).

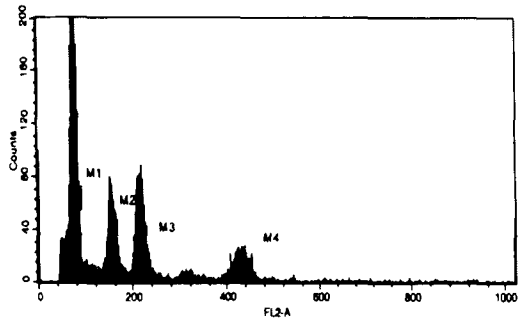
Direct comparisons were then made between diploid and tetraploid maize. Here, three peaks would be expected, one representing the non-dividing cells of the diploid, one representing the dividing diploid cells and the non-dividing cells of the tetraploid, and one representing the post S-phase cells of the tetraploid (Table 6B). Following confirmation of a true tetraploid by this means, samples of other tetraploids were pooled, to ensure there was only one dividing and one non-dividing peak, and a co-efficient of variance of a similar size to plants in isolation. This was the case for all pooled samples of tetraploids (data not shown).

A very low percentage of plants could be recovered from interploidy crosses. To ensure they were the products of true interploidy crosses, rather than from stray pollen of the same ploidy as the maternal parent, leaf material was harvested. It would be expected that the non-dividing peak of nuclei from this plant would be located exactly half way between the diploid and tetraploid parents. Table 6C displays a flow cytometry trace in which diploid, tetraploid and triploid plants were run simultaneously. As can be observed, three major peaks are observed. The predicted ratios between the positions

[4] http://www.rbgkew.org.uk/cval/

of the peaks on the x-axis closely match their expected values: the ratio between the triploid peak and diploid peak is 1.45 (expected value: 3/2 or 1.5), whereas the ratio between the tetraploid and diploid is 1.38 (expected: 4/3 or 1.33).

Diploid Maize and Tomato

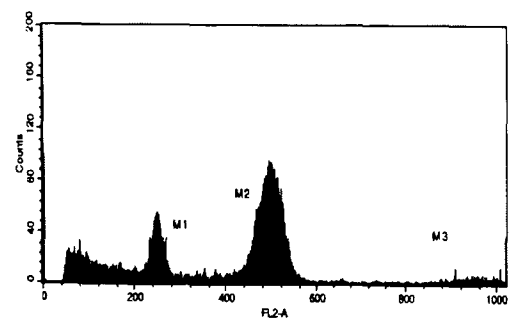


	Tomato		Maize 2x	
	NDP	**DP**	**NDP**	**DP**
Peak	**M1**	**M2**	**M3**	**M4**
Mean	79.41	160.00	219.50	432.40
SD	0.20	0.94	0.43	1.50

M2/M1	2.01
M4/M3	1.97
Maize/ Tomato	2.76

A

Diploid and Tetraploid Maize

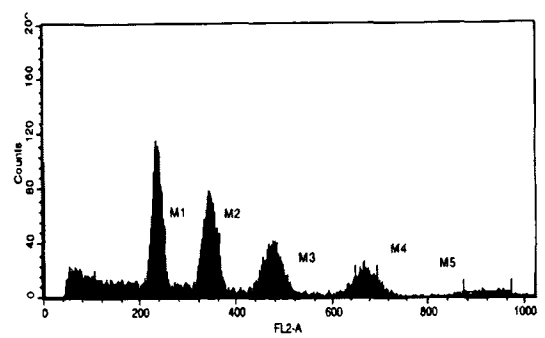


	Maize 2x		Maize 4x	
	NDP	**DP**	**NDP**	**DP**
	M1	**M2**	**M2**	**M3**
Mean	252.50	500.50	500.50	954.90
SD	0.71	1.49	1.49	5.47

M2/M1	1.98
M4/M3	1.91
M3/M1	3.78

B

Diploid, triploid and tetraploid maize



	Maize 2x		Maize 3x		Maize 4x	
	NDP	**DP**	**NDP**	**DP**	**NDP**	**DP**
	M1	**M3**	**M2**	**M4**	**M3**	**M5**
Mean	239.90	482.90	349.90	675.30	482.90	930.80
SD	2.32	4.65	2.76	3.58	4.65	10.14

M3/M1	2.01	**M2/M1**	1.45
M4/M2	1.93	**M3/M2**	1.38
M5/M3	1.93		

C

Table 6 Sample of flow cytometry results

Images show sample traces from flow-cytometry runs in which the plant lines mentioned were run together. Tables below show the mean peak size (FL2-A, linear scale), and the standard deviation between three separate runs. NDP = Peaks representing non-dividing cells, DP = Peaks representing dividing cells.

3.2.2 Flowering time of maize lines

In order to assess similarities in the biology of the plants used in this work, the time taken from sowing seeds to flowering was measured. The hybrid plants, the pVp1::GUS line and pBETL::GUS lines, were very similar in habit (Figure 3-1A), which is unsurprising, as they have the same genetic origin and were both introgressed into the same genotype by several generations. The plants were tall, and healthy with broad green leaves, producing an abundance of pollen at anthesis. The HybT line was also vigorous, growing to a similar height as the two reporter lines, and producing an large amount of pollen (Figure 3-1D). The H99-2x inbred line was shorter than the reporter line diploid plants (Figure 3-1B), and was generally less vigorous. Most revealing was the phenotypes of the H99-4x plants (Figure 3-2C), which despite being genetically identical to the H99-2x diploid plants, were considerable shorter, showed low fertility, producing only small amounts of pollen, and often failed to produce a receptive ear. *In vitro* germination techniques showed that the number of pollen grains germinated from these tetraploids was comparable to the diploid lines used in this study (not shown). However, when applied to pVp1::GUS line and pBETL::GUS line ears in paternal excess crosses, only around 10-20 kernels were obtained per ear, (representing only 5-10% of possible fertilization events) pointing to lowered pollen tube viability. Figure 3-2 shows the mean time taken until initial shedding of tassels and receptivity of ears for the all plants used in these studies. The three diploid lines and the HybT tetraploid flowered after a similar period (although the latter did take slightly longer), and the time between anthesis and female ear receptivity was longer in the tetraploid line. The H99-4x plants also took, on average, almost two weeks longer than their diploid progenitors to start shedding pollen, and in the study used for this graph, none produced fertile ears.

Kernels resulting from interploidy crosses were sown, to confirm if any plants obtained had the predicted 3x constitution. A few plants were recovered, which were healthy and produced viable pollen and ears, although the mean time to flowering was longer than for either of their parents Figure 3-2. These plants were fertile, and crosses took well both as male and female parents. However, the kernels produced from crosses with diploids and tetraploids showed a range of phenotypes ranging from almost normal to either maternal excess or paternal excess, depending on the cross. This suggests that during meiosis in these triploid plants chromosome segregation is disturbed, resulting in gametes carryingvariable proportions of the genome.

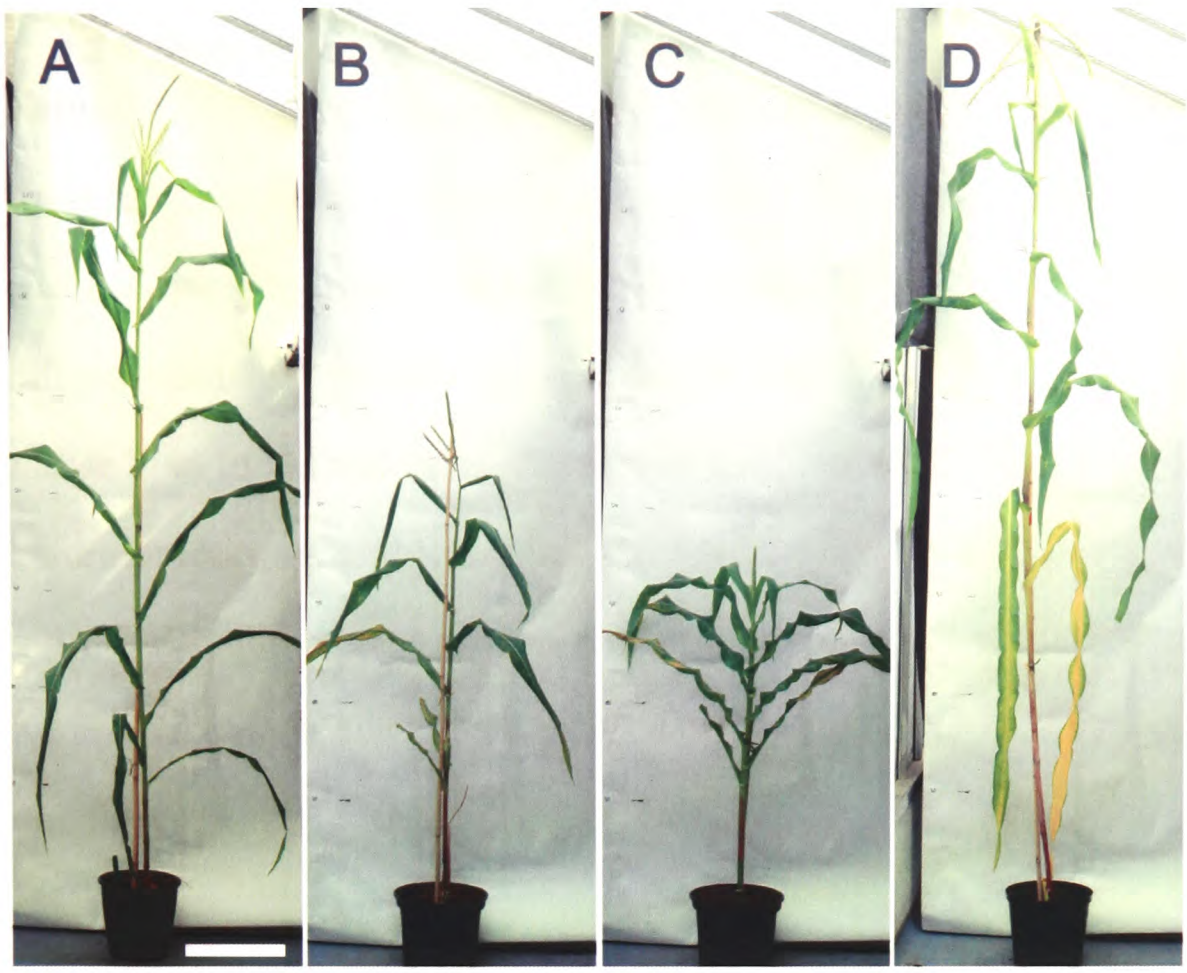


Figure 3-1 Images of mature, flowering plants .

Flowering plants were photographed at maturity **a**)pVp1::GUS line, **b**)H99-2x **c**)H99-4x **d**)Hyb-T. Scale bar = 30cm.

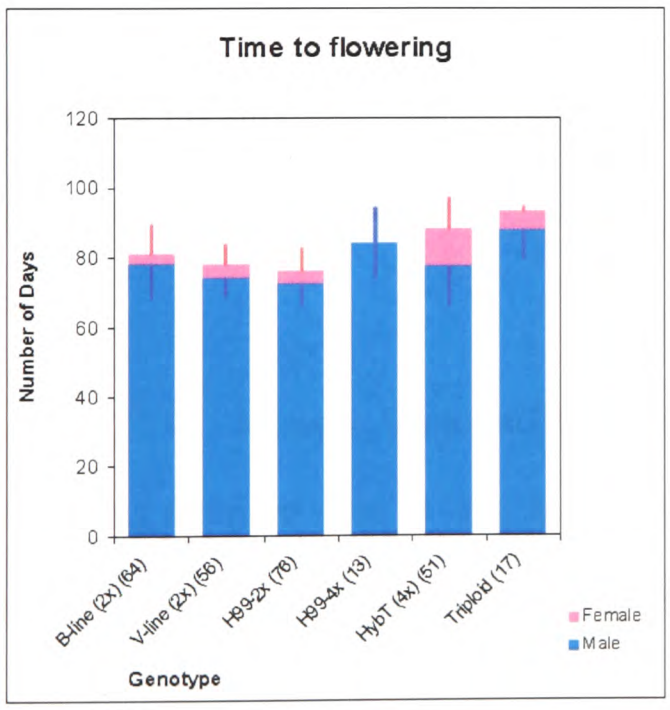


Figure 3-2 Time to flowering

Graph shows the time taken to flowering, in days from sowing until initial shedding of pollen (male), and initial receptivity of the ear (female). Error bars shows the standard deviation mean number of days to flowering, and the number in brackets on the x-axis details the sample size

3.2.3 Seed mass variation between lines in balanced crosses

Natural variation within maize, and artificial selection performed to improve kernel traits has generated enormous heterogeneity between different lines. For this reason, it was necessary to assess the variability within the lines used in this thesis so that any differences in interploidy crosses can be attributed to genuine interploidy effects - rather than to natural variability within the lines used. Harvested seeds were weighed before processing, and the data for seeds between 6 and 24 DAP are shown in Figure 3-3A and B for crosses with the pVp1::GUS line and pBETL::GUS line respectively. Additional data is displayed for selfed-crosses of the Hyb-T tetraploid and the H99-2x diploid.

All lines show a similar steady increase in seed mass, although there is variation in both final seed mass and rate of filling. The graph demonstrates that when H99-2x plants are crossed as a female parent with either reporter line, the resultant seed weigh less than in the reciprocal cross, varying from 50 to 83% of the higher seed mass, (average 73%). Interestingly, however, when the pVp1::GUS line is selfed (Figure 3-3A), at 24 DAP the kernels are only 4% heavier than kernels from crosses where it was used to pollinate H99-2x plants.

At 24 DAP, kernels from selfed H99-2x plants are 50% of the weight of kernels resulting from crosses in which H99-2x is used as a female parent, and fertilised with pBETL::GUS line pollen (249 versus 126 mg).

Kernel masses for selfed HybT × HybT plants demonstrate seed masses generally comparable with crosses between the two reporter lines when pollinated by H99-2x plants.

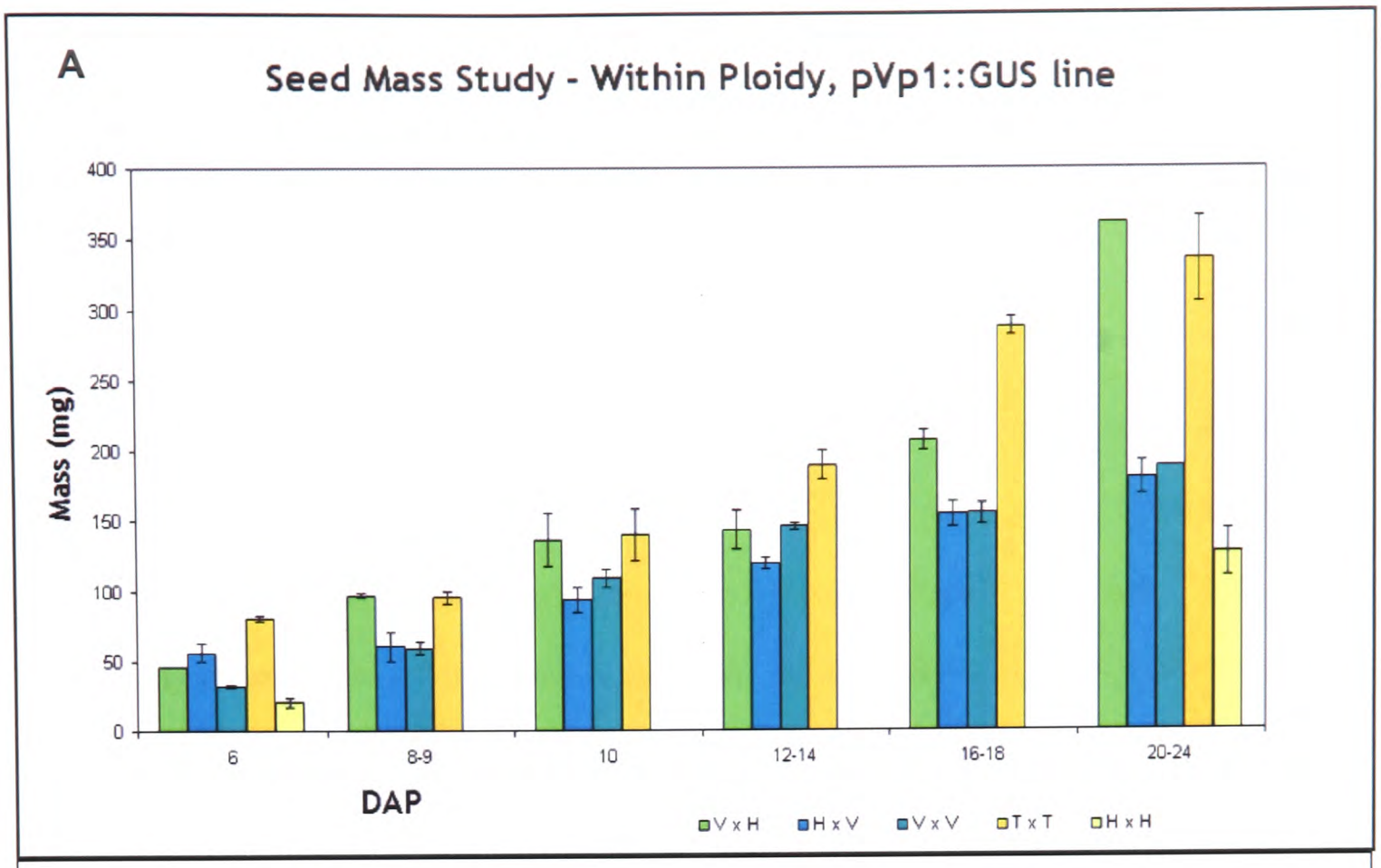


Figure 3-3 Increase in seed mass following within ploidy pollinations

Individual or pooled kernels were weighed and an average mass obtained, for seeds obtained straight from ears at the timepoints indicated on the x-axis. Error bars show the standard deviation of the mean. V=pVp1::GUS line, B = pBETL::GUS line, H=H99-2x, T=HybT tetraploid. Directions of crosses displayed in the legend refer to the maternal parent first and the paternal parent second. **A**) Crosses involving the pBETL::GUS line, **B**) Crosses involving the pVp1::GUS line

3.2.4 Interploidy crosses result in profoundly reduced seed masses

Figure 3-4A and B show the summary of seed masses generated by interploidy crosses using either the pVp1::GUS line or pBETL::GUS line as the diploid parent, and diploid and tetraploid selfed controls, presented previously in Figure 3-3, for comparison.

These graphs clearly demonstrate the difference between interploidy crosses and balanced controls. Instead of a rapid rate of increase in mass of several hundred percent between 6 and 25 DAP, as in the control crosses (Figure 3-3), both directions of interploidy cross show a very limited increase. At the earliest stages of development, the maternal excess kernels are comparable with and even heavier than those of the controls in crosses involving the pVp1::GUS line. But after 12 DAP, the control seeds increase in mass at a rate far greater than the maternal excess kernels. At every time-point observed the paternal excess kernels weigh less than the balanced controls.

Figure 3-5 also displays a direct comparison of the two reciprocal interploidy crosses for those crosses in which the pBETL::GUS line was used as the diploid parent. This graph clearly demonstrates the differences between the two interploidy directions. At earlier stages of development (until 14 DAP), the seed masses are comparable, but during the later stages, the maternal excess kernels actually decrease in mass, whereas the paternal excess kernels continue, on average, to increase in mass. However, the error bars also demonstrate the striking variability in the seed mass of paternal excess endosperm when compared with maternal excess endosperm.

3.2.5 Interploidy kernels show low germination rates

Seeds which had been left on the ear to mature for at least 40 DAP and further left to dry at room temperature for several weeks were weighed and then sown to observe germination rates. Images of dried kernels are shown in Figure 3-6, where it is clear that the same ploidy crosses (a and b), show plump, round healthy kernels, whereas the two interploidy crosses (c and d) show small shrivelled kernels. Individual kernels were weighed, and the mass of these are presented in Figure 3-7.

Kernels were sown and rates of germination success were monitored between seeds from the two interploidy directions. On average 1.7% of kernels from paternal excess crosses (2x × 4x) germinated and produced healthy plants, whereas 0.83% of maternal excess seeds produced viable seedlings (Figure 3-8). Often more kernels would germinate, but shortly die before the production of the first true leaf, and sometimes some seedlings . Although there was some variability between the mass of the kernels, there was no strict correspondence between seed size and likelihood of germination.

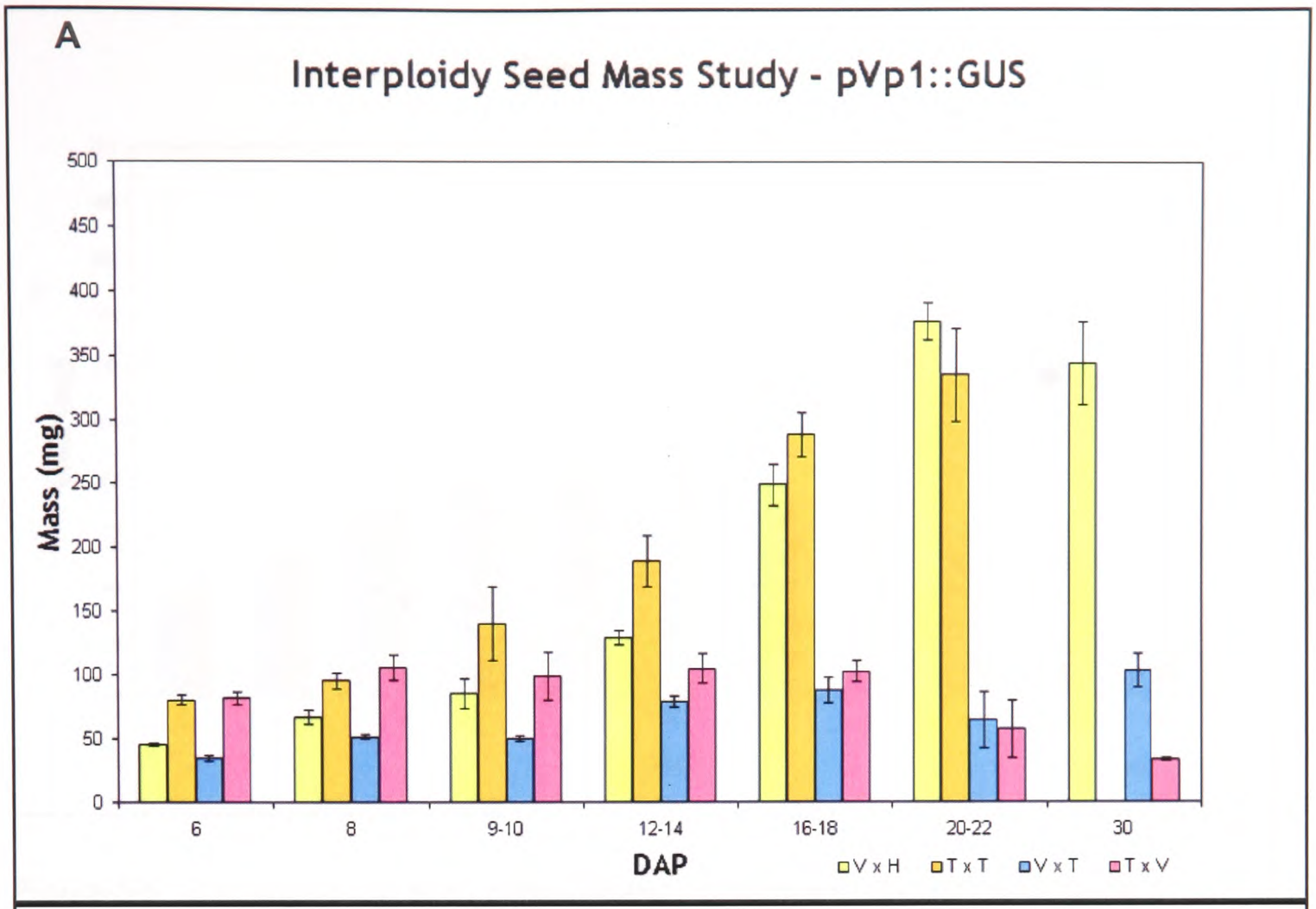


Figure 3-4 Increase in seed mass for interploidy crosses.

Individual or pooled kernels were weighed and an average obtained, for seeds obtained straight from ears at the timepoints indicated on the x-axis. V = pVp1::GUS line, B=pBETL::GUS line, H=H99-2x, T=HybT tetraploid. Blue bar represent 'paternal excess' seeds, and pink bars represent 'maternal excess seeds'. Orange and yellow bars are for within ploidy controls. Error bars show the standard deviation of the mean. **A**) Crosses involving the pBETL::GUS line, **B**) Crosses involving the pVp1::GUS line.

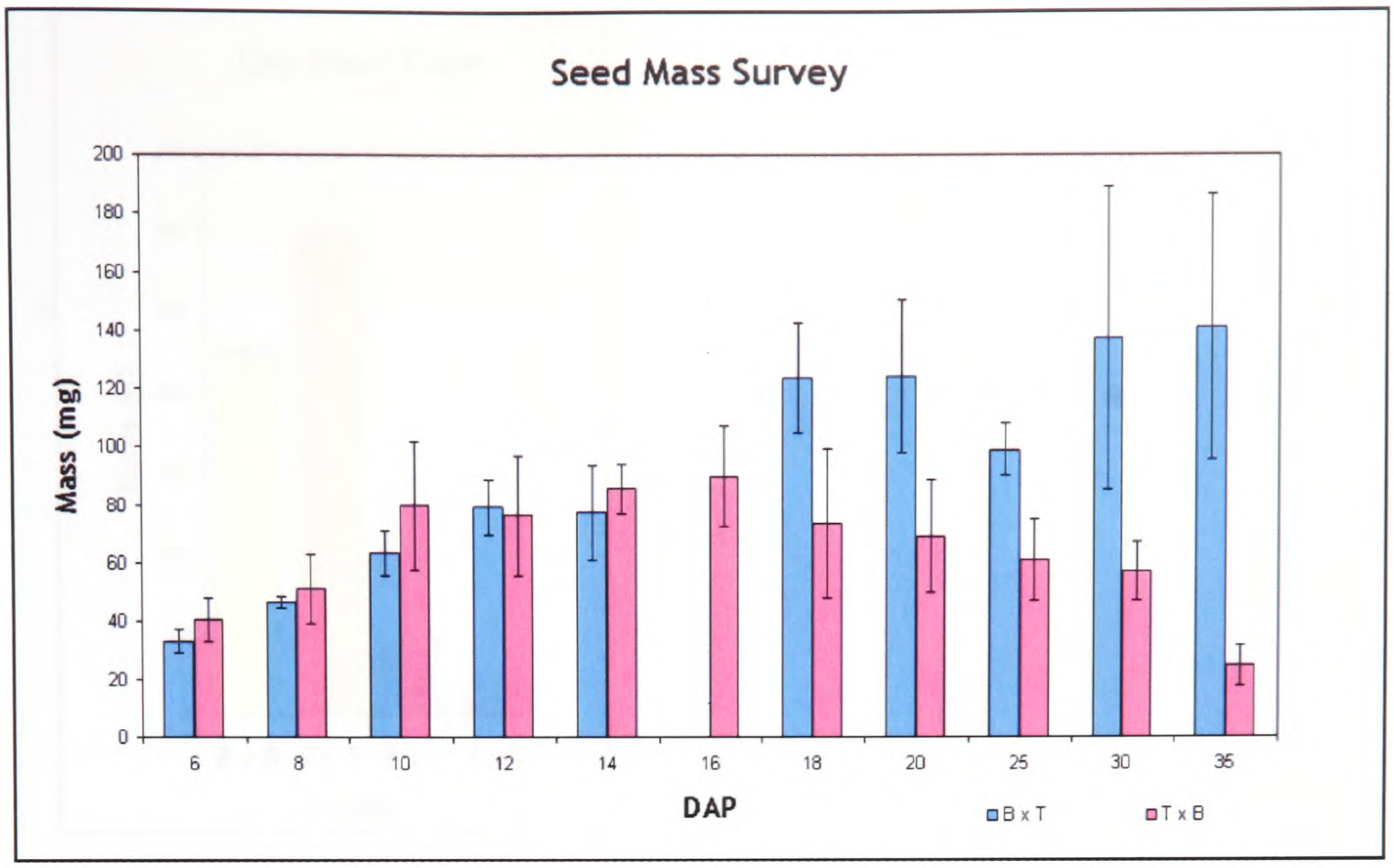


Figure 3-5 Comparison of seed mass between reciprocal interploidy crosses.

Individual or pooled kernels were weighed and an average obtained, for seeds at the timepoints indicated on the x-axis. B=pBETL::GUS line, T=HybT tetraploid. Blue bar represent 'paternal excess' seeds, and pink bars represent 'maternal excess seeds'. Error bars show the standard deviation of the mean.

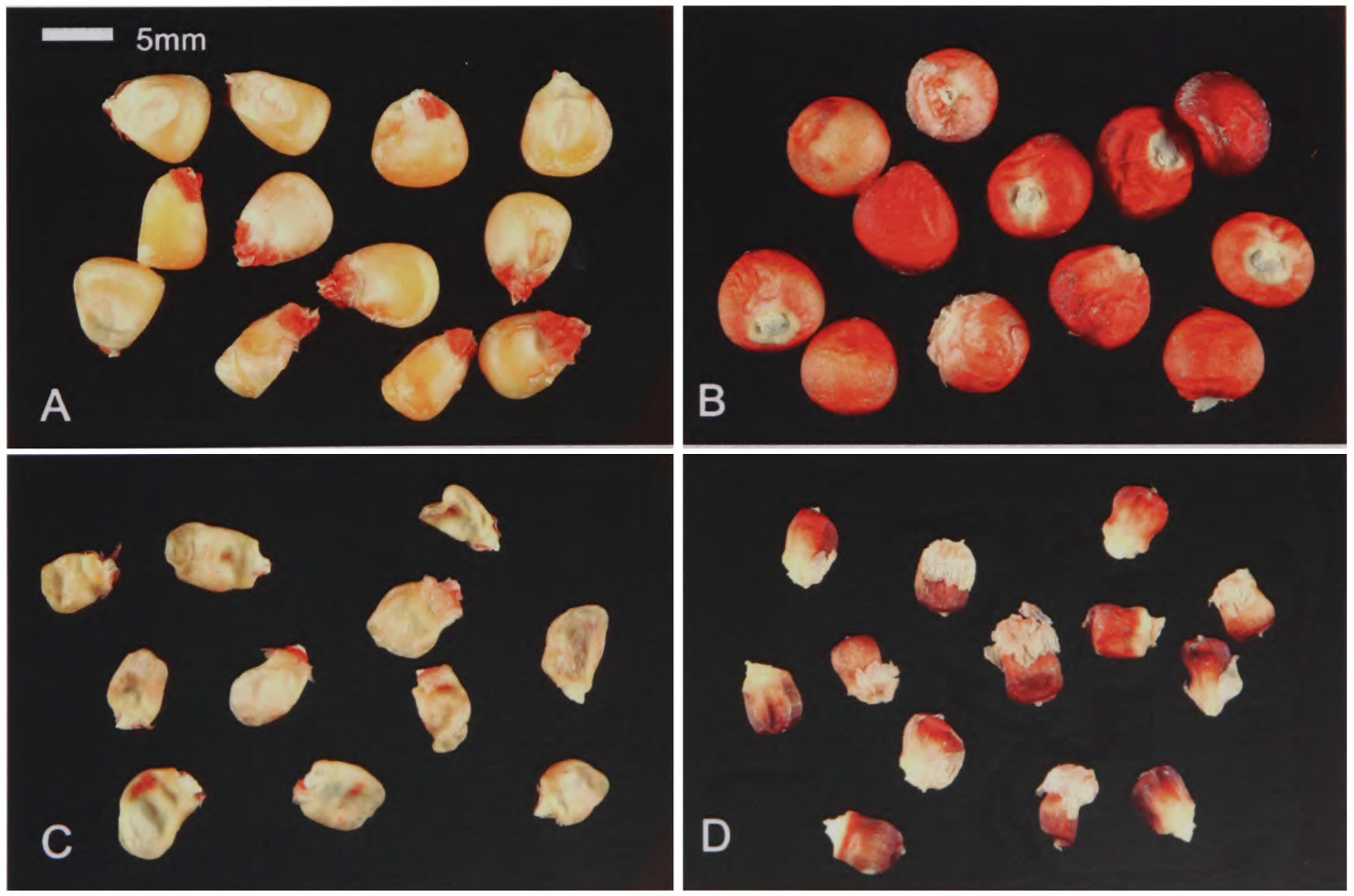


Figure 3-6 Mature, dried kernels

A) 2x × 2x (pBETL::GUS line selfed), **B**) 4x × 4x (HybT tetraploid selfed), **C**) 2x × 4x (pBETL::GUS line x HybT-4x), **D**) 4x × 2x (T-line x pBETL::GUS line).

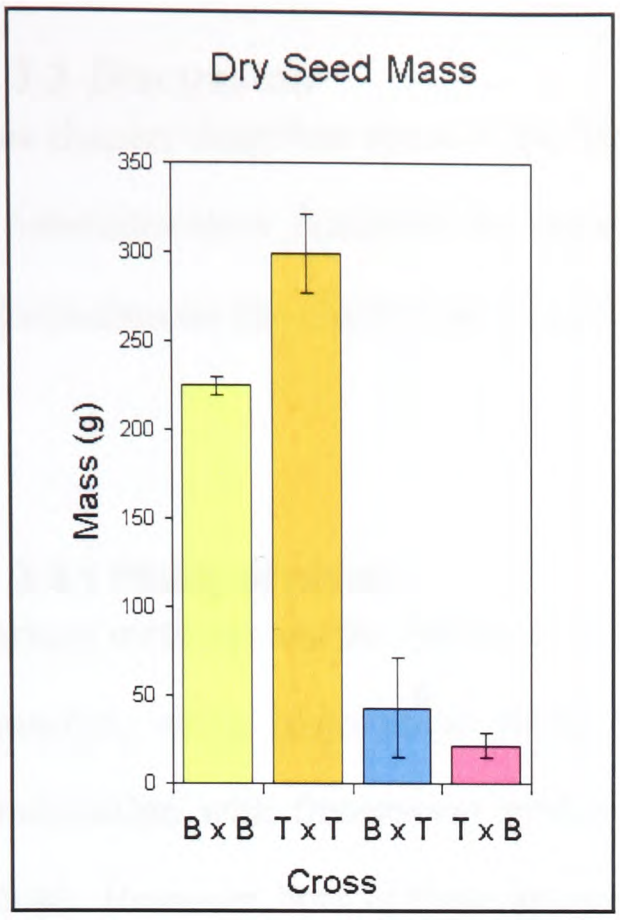


Figure 3-7 Seed mass of mature, desiccated kernels

Kernels were left on the plant for 40 days, and allowed to dry further on the ear for a month. Individual kernels were weighed. **a**) 2x × 2x (yellow), **b**) 4x × 4x (orange), **c**) 2x × 4x (blue, 'paternal excess'), **d**) 4x × 2x (pink, 'maternal excess'). Error bars show the standard deviation of the mean.

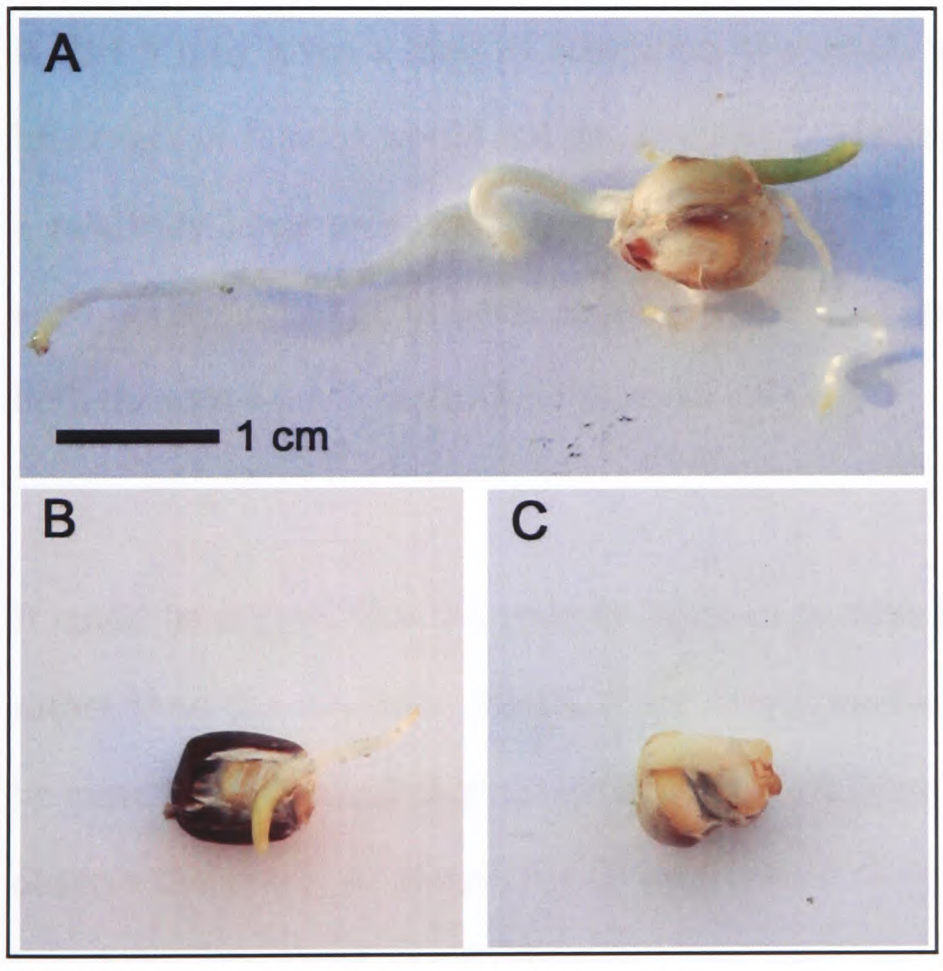


Figure 3-8 Germinated triploids kernels

Kernels were sown in sand and observed after 10 days. **A**) Kernel from a balanced diploid cross, **B**) Kernel from a maternal excess cross, **C**) Kernel from a paternal excess cross.

3.3 Discussion

This chapter describes some of the biology of the plant lines used in the work reported to determine their suitability for the later experiments. It has also provided some novel information on the effect of interploidy crosses on seed mass dynamics.

3.3.1 Ploidy of plants

Various methods can be employed to verify the ploidy of plants, including chromosome squashes, using histological stains for chromosome counting (Kato, 1997), or in combination with fluorescent probes to regions of the chromosomes (Walling et al., 2005). However, both of these approaches are more time-consuming, both to set up an initial working protocol, and in terms of experimental time. Flow cytometry on leaf material, however, has provided a rapid method to screen plants, and it has also proved possible to screen multiple plants simultaneously. The main drawback of this method is that it only gives a limit of resolution of a single genome, hence any chromosome breakages or fusions would not be recognized. Moreover, the system is known to have a relatively large technical variance (Dolezcaronel et al., 1998), so in theory partial aneuploidies could be missed. However, the latter argument could be levelled against both the root squash methods mentioned above.

It could be argued that it would be more important to study the ploidies of gametes, rather than the vegetative tissue, since chromosomal translocations and losses would be more likely to occur during reproductive development. Most instructive would be to observe the ploidy or chromosomal appearance of the endosperm itself, as this would reveal the actual part of the genomes which were successfully transmitted. This may be an appropriate method in future experiments in which one half of the endosperm would be saved and referenced, and the other half used in alternative studies such as gene expression analysis. Any inter-kernel variability could then be correlated with the

actual genomic composition of the endosperm. However, for the broad survey of interploidy effects reported here, this would have been unnecessary. The effects of interploidy crosses have been demonstrated to result from the changes to the distorted genome balance, rather than the partial transmission of genomes through the pollen of tetraploid plants (Punyasingh, 1947). Furthermore, complete genome complements may not even be necessary to generate the major effects of interploidy crosses, if the factors responsible (such as imprinted genes) are numerous and located throughout the genome as suggested the reduced seed size caused by maize chromosomal arm deletions (Birchler, 1993; Lin, 1982).

In summary, the methods employed confirmed the ploidy of the plants to a resolution necessary for the subsequent research. Confirmation that the viable seed recovered from interploidy crosses were triploid (Table 6) further supports the applicability the approach adopted for this work.

3.3.2 Seed mass effects between diploid plants demonstrate parental and hybrid effects

Consistent with the genetic variability reported in maize, especially for seed traits (Soave and Salamini, 1984), crosses between different diploid lines displayed different rates of filling and final seed masses. Most interestingly, however was the variation in behaviour of crosses between different plant genotypes, which suggested effects such as hybrid vigour and female control of seed development, as have been reported to be important in seed development (Chaudhury et al., 2001). The smallest kernels observed were selfed crosses between the H99-2x plants, and after this were selfs between the pVp1::GUS line plants. Although having a hybrid origin, the pVp1::GUS line had been introgressed for five to six generations into W23, and so comprised a genome which was over 95% inbred origin. Crosses between these two lines, however

resulted in considerably larger kernels than either, suggesting hybrid vigour. Intriguingly, however, these effects were most pronounced when the pVp1::GUS line plants were used as the female parent, reflecting the parental differences in control over seed development. This supports the recognized importance of the female parent, controlling seed size, but still demonstrates that the paternal genome does have an effect (Birchler, 1993).

3.3.3 Seed mass following interploidy crosses

In spite of several previous studies of interploidy effects in maize which suggest that the ultimate seed size is reduced, none have provided an indication of the dynamics of increase in mass. The current work has provided an illuminating insight into grain filling following maize interploidy crosses.

Seeds from paternal excess crosses weigh as little as 20% of the final weight of selfed diploids from crosses with the same maternal genotype. The effect is even more extreme for maternal excess kernels, which show a seed mass as low as 7% of the mass of a tetraploid sharing the same maternal genotype. Although grain filling is initially more rapid for the maternal excess crosses, seeds appear to peak in mass around 16 DAP, and then start to decrease. In normal kernel development a period of increase in mass occurs followed by maturation, during which there is a decrease in mass through dehydration. The main grain filling period, during which mass increases most dramatically, would normally start around 16 DAP and dehydration from around 30 DAP. Hence, it seems that the 'normal processes' of seed development have been accelerated in the maternal excess cross.

In contrast, the mean mass of kernels from paternal excess crosses increases throughout the period investigated, with no evidence of a desiccation period. There is much greater variability in the seed mass of these kernels.

The profiles of these mass accumulation graphs mirror closely the profiles of endoreduplication activity in maize interploidy crosses (Leblanc et al., 2002). Apparently, female excess kernels display an acceleration of the endoreduplication process, although never reach the levels of the diploids, in contrast the process is delayed in the paternal excess.

4 Specification of two cell domains and cellular effects of balanced and interploidy crosses

4.1 Introduction

Progression of development of maize kernels has been previously charted rigorously (Kiesselbach, 1949b), and is summarized in Chapter 1.2.4. The key phases of seed formation are: karyogamy, syncitial development, alveolation and cellularization, growth by centrifugal cell division, outwards growth by cell expansion, and maturation. Elaboration of the four functional domains of the endosperm, detailed in 1.2.4, commences from after cellularization of the coenocyte, and these regions persist into maturity, since each has developmentally important functions. Hence, charting how these domains are disrupted following interploidy crosses should give an indication into the disturbances which such crosses cause.

4.1.1 The *Vp1* Gene as a marker for the aleurone layer

The reporter gene which marks the aleurone is controlled by the promoter sequence for the *Vp1* gene (3.1.1.4). The Vp1 protein is a transcriptional regulator involved in the sensitivity of the kernel to absicissic acid (ABA), and hence important in seed dormancy and germination (McCarty et al., 1991), v*p1* mutants lead to precocious seed germination on the ear (Robichaud et al., 1980). The endogenous gene is expressed in both the aleurone layer of the endosperm, where it is first detectable at 6 DAP, but also in the embryo, in which expression occurs first in the mid-globular phase, from 2 DAP.

4.1.2 The *BETL1* gene as a marker for the basal endosperm transfer layer

The *BETL1* gene, the promoter of which drives expression of this domain marker (3.1.1.3), is a gene of unknown function, although may have a role in protection from pathogens, or possibly signalling (Hueros et al., 1995). The endogenous gene is first expressed from 7DAP, and its protein is first detectable from 10 DAP. A previous

report into this reporter line suggested that GUS signal was first detectable at 12 DAP (Hueros et al., 1999a).

4.1.3 Scope of this chapter

Clearly it is important to have a more detailed context of BETL and aleurone location than is currently available, and for that reason this chapter details the patterns of reporter activity in balanced diploid crosses with reporter lines for these cell domains. The reciprocal cross was deemed essential as the reporter lines will be used as both male and female parents when crossed with tetraploids, so it is importance to see whether the reporter is active when inherited from both the female and male parents alike . Finally, cell domain localization in interploidy crosses will be reported to determine how a disturbance in the parental genomes affects these two cell domains.

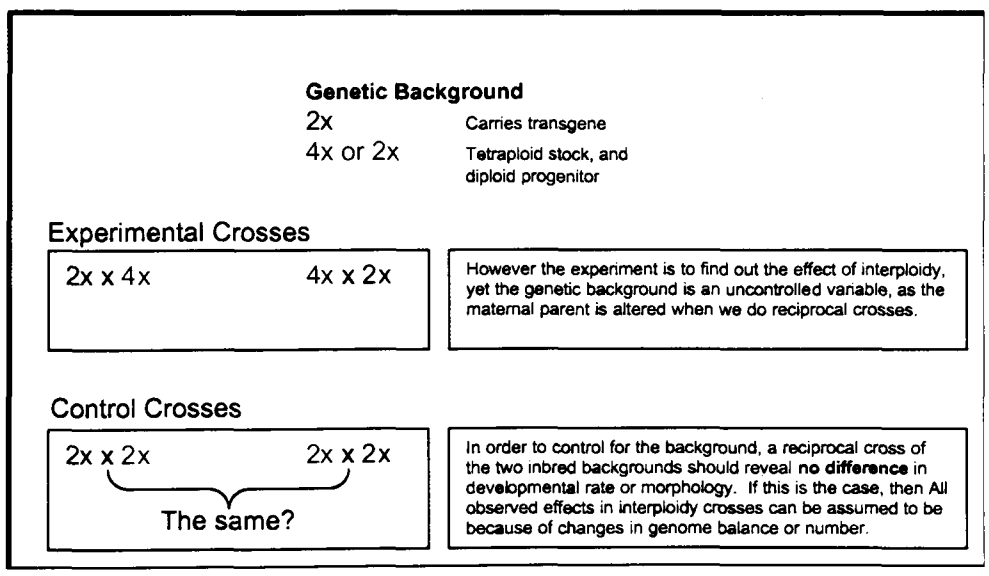


Figure 4-1 Logic for transgene parental origin control crosses

4.2 Results

4.2.1 Activity of the pVp1::GUS reporter gene in balanced crosses

Data is presented for reciprocal crosses between the pVp1::GUS line and H99-2x (Figure 4-3, pVp1::GUS line x H99-2x; Figure 4-4, H99-2x x pVP1::GUS).

GUS activity was first seen in crosses at 5 DAP (data not shown), where it was restricted to the embryo, with no sign of activity in the future aleurone. Embryo staining persists throughout the entire period of kernel development monitored in these crosses, until 24 DAP (pVp1::GUS x H99-2x, Figure 4-3, 24 DAP) or 34 DAP (H99-2x x pVp1::GUS, Figure 4-4, 34 DAP), and is seen throughout the embryo. The percentage of kernels which stain is close to the expected value of 50% on the basis of transgene segregation (Figure 4-2).

By 6 DAP GUS activity was first observed in the aleurone (Figure 4-3, 6DAP). At this stage GUS activity did not surround the entire endosperm periphery, but was restricted to regions within it - predominantly at the apex, which, at this stage of development, is still conical. In some samples, staining is also observed along the adgerminal endosperm margin. Sectioning of these samples combined with dark field microscopy revealed that although the GUS precipitate was generally localized to the periphery, it was also apparent in the underlying cells, presumably through diffusion of the product, since the intensity of the signal decreases with distance from the endosperm margin.

Between 6 and 9 DAP (Figure 4-3) isolated endosperms varied in the degree to which the periphery stains blue. At each successive time point the proportion of the outer margin which stains increases, apparently from the apex down the periphery. By 8 DAP most kernels showed complete peripheral staining, except for the area in contact

with maternal conducting tissue, which does not develop into aleurone tissue. Embryo staining also persists during this developmental phase.

Between 10 and 20 DAP (Figure 4-3, 10 - 20 DAP), GUS activity remains confined to the embryo and aleurone, or the endosperm periphery in spite of very drastic developmental changes taking place in terms of seed expansion. Surprisingly between 12 and 14 DAP a streak of GUS activity is observed along the central strand of the starchy central endosperm, a phenomenon which has not been reported before.

As the endosperm undergoes maturation, the GUS reporter continues to define the outer margin of the endosperm and the embryo, until as late a stage as 34 DAP (Figure 4-4, 34 DAP).

The developmental progression of pVp1::GUS aleurone reporting is very similar between the two reciprocal crosses, suggesting that by the time at which the promoter is active, there are limited differences in the expression of the transgene, irrespective of whether it originated from the paternally or the maternally inherited genome. There do not appear to be any differences in the location of the GUS activity.

4.2.2 pVp1::GUS activity precedes aleurone morphology in balanced crosses

GUS activity is first observed in the endosperm periphery at around 6 DAP. In order to investigate whether this expression pattern is concomitant with the assumption of typical aleurone cell morphology, samples were sectioned and stained with PAS for carbohydrate (Figure 4-5). These images reveal that the initial stages of GUS activity occurs prior to the adoption of aleurone morphology as the GUS precipitate (which appears red in dark field, DF, microscopy) is found throughout the periphery and

underlying cells (presumably through diffusion) of the 6 DAP endosperm (Figure 4-5A), yet PAS staining reveals that the cells located in this region are histologically identical to the underlying cells (Figure 4-5B). By 8 DAP the cells still show GUS in underlying cells, although it is not found as deeply towards the centre. Although at this time the peripheral cells are still not typical of mature aleurone cells, their morphology is now distinct from those beneath: being smaller in size, more regular in shape and having more substantial cell walls (Figure 4-5C and D). Only by 10 DAP are the peripheral cells unambiguously aleurone-like, being rectangular, compact, and forming a distinct layer they also stain intensely for GUS (Figure 4-5 E and F).

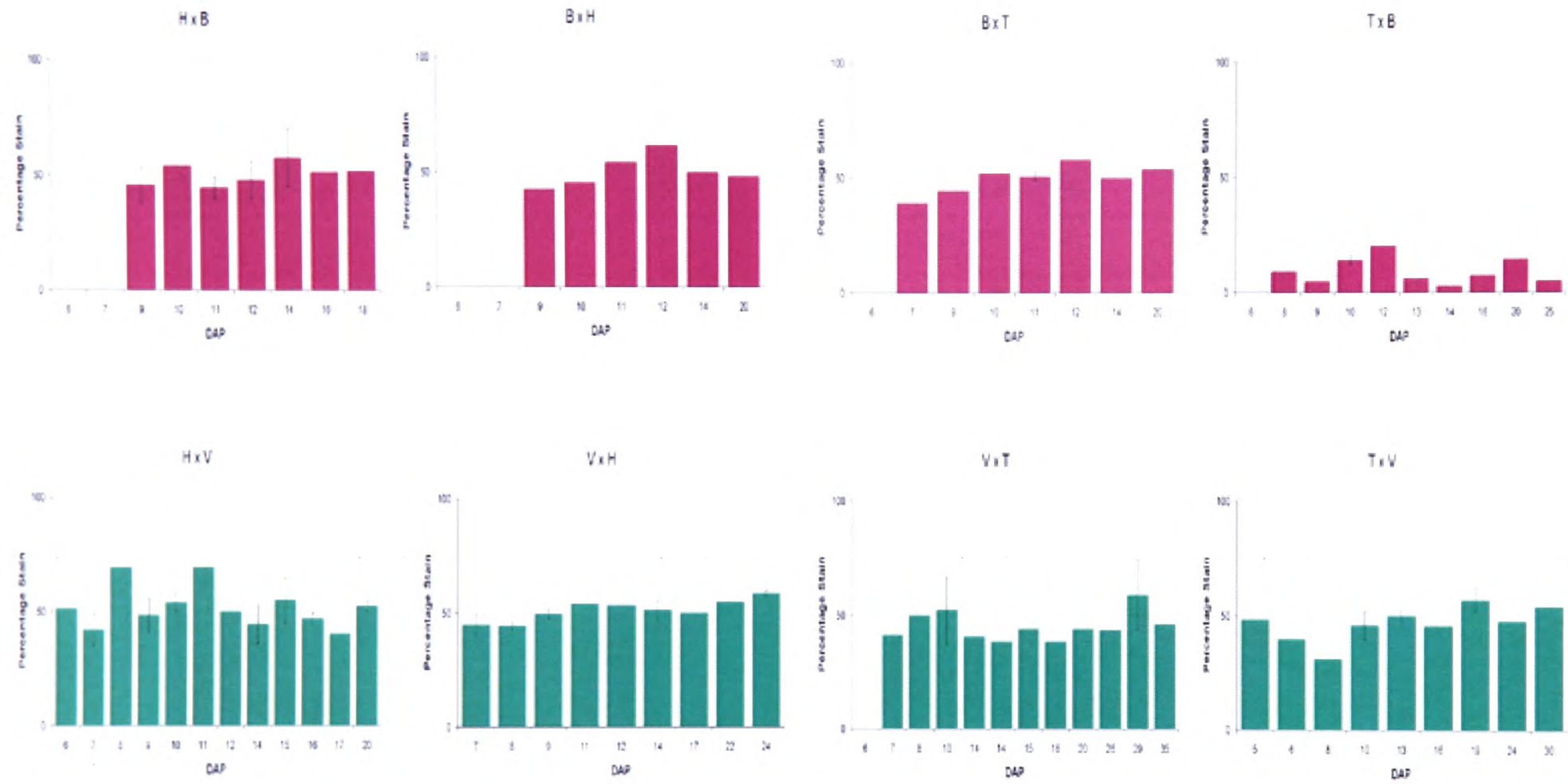


Figure 4-2 Frequency of kernels that stain for GUS

Graphs show the percentage of kernels which stain for GUS of those isolated at the time point displayed on the X-axis. Although sample sizes vary between different crosses and time-points, depending on the sample sizes never represented fewer than 30 kernels.

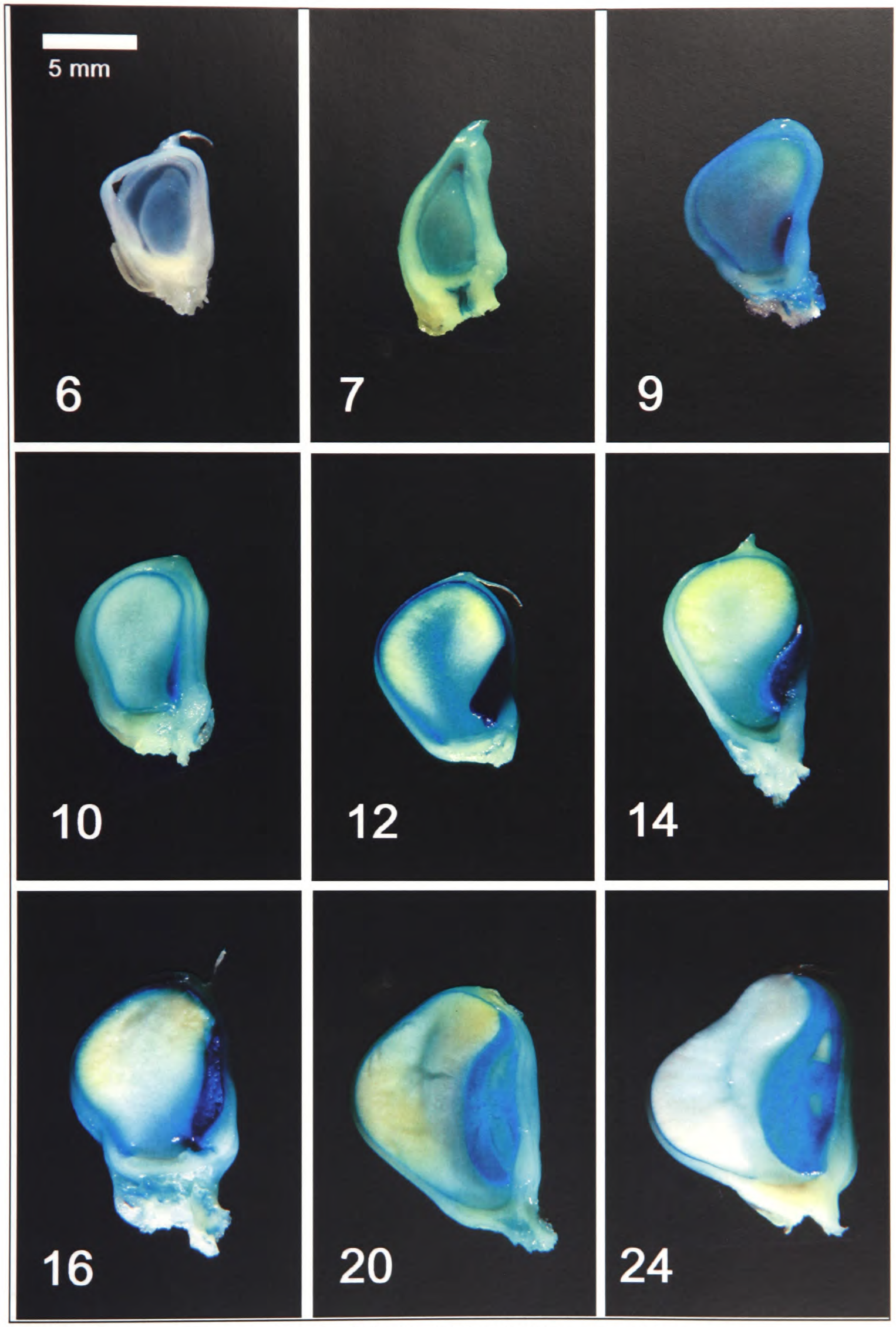


Figure 4-3 Progression of GUS activity in pVp1::GUS line x H99-2x crosses

Kernels at increasing DAP from within ploidy crosses, showing GUS activity for the aleurone marker pVp1::GUS reporter line when introduced as a female parent in a cross with the H99-2x diploid line. All images are to the same scale, and the scale bar represents 5mm. Numbers represent the number of DAP at which kernels were harvested and stained for GUS activity.

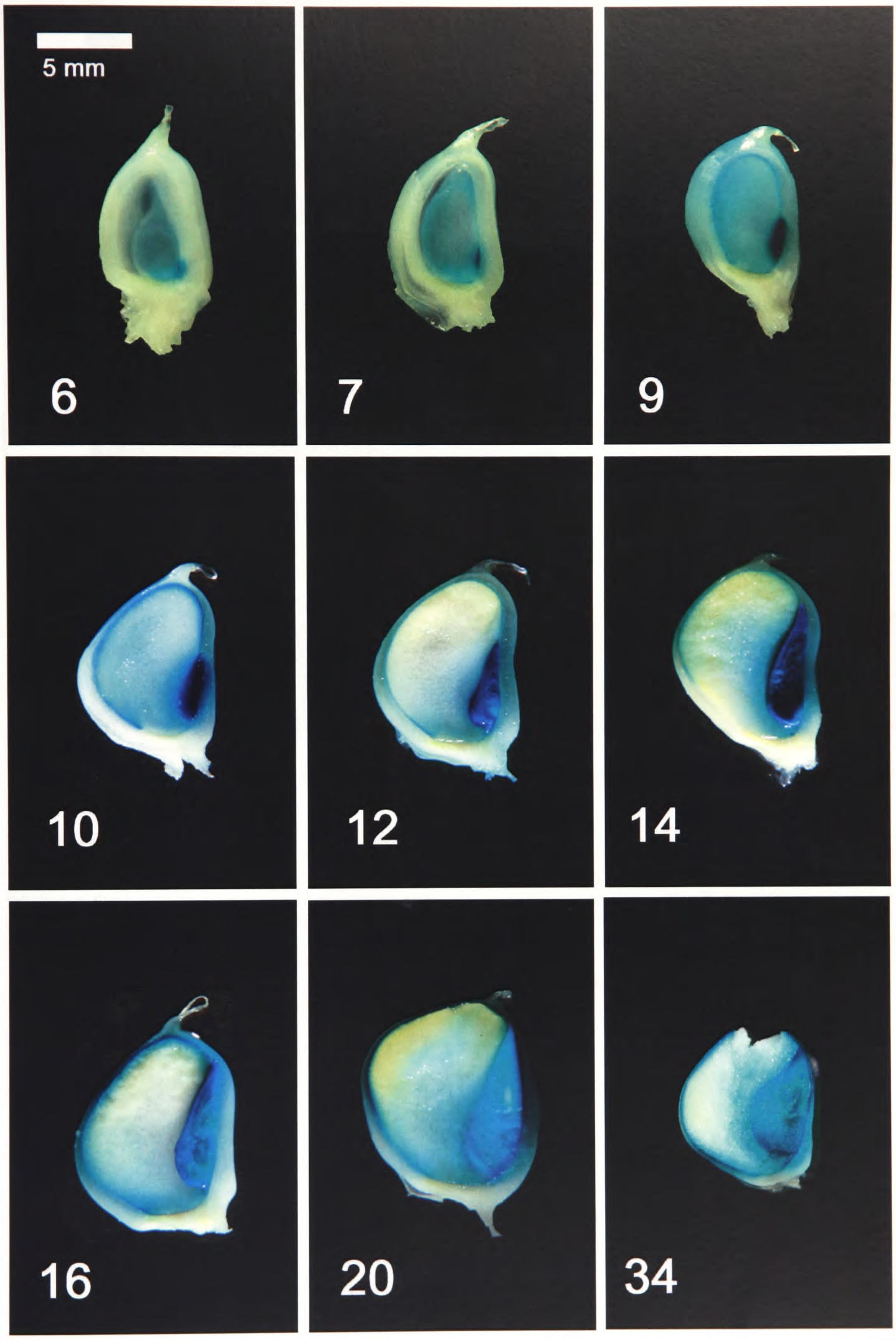


Figure 4-4 Progression of GUS activity in H99-2x x pVp1::GUS line crosses

Kernels at increasing DAP from within ploidy crosses, showing GUS activity for the aleurone marker pVp1::GUS reporter line when introduced as a male parent in a cross with the H99-2x diploid line. All images are to the same scale, and the scale bar represents 5mm. Numbers represent the number of DAP at which kernels were harvested and stained for GUS activity.

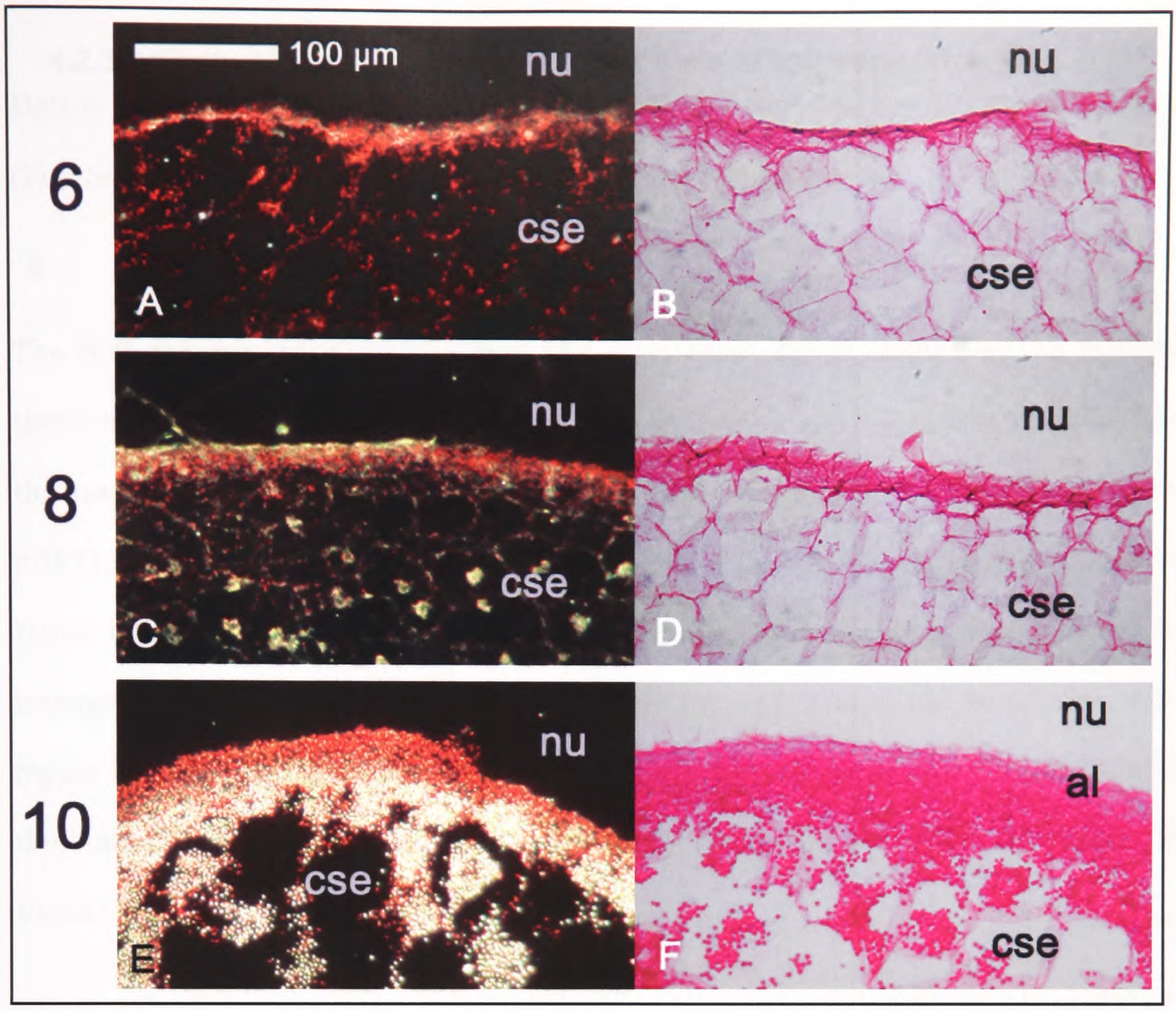


Figure 4-5 Comparison of aleurone-specific promoter activity and adoption of aleurone identity.

Images are from balanced crosses in which the pVp1::GUS reporter line is crossed with a H99-2x pollen parent and show the peripheral region of the endosperm from sagittal sections, just adgerminal to the crown, . **A**,**C**,**E** = dark field microscopy (DF), **B**,**D**,**F** = light microscopy following PAS staining for total carbohydrate. Days after pollination are labelled in the left hand margin. GUS precipitate appears red in DF images, revealing diffusion into underlying cells at all time points, but greatest at 6 and 8 DAP. At 6 DAP peripheral cells are identical to underlying cells(**B**), although they clearly express GUS. At 8 DAP peripheral cells have more defined cell walls, and appear to be adopting a distinct morphology to underlying cells, although not typically aleurone (**D**). By 10 DAP peripheral cells are typically aleurone-like with a box like morphology forming a single peripheral layer, intensely staining for GUS (**F**). nu = nucellus, al = aleurone, cse = central starchy endosperm.

4.2.3 Activity of the pBETL-GUS reporter gene in balanced crosses

Data is presented for both reciprocal crosses between the pBETL::GUS line and H99-2x (Figure 4-6, pBETL::GUS line x H99-2x; Figure 4-7, H99-2x x pBETL::GUS line).

The first appearance of GUS is probably artefactual, being located in the maternal tissue near the base from 6 DAP (Figure 4-6), as a small spot consistently located near the suspensor of the embryo. It is present in both crosses when either H99-2x or the pBETL::GUS line is used as the maternal parent, and at a frequency of almost 100% rather than the expected 50% in the latter case, so cannot be due to expression of the transgenic GUS gene. The existence of an endogenous β-glucuronidase activity in this region has been reported previously (Muhitch, 1998), and the addition of methanol to the staining mix reduced, although not entirely eliminated the effect (Hueros et al., 1999a).

The first endosperm-localized appearance of GUS activity is at 9 DAP, in the cross with the transgenic line as the female parent (Figure 4-6). The first unambiguous stain in the reciprocal cross does not appear until 10 DAP. In both these cases, the stain appears dramatically as a wide patch at the base of the endosperm where it overlies the maternal pedicel. The frequency of kernels that stain is close to what would be expected on the basis of transgene segregation (Figure 4-2). As the endosperm matures and grows, GUS activity remains restricted to the basal region, although this domain progressively increases in width. The GUS signal persists until at least 20 DAP in this region, although its activity has decreased by this stage, peaking around 16-18 DAP.

There is no significant difference in the timing or localization of the pBETL::GUS reporter, irrespective of whether it was inherited maternally or paternally.

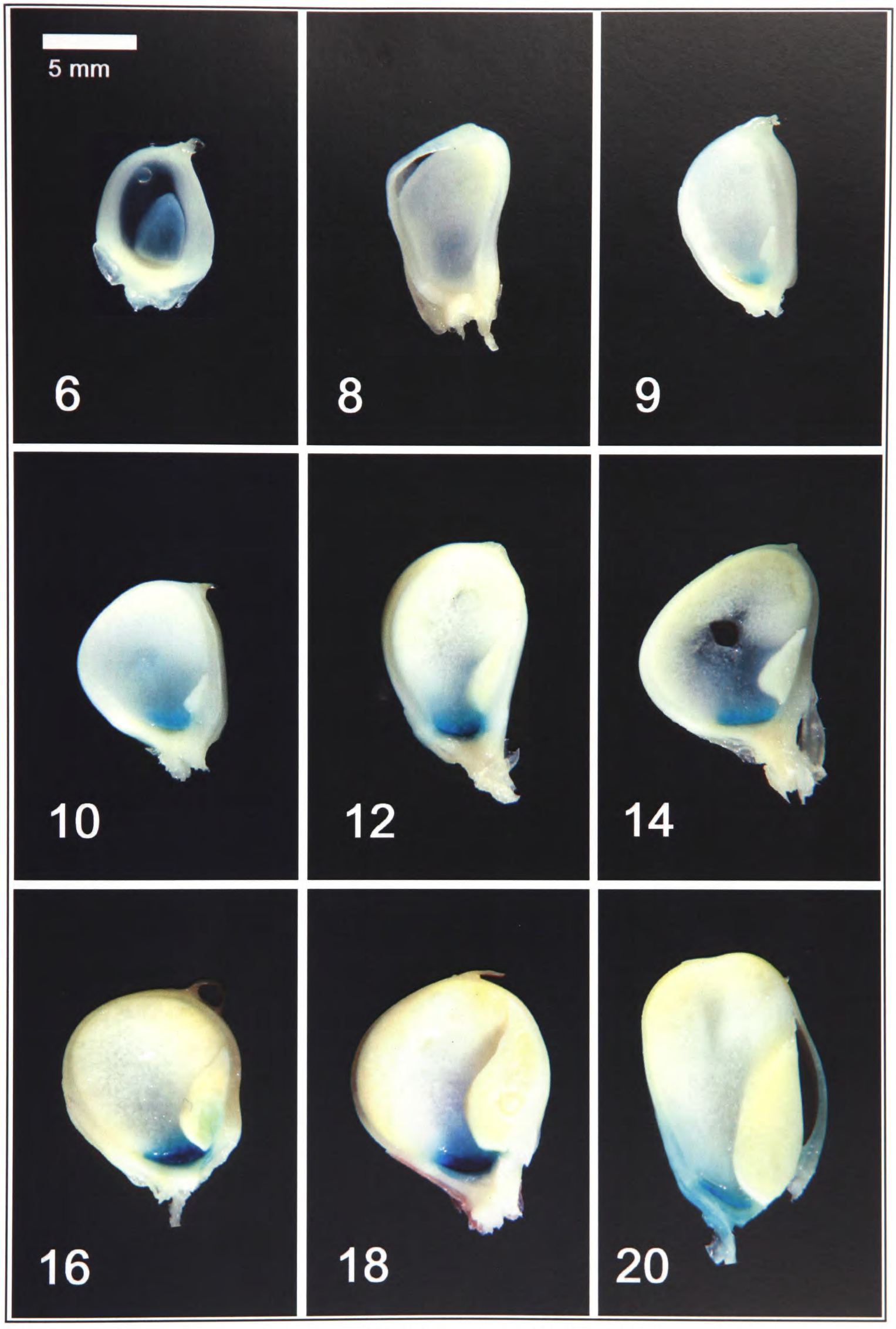


Figure 4-6 Progression of GUS activity in pBETL::GUS line x H99-2x crosses

Kernels at increasing DAP from within ploidy crosses, showing GUS activity for the transfer layer marker pBETL1::GUS reporter line when introduced as a female parent in a cross with the H99-2x diploid line. All images are to the same scale, and the scale bar represents 5mm. Numbers represent the number of DAP at which kernels were harvested and stained for GUS activity.

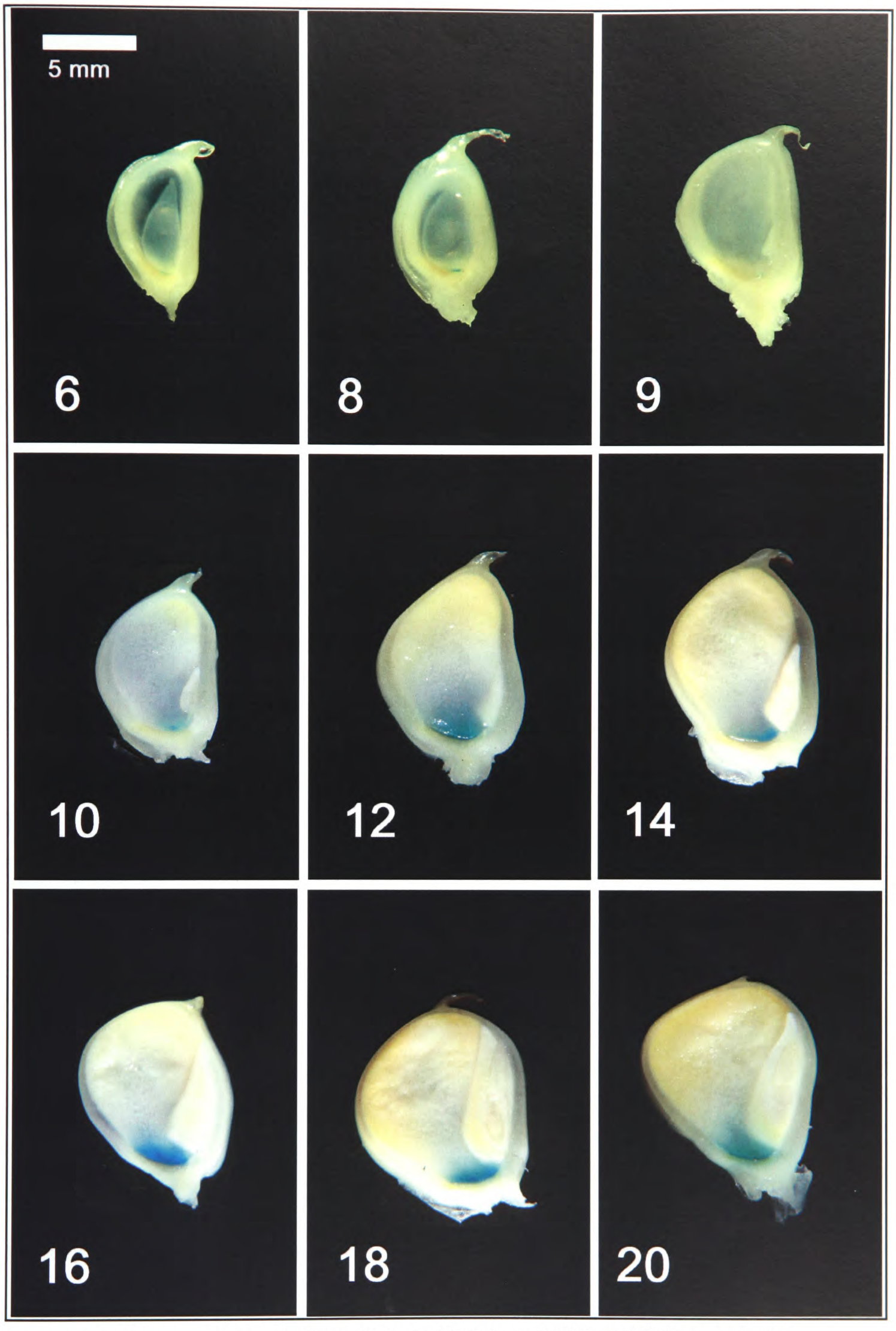


Figure 4-7 Progression of GUS activity in H99-2x x pBETL::GUS line crosses

Kernels at increasing DAP from within ploidy crosses, showing GUS activity for the transfer layer marker pBETL1::GUS reporter line when introduced as a male parent in a cross with the H99-2x diploid line. All images are to the same scale, and the scale bar represents 5mm. Numbers represent the number of DAP at which kernels were harvested and stained for GUS activity.

4.2.4 BETL cell morphology precedes pBETL1::GUS activity in balanced crosses

In order to investigate whether the first signs of GUS expression were concomitant with the appearance of typical BETL cell morphology, samples were fixed, sectioned and stained with PAS, which clearly reveals the cell wall invaginations in GUS reporting cells. These samples suggest that typical BETL morphology (i.e. feather-like striations perpendicular to the endosperm/pedicel margin, representing cell wall ingrowths) is first evident at 6 DAP, when interspersed cells display this morphology (Figure 4-8A). By 7 DAP, all cells along the basal layer display this morphology (Figure 4-8B). Apparently, differentiation of this layer occurs slightly earlier in crosses in which H99-2x is used as the female parent (Figure 4-8C), as at 7 DAP, BETL cells are more intensely staining, and there are more cell layers (shown by red arrow).

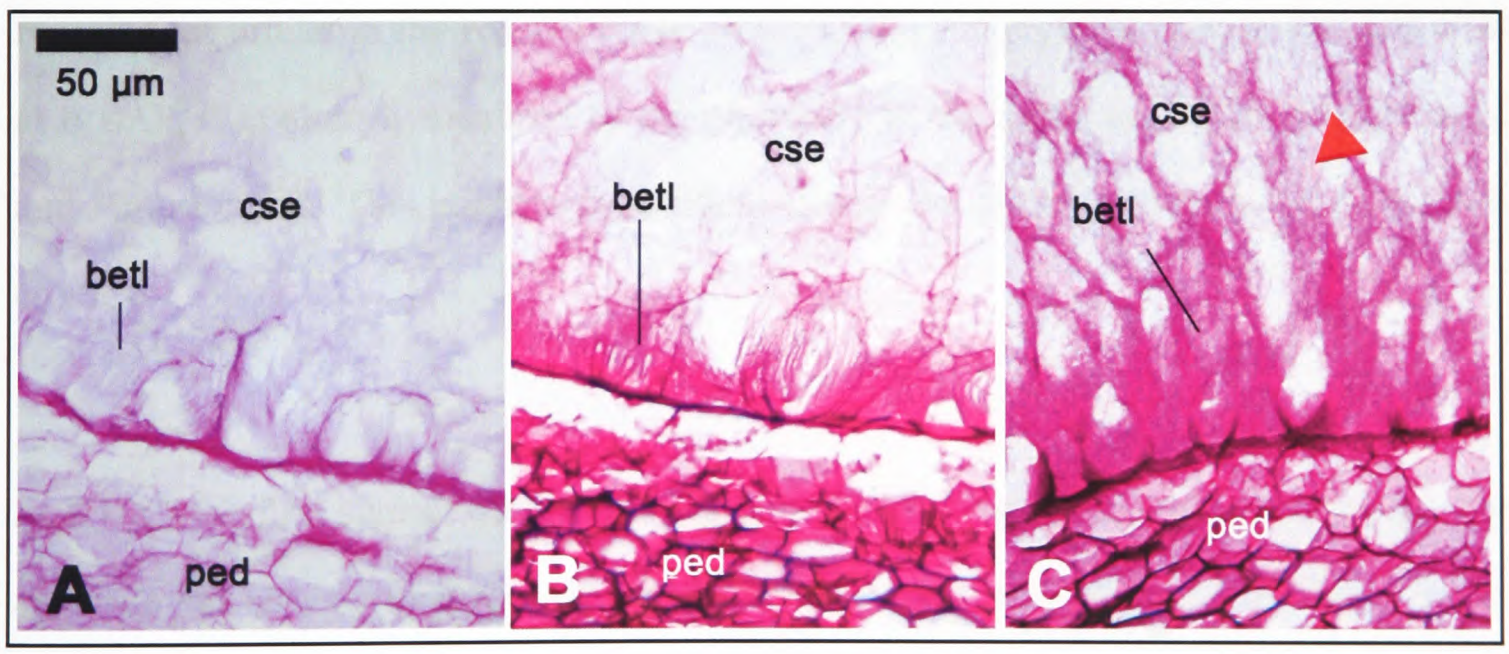


Figure 4-8 Initial appearance of transfer cell morphology in basal cells

Samples were stained with PAS, and imaged using bright-field microscopy. Samples from crosses between a female pBETL::GUS plant and a male H99-2x plant at 6 DAP (**A**), and 7 DAP (**B**). Samples from H99-2x x pBETL::GUS plants are a little more advanced: (**C**) is taken at 7 DAP. cse= central starchy endosperm, betl = BETL cell, ped= pedicel.

4.2.5 Activity of the pVp1::GUS reporter gene in maternal excess crosses.

Reporter lines carrying the pVP1-GUS transgene were used to pollinate plants of the tetraploid line, HybT, and the pattern of reporting assessed between 4 and 26 DAP (Figure 4-9).

The first appearance of staining is at 4 DAP (Figure 4-9, 4 DAP), and is exclusively localized to the embryo, which is one day sooner than in the control crosses. Between 6 and 8 DAP there is some variability in the extent of GUS activity – some samples have only a small patch on the germinal margin, whereas some have a more extensive, yet still incomplete pattern of staining, with GUS activity located to varying degrees in the epidermis near the germinal corner, or in the crown, and others show complete peripheral GUS activity. Together these observations suggest that, in this direction of interploidy cross, the time to complete peripheral GUS activity is between 24 and 48 hours, and this activity probably starts at the germinal corner, or the apex, with these domains expanding along the flanks of the endosperm. Histological investigation reveals that unlike in the control cross, where GUS product leaks into underlying cells at 6 DAP, it is sharply localized to the periphery in the maternal excess (Figure 4-25 and Figure 4-27). This might reflect differences in the cell structures between the two crosses.

At 8 DAP the first sign of deviation from normal development is already apparent, the aleurone has already formed a clearly defined cell layer (Figure 4-25 and Figure 4-27). An even more profound difference is also apparent - GUS activity is present in the region which normally develops into the BETL layer. In the control endosperm GUS activity is present along the entire periphery except for the region just above the maternal pedicel, where the BETL layer is located. However, in maternal excess crosses pVp1::GUS activity suggests that this region which should be the BETL has developed

into an extension of the aleurone. The same phenomenon was also reported when the H99-4x plant was used as the female parent (Figure 4-10). This 'ectopic aleurone' layer shows variability in the extent to which it invades the region above the pedicel, as some endosperms show a non-staining disc-like region of variable size within the ectopic aleurone (Figure 1-1). Histological observation of the cells in this region suggested that cells lacking GUS are mostly not BETL-like in appearance, but have an intermediate appearance between aleurone and transfer layer cells. However, a limited number of interspersed cells do show the typical striations of BETL cells, although these never extended more than one cell deep (Figure 4-12).

The exclusive localization of GUS activity in the endosperm periphery, including activity in the basal region continues through development until 20 days after pollination, mimicking the peripheral and embryo localization in the balanced crosses, even though the endosperm is morphologically quite different. By 26 days after pollination, GUS activity remains located in the cells mentioned, but by this stage it is much fainter, and by 30 DAP only the embryo continues to stain (Figure 4-9, 26 and 30 DAP).

Samples taken at 30 DAP were sectioned, and stained with toluidine blue O, which reveals details of cell structure (Figure 4-27). It is evident that although a single layer of peripheral cells is found, they are histologically different from the aleurone cells of the controls. Instead of being cubodial, with the long axis at a right-angle to the endosperm edge, they are flatter, and more rounded in profile, having curved corners.

The frequency of staining for GUS throughout all stages is close to the expected proportion of 50% (Figure 4-2).

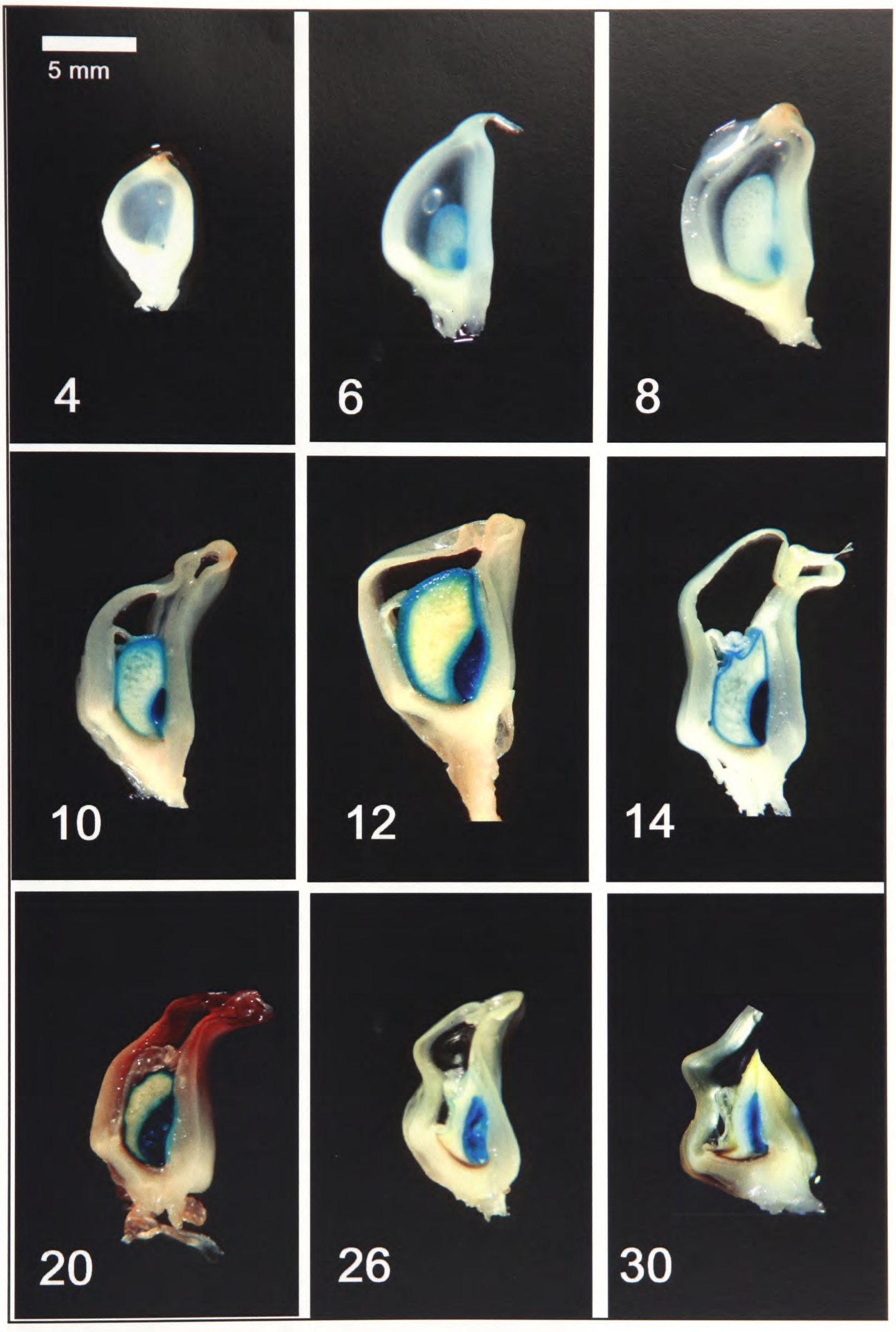


Figure 4-9 Progression of GUS activity in HybT x pVp1::GUS line crosses.

Kernels at increasing DAP from maternal excess interploidy crosses, showing GUS activity for the transfer layer marker pVp1::GUS reporter line when introduced as a male parent in crosses with the HybT tetraploid line. All images are to the same scale, and the scale bar represents 5mm. Numbers represent the number of DAP at which kernels were harvested and stained for GUS activity.

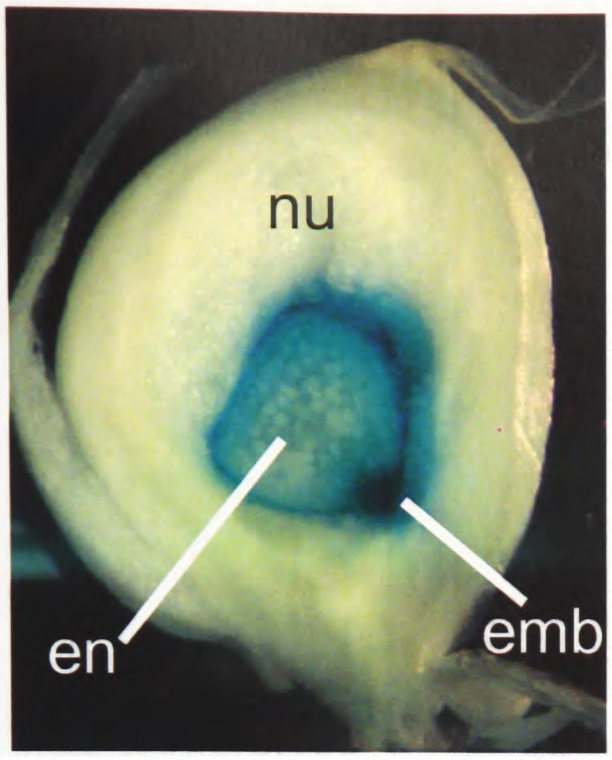


Figure 4-10 Maternal excess Vp1 reporting pattern in a H99-4x maternal genotype

Hand cut kernel, showing endosperm (en) and embryo (emb) within the nucellus (nu) of a kernel. The maternal excess endosperm stains for the aleurone and embryo (both blue), displaying the similarity of phenotype as for a cross with a HybT maternal parent. As the sample was harvested at 10 DAP, it clearly demonstrates the delay in growth of the kernel in this sample, which by this stage would normally occupy the entire region of the nucellus in a balanced cross.

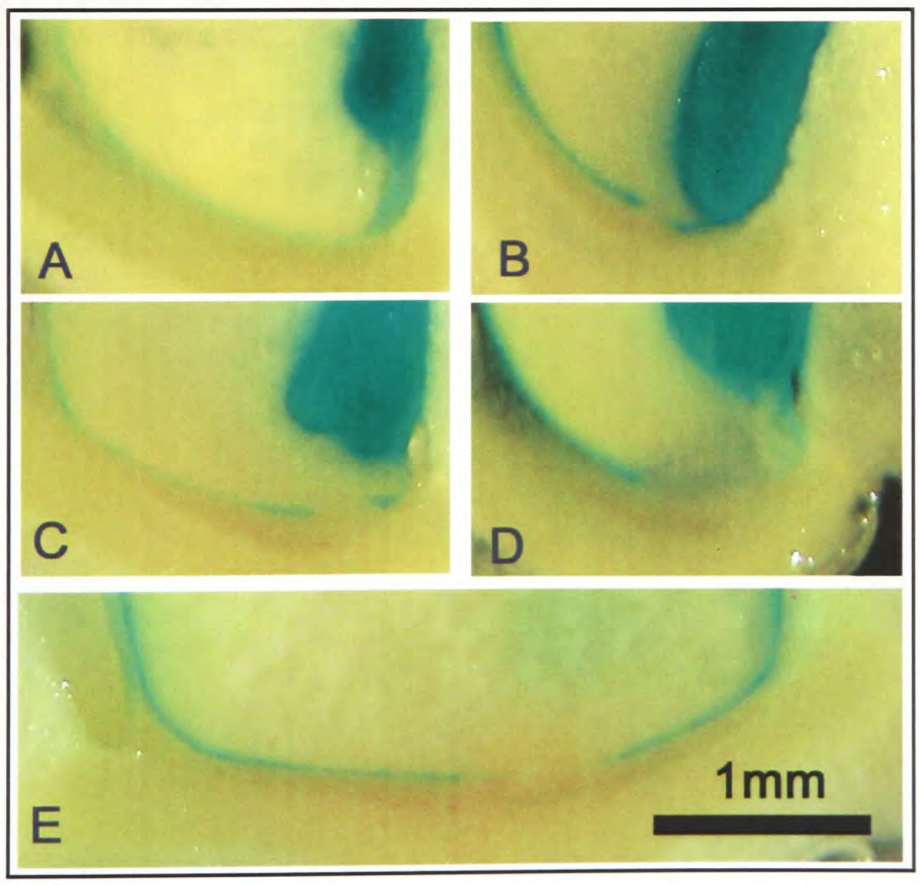


Figure 4-11 Gaps of varying size within the ectopic basal 'aleurone'

A-D)Sagittal hand-cut sections of 13 DAP maternal excess kernels, stained overnight with GUS to assess activity of the pVp1::GUS transgene carried through the male parent. **E**) Transverse section hand cut section from same cross, sectioned just behind the embryo.

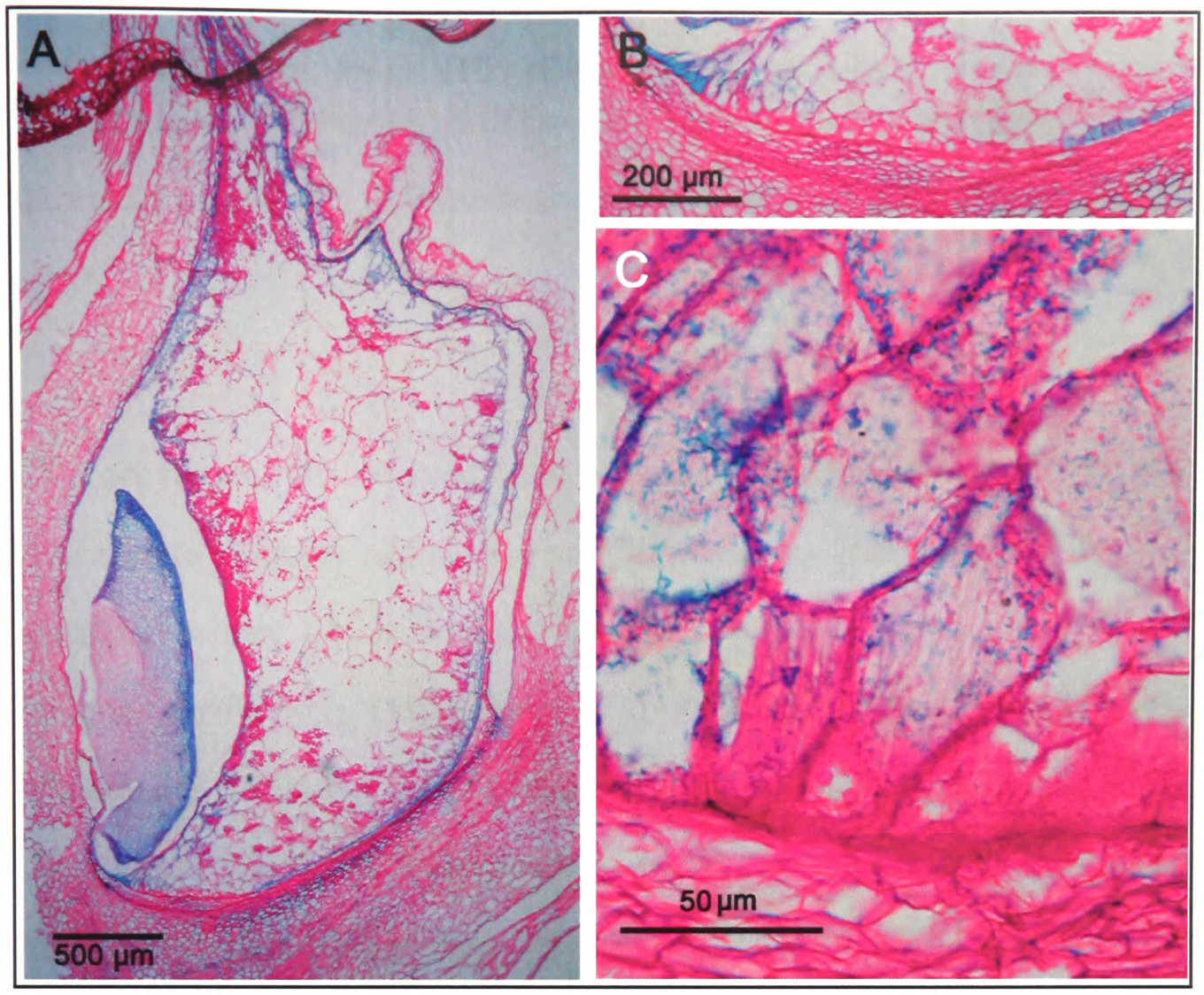


Figure 4-12 PAS staining of non-GUS staining gaps in basal 'aleurone'

10 DAP Maternal excess kernels which had a gap in the blue GUS staining were sectioned and stained with PAS for total carbohydrate (**A**). Most cells in this area show neither clearly BETL nor aleurone morphology (**B**). A limited numbers of cells in the area basal lacking GUS staining display correct BETL morphology, although usually such cells are isolated, and only one cell deep (**C**).

4.2.6 Activity of the pBETL1::GUS reporter gene in maternal excess crosses

Reporter lines carrying the pBETL::GUS transgene were crossed used to pollinate HybT tetraploid plants, and reporter activity was monitored between 6 and 25 DAP (Figure 4-13).

GUS activity is first observed from the normal time-frame of 10 DAP, however, it is restricted to a smaller region than normal, forming a domain consistently located at the germinal corner, which occupies around 1/3rd of the region that would normally become transfer layer (Figure 4-13, 10 DAP). Significantly, GUS activity was observed in a smaller fraction of kernels than would be expected on the basis of segregation of the transgene (between 5 and around 15%, Figure 4-2). Furthermore the strength of the GUS signal is weaker than normal, being only a faint blue, so it was much more difficult to observe than normal.

Between 12 and 30 DAP a very similar pattern of GUS activity is observed with a transfer layer found at the basal region, but only in a restricted domain consistently located at the germinal corner, and also at a lower frequency than expected of transgene segregation, the majority of kernels lacking any GUS staining at all (Figure 4-2).

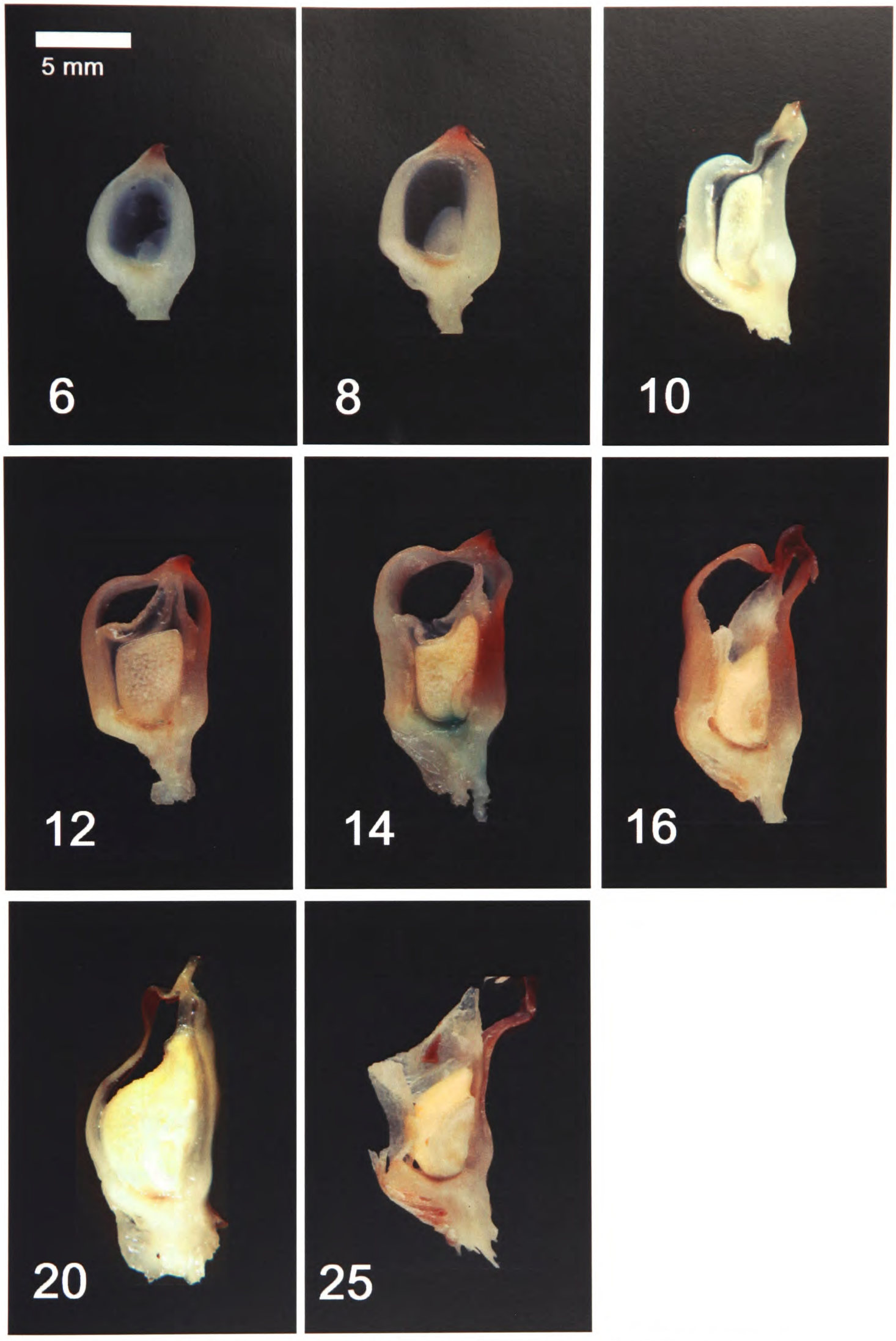


Figure 4-13 Progression of GUS activity in HybT x pBETL::GUS line crosses.

Kernels at increasing DAP from maternal excess interploidy crosses, showing GUS activity for the transfer layer marker pBETL1::GUS reporter line when introduced as a male parent in crosses with the HybT tetraploid line. All images are to the same scale, and the scale bar represents 5mm. Numbers represent the number of DAP at which kernels were harvested and stained for GUS activity.

4.2.7 Activity of the pVp1::GUS reporter gene in paternal excess crosses.

Diploid reporter lines plants carrying the pVP1::GUS transgene were crossed with pollen from HybT plants, and staining patterns recorded between 4 and 30 DAP, as presented in Figure 4-14.

To observe aleurone development in the case of paternal genomic excess, diploid plants carrying the reporter gene were crossed with the two tetraploid lines. The cross with H99-4x was investigated at 10 DAP, and with the HybT line plant between 4 and 30 DAP. A summary of the development of kernels from crosses between the pBETL::GUS line and HybT are summarized in Figure 4-14.

Unlike in maternal excess endosperms, at 4 DAP there is no sign of GUS activity in the embryo.

Between 7 and 10 DAP GUS activity is seen throughout the endosperm, but is greatest on the germinal side and near the crown. GUS activity was not seen in the centre of the endosperm in the control or the maternal excess crosses. Observation of kernels stained for only ½ hour displayed GUS activity exclusively located in the periphery, but after further incubation in the stain, it was located throughout the endosperm, indicating a rapid diffusion of the coloured substrate through the kernel (Figure 4-15). Since an overnight stain never showed central endosperm GUS staining in the control cross, it seems that this rapid diffusion is due to a qualitative difference between the centrally located cells of the paternal excess endosperm and those from balanced crosses. This diffusion of the GUS product was quite apparent through sectioning of the samples combined with dark field microscopy (Figure 4-26), which only starts to become less severe by 14 DAP.

At 14 DAP, GUS activity is more tightly restricted to the peripheral domain, where the aleurone would normally lie. Also at this time-point the embryo, which is normally conspicuous and stains intensely for GUS, is not visible at all in hand sections, suggesting that its development has been perturbed in this cross.

By 20 DAP the endosperm has lost most of its translucency, and apparent diffusion of GUS precipitate seen at earlier time-points has ceased (Figure 4-14, 20 DAP). It is interesting to note that many of the kernels from this cross show an irregular outer-surface, which instead of being smooth and almost spherical, is convoluted, often displaying deep cavities into the adgerminal face. The outer margin of the aleurone is therefore highly irregular and variable between kernels, and there are numerous surface invaginations creating 'ectopic surfaces' deep within the endosperm. In spite of their unusual location, these surfaces also stain for GUS activity, suggesting they are also aleurone-like cells (Figure 4-16).

Strikingly by 30 DAP, not only does the periphery of the endosperm stain but also random locations throughout the tissue (Figure 4-14, 30 DAP). The aberrant staining domains are also of varying sizes suggesting that the events which induced GUS activity in these clusters may not have occurred at the same time.

At every developmental stage investigated, a proportion of endosperms close to the expected 50% stained for GUS (Figure 4-2).

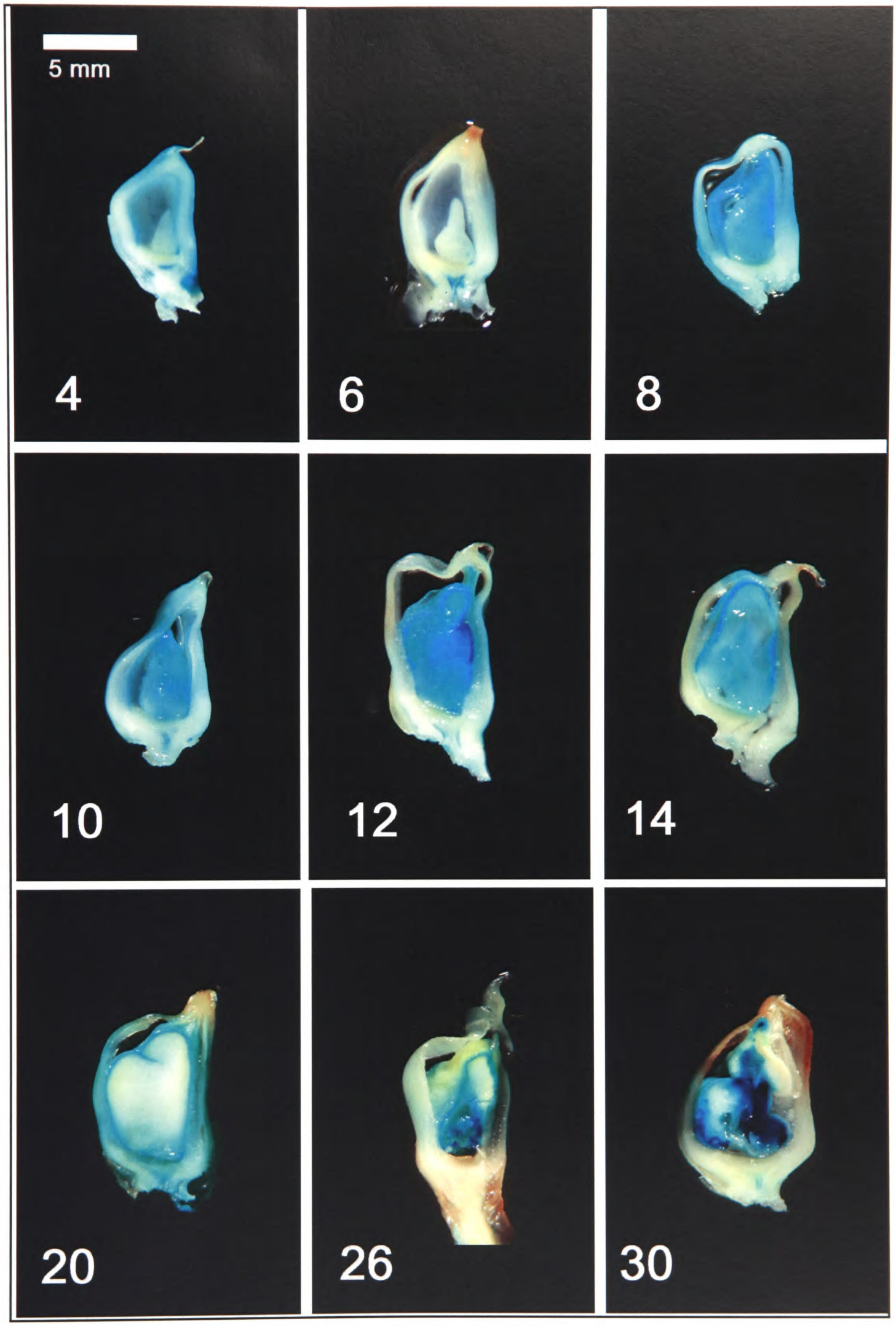


Figure 4-14 Progression of GUS activity in pVp1::GUS line x HybT crosses.

Kernels at increasing DAP from maternal excess interploidy crosses, showing GUS activity for the transfer layer marker pVp1::GUS reporter line when introduced as a female parent in crosses with the HybT tetraploid line. All images are to the same scale, and the scale bar represents 5mm. Numbers represent the number of DAP at which kernels were harvested and stained for GUS activity.

4.2.8 Aleurone morphology is severely disrupted in paternal excess crosses

30 DAP samples were sectioned and stained with toluidine blue (Figure 4 27) to observe aleurone morphology. Although an aleurone layer is reported by the GUS reporter, it is quite clearly aberrant. Unlike normal aleurones, which comprise only a single cell layer, the paternal excess aleurone has an uneven layer comprising between one and three cells. These cells are irregularly shaped, and of variable size.

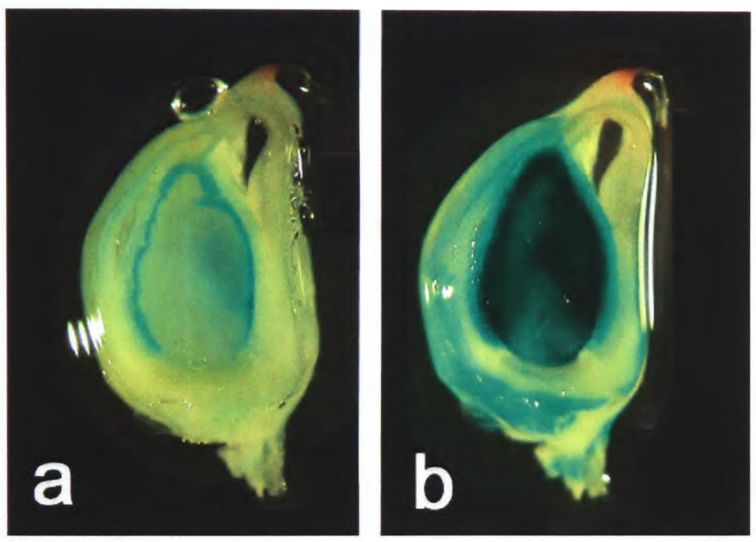


Figure 4-15 Diffusion of GUS precipitate into underlying cells in paternal excess crosses

One paternal excess kernel was stained for GUS activity at 14 DAP and photographed after three hours (**a**): staining is located in the outer margin. It was then left to stain overnight (**b**): after this period the GUS was intense and located throughout the endosperm.

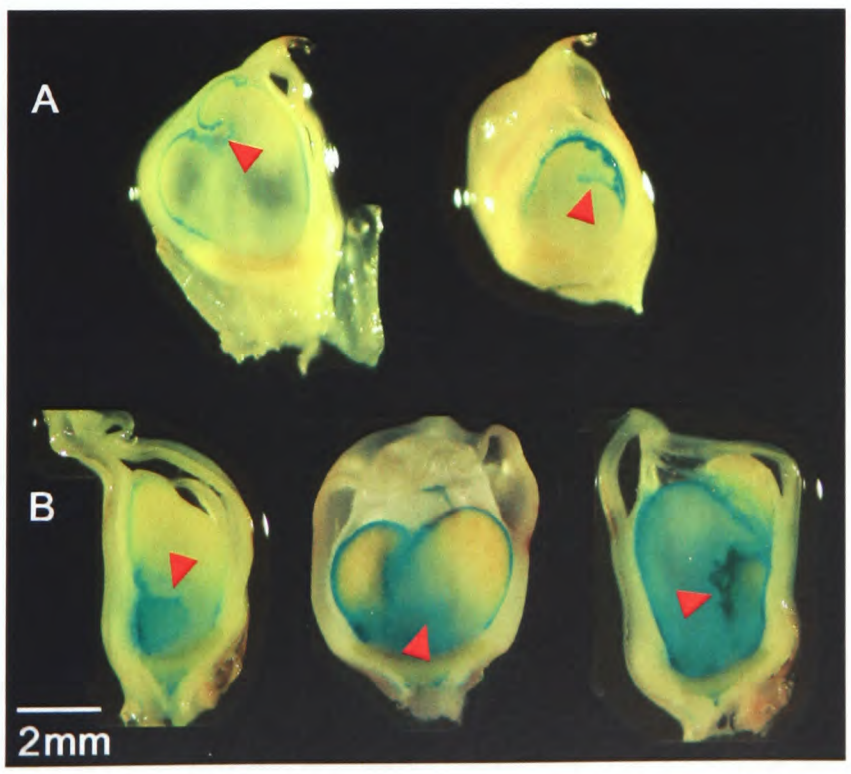


Figure 4-16 Vp1::GUS activity on ectopic surfaces

Kernels from paternal excess crosses in which the diploid paternal parent carried the pVp1::GUS transgene were stained for GUS activity. Samples were harvested at 14 DAP (**a**), or 20 DAP (**b**), in spite of the endosperm edges being located ectopically (red arrows), they stain for GUS, suggesting they are 'aleurone' cells.

4.2.9 Activity of the pBETL1::GUS reporter gene in paternal excess crosses.

Crosses between the transgenic diploid pBETL::GUS line, used as a female parent, and the tetraploid HybT as pollen donor were analysed at a range of time points as shown in Figure 4-17. The tetraploid H99-4x was also used as a pollen parent for samples harvested at 10 and 15 DAP.

The first sign of GUS activity was at 10 DAP, as in the control. However, the basal region where contact is made between the endosperm and the pedicel lacks the complete layer of GUS activity seen in the controls. Instead kernels usually contain multiple small spots of GUS activity, of the order of one to five. The location of these spots are sporadically distributed across the area which would normally form the BETL region. Surprisingly, there is a small difference between the two different tetraploids used. When H99-4x is used as a male parent, in addition to the appearance of small clusters of cells along the would-be BETL layer, a small percentage of kernels also display a small focus of blue–staining cells at the apex of the endosperm (Figure 4-18). This ectopic localization of GUS activity is not observed in the crosses with HybT plants at this time-point. Sectioned samples were stained with PAS to determine whether either the spots located in the base, or those located ectopically, appear like normal transfer-like cells with typical striations perpendicular to the base. Although the cells located at the base were morphologically different to the centrally located cells, and stained more intensely, suggesting a more extensive carbohydrate composition, they showed no evidence of the typical pattern of the BETL striations at this stage. Ectopically located cells showed no features suggestive of BETL identity (Figure 4-18).

By 12 DAP an abnormal basal layer is still found in paternal excess endosperms, again showing punctuate staining randomly dispersed along the base where the BETL should be. However, the size of these clusters of GUS expressing cells domain is larger.

Between 12 and 18 DAP there is a range of phenotypes of endosperm, from those which that have only a few small patches of GUS activity along the base, via those that have very wide patches, to those which show an almost complete band of GUS activity as in the control crosses. Generally, the number of spots tends to increase with the number of DAP, and they are of variable size, as shown in (Figure 4-22). However, it seems likely that high rates of diffusion through the cells of the base, which remain translucent and presumably lack mature cell walls, accounts for the broader staining patterns. PAS-stained sections at 14 DAP suggested that whilst most cells stained intensely, but lacking typical BETL cell wall projections, a few cells did show some evidence of this modification, although the length of these projections were shorter than normal, suggesting the cells had not formed properly or were positioned obliquely Figure 4-23. Most intriguingly, around these timepoints the first evidence of ectopic staining in locations randomly located through the endosperm is apparent in crosses with HybT as the male parent (similar to what occurs at an earlier time than in crosses with H99-4x) (Figure 4-20).

By 18 DAP some kernels show strong staining around the base, and appear similar to wild type; however closer observation shows that the blue stain is patchy and smaller domains within those larger staining regions can be seen also evident (Figure 4-21), moreover embryos are frequently absent. By 24 DAP the endosperms stain only faintly in the bottom layer, although by now this stain occupies most of this region. Closer analysis shows that once again there are smaller intensely staining patches within broader staining domains (Figure 4-19) of cells intense patches of cells. Ectopically located spots are found in a larger proportion of kernels.

Throughout development a proportion of kernels stain for GUS which is close to the expected figure of 50% (Figure 4-2).

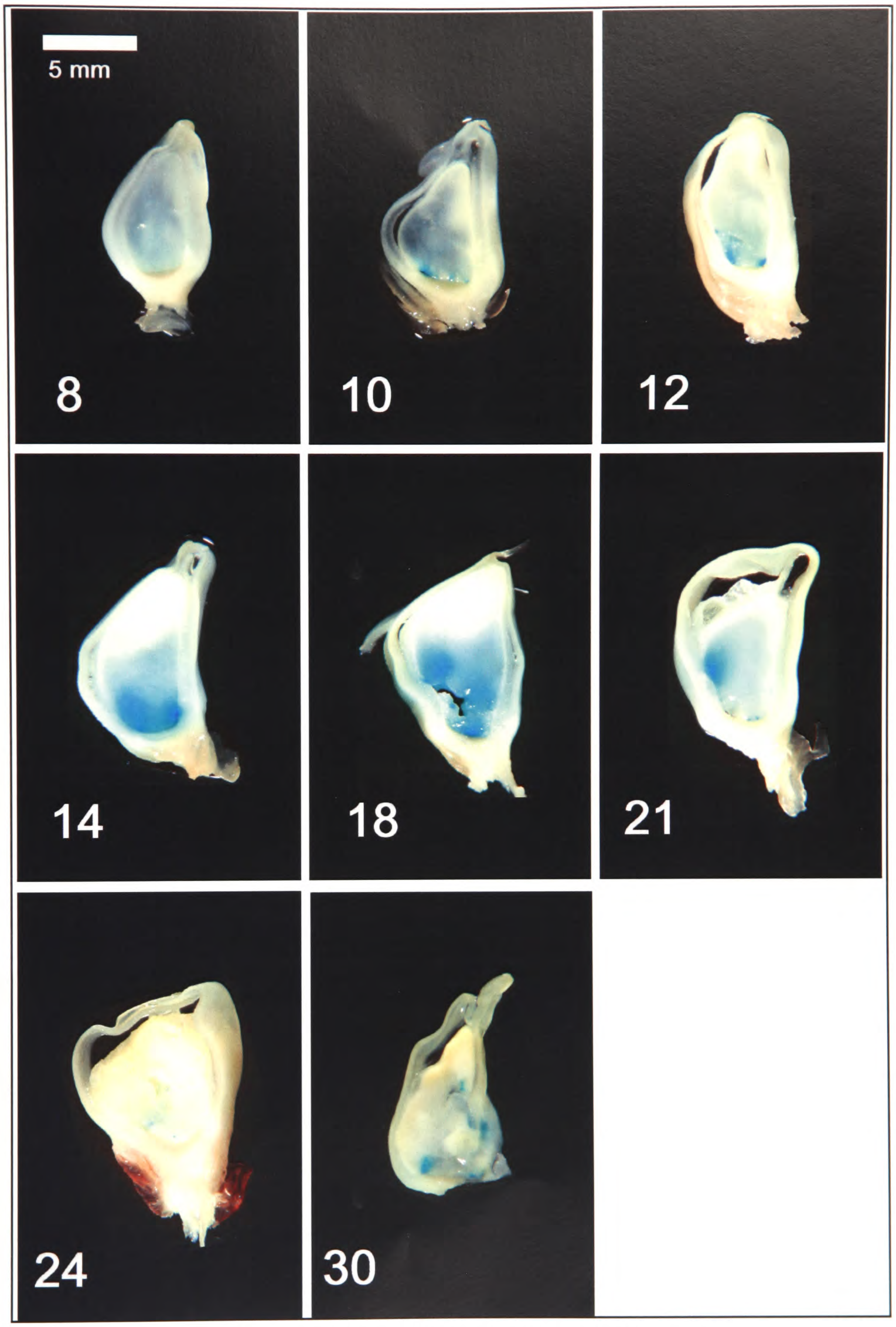


Figure 4-17 Progression of GUS activity in pBETL::GUS line x HybT crosses.

Kernels at increasing DAP from maternal excess interploidy crosses, showing GUS activity for the transfer layer marker pBETL1::GUS reporter line when introduced as a male parent in crosses with the HybT tetraploid line. All images are to the same scale, and the scale bar represents 5mm. Numbers represent the number of DAP at which kernels were harvested and stained for GUS activity.

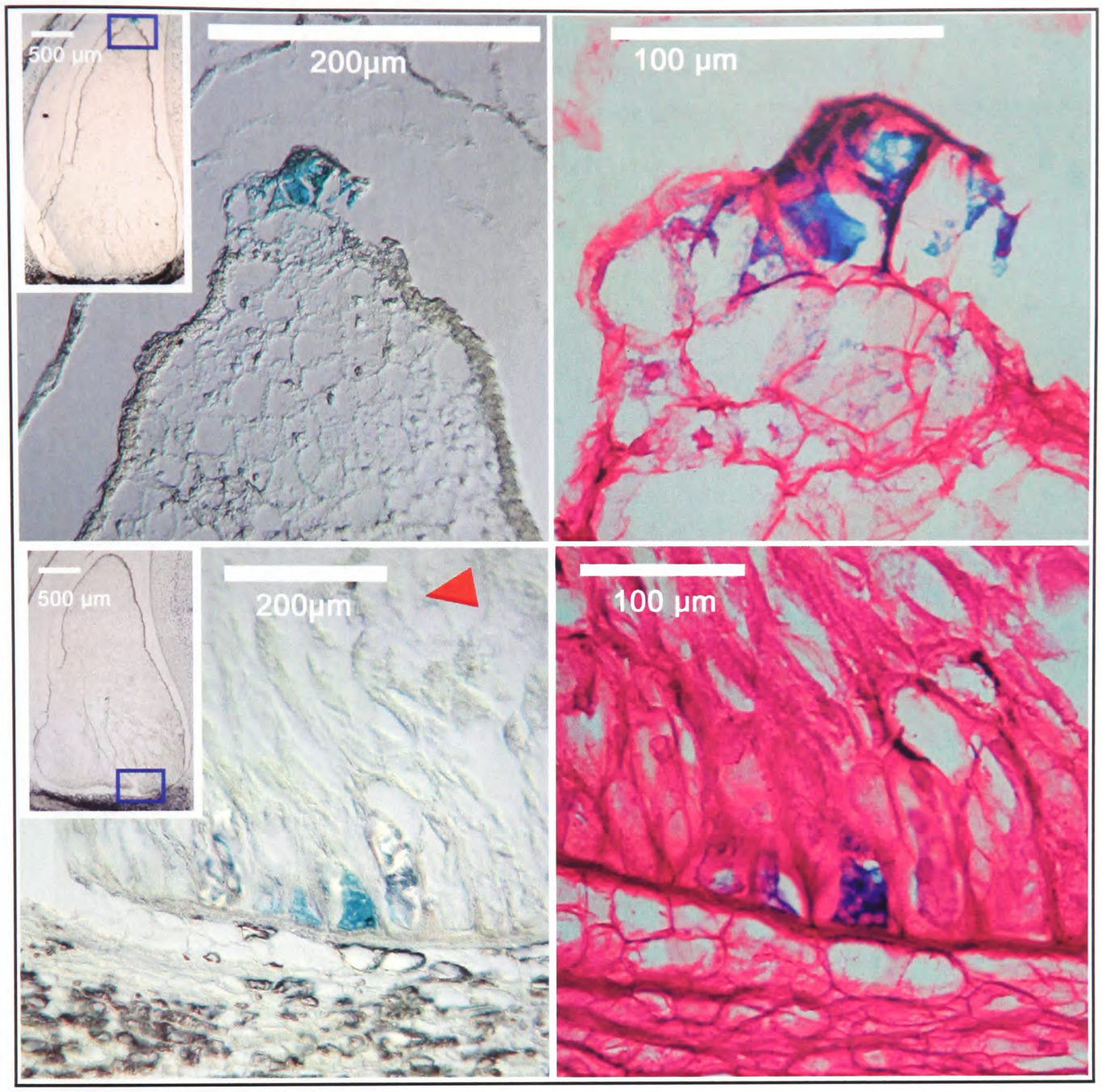


Figure 4-18 Apically located GUS activity in a pBETL::GUS x H99-4x cross

Samples stained for GUS had ectopic GUS activity in the apical portion of the endosperm at 10 DAP in a small percentage of kernels if H99-4x was the pollen parent. Staining of samples with PAS suggested that these cells did not have normal BETL morphology.

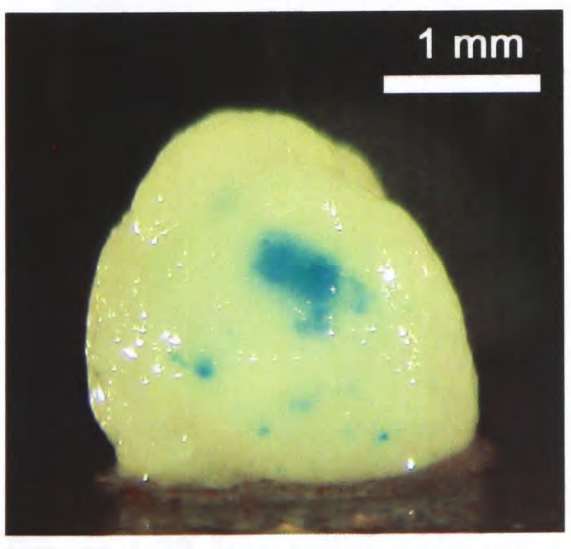


Figure 4-19 Patches of GUS activity on base of paternal excess endosperm

14 DAP endosperm were removed from the surrounding nucellus and pericarp by careful incisions, and stained for GUS activity. This kernel has been imaged from the base of the endosperm, which clearly demonstrates a range of sizes of GUS patches.

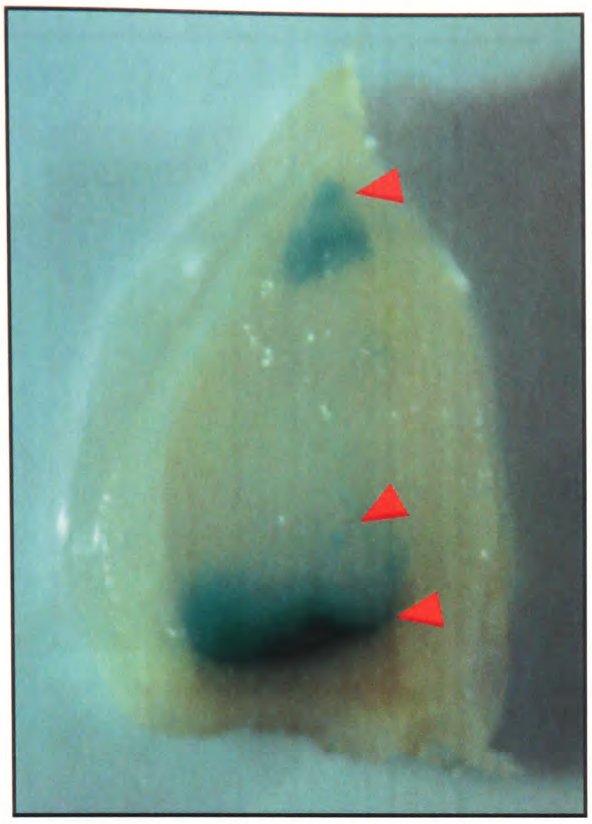


Figure 4-20 Ectopic localization of GUS in pBETL::GUS x HybT plants

A hand-cut paternal excess endosperm, carrying the pBETL::GUS reporter at 14 DAP, following GUS staining. A small proportion of kernels show ectopic localization of GUS activity, such as at the apex, in internally located cells and at the base as in this kernel, as indicated by the red arrows.

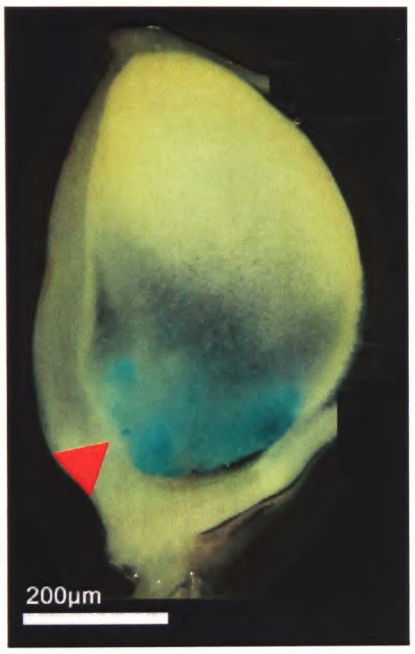


Figure 4-21 Kernel demonstrating patches within larger patches of GUS activity.

A hand-cut paternal excess kernel of 14 DAP, this kernel is relatively normal looking, showing a plump, well filled starchy endosperm, however, staining for pBETL::GUS activity still demonstrates a highly irregular BETL, showing blue patches of varying size, including smaller patches within larger one, suggesting different ages from when one progenitor cell started expressing GUS.

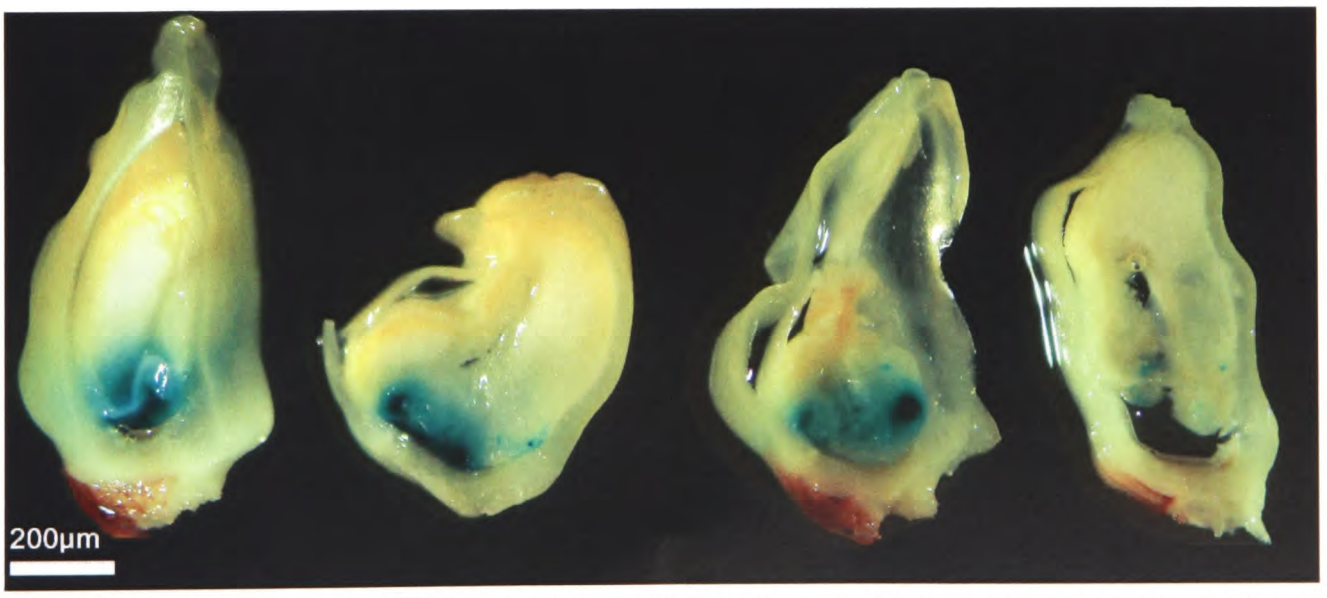


Figure 4-22 Variability between pBETL::GUS activity in paternal excess kernels

Several hand-cut paternal excess kernels taken at 20 DAP, demonstrate the range of morphology apparent in these kernels, and the variability in pBETL::GUS expression.

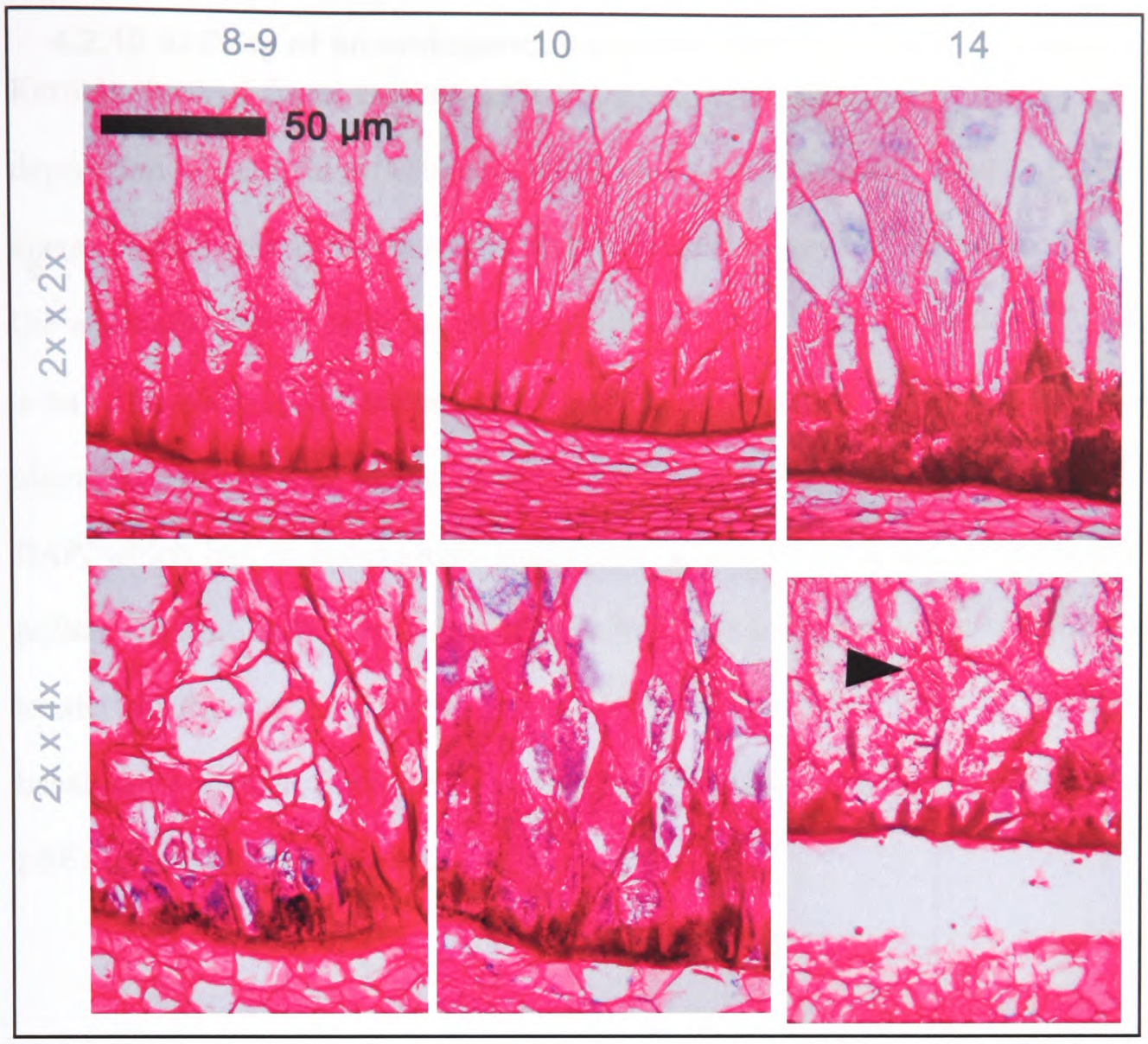


Figure 4-23 PAS staining of basally located cells

Control and paternal excess endosperm are displayed, at the number of DAP listed above the panels. Typical cell wall modifications are apparent in the diploid control at all stages, but are only apparent by 14 DAP in the paternal excess, and most cells still lack this modification. Whereas the striations are long and project at right angles to the endosperm base, in the paternal excess they are very short and at a skewed angle, suggesting the cells may be at oblique angles.

4.2.10 Activity of an endogenous gene in paternal excess crosses

Kernels derived from crosses with the pVp1::GUS line as the female parent showed deposition of an anthocyanin pigment in both wild type and interploidy crosses at late stages of endosperm development (from 16 DAP onwards). This allowed the analysis of the aberrant effects of these 'paternal excess' crosses for an endogenous gene (Figure 4-24). In normal diploid-self crosses, this pigment would be seen exclusively in the aleurone layer as seen in the 'escaped' kernel on the far right panel of Figure 4-24 at 18 DAP, which has spontaneously undergone a normal program of development, or is a pollen contaminant. In paternal excess endosperms, although still predominantly located in the aleurone layer, anthocyanin is found ectopically in the CSE and in the basal layer of the aleurone in a punctuate fashion, similar to that reported for the pBETL1::GUS reporter gene.

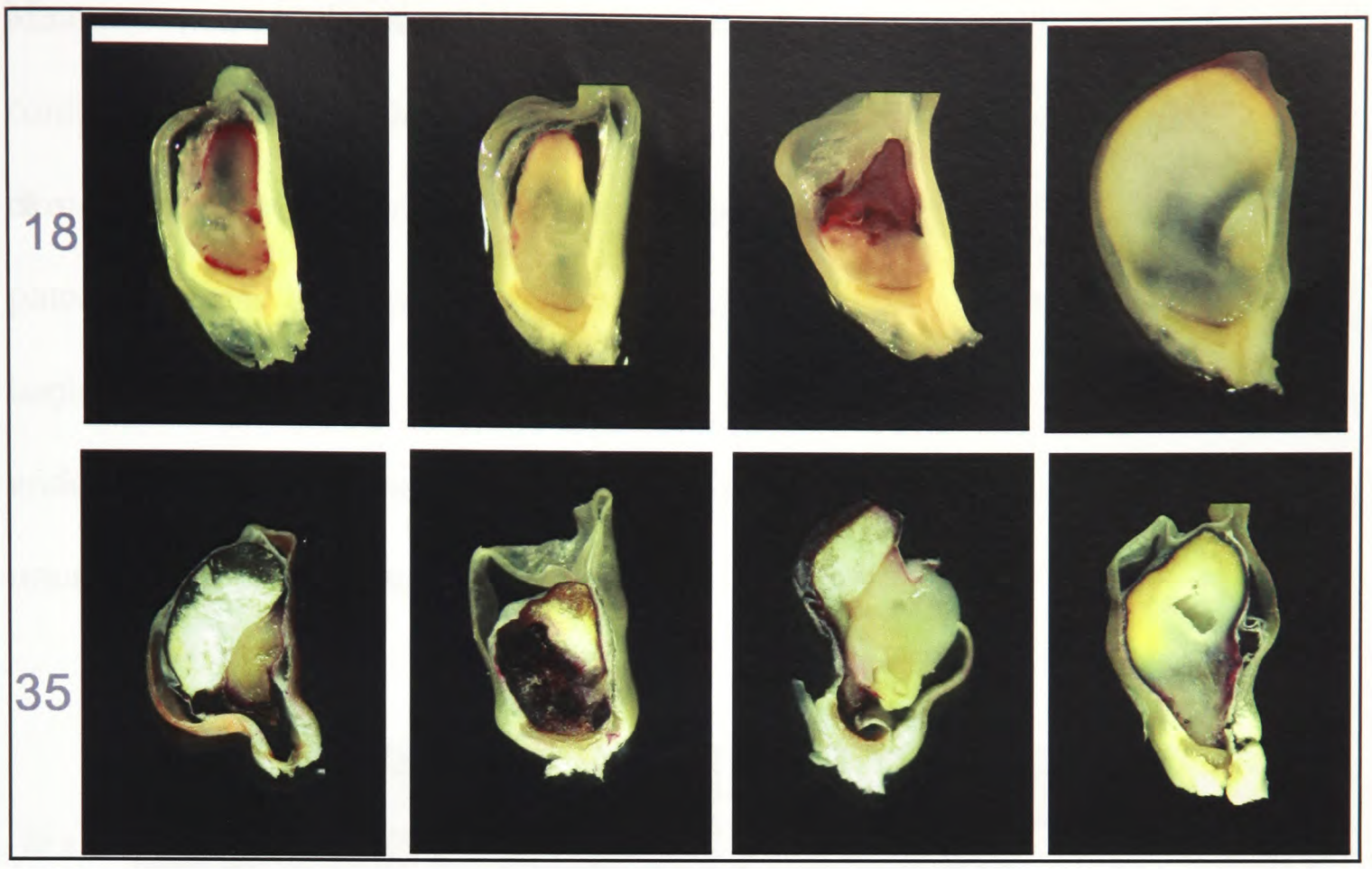


Figure 4-24 Ectopic localization of anthocyanin in paternal excess crosses

Endosperm samples were cut by hand to reveal a range of different localization of anthocyanin pigment. Anthocyanin was observed from 18 DAP until the latest stages of development, two time points are shown here, as listed in DAP on the left of the figure. Anthocyanin is normally confined to the peripheral aleurone layer (as shown for the 'escaped' regular looking kernel on the right-hand side of the 18 DAP samples. Scale bar represents 5 mm

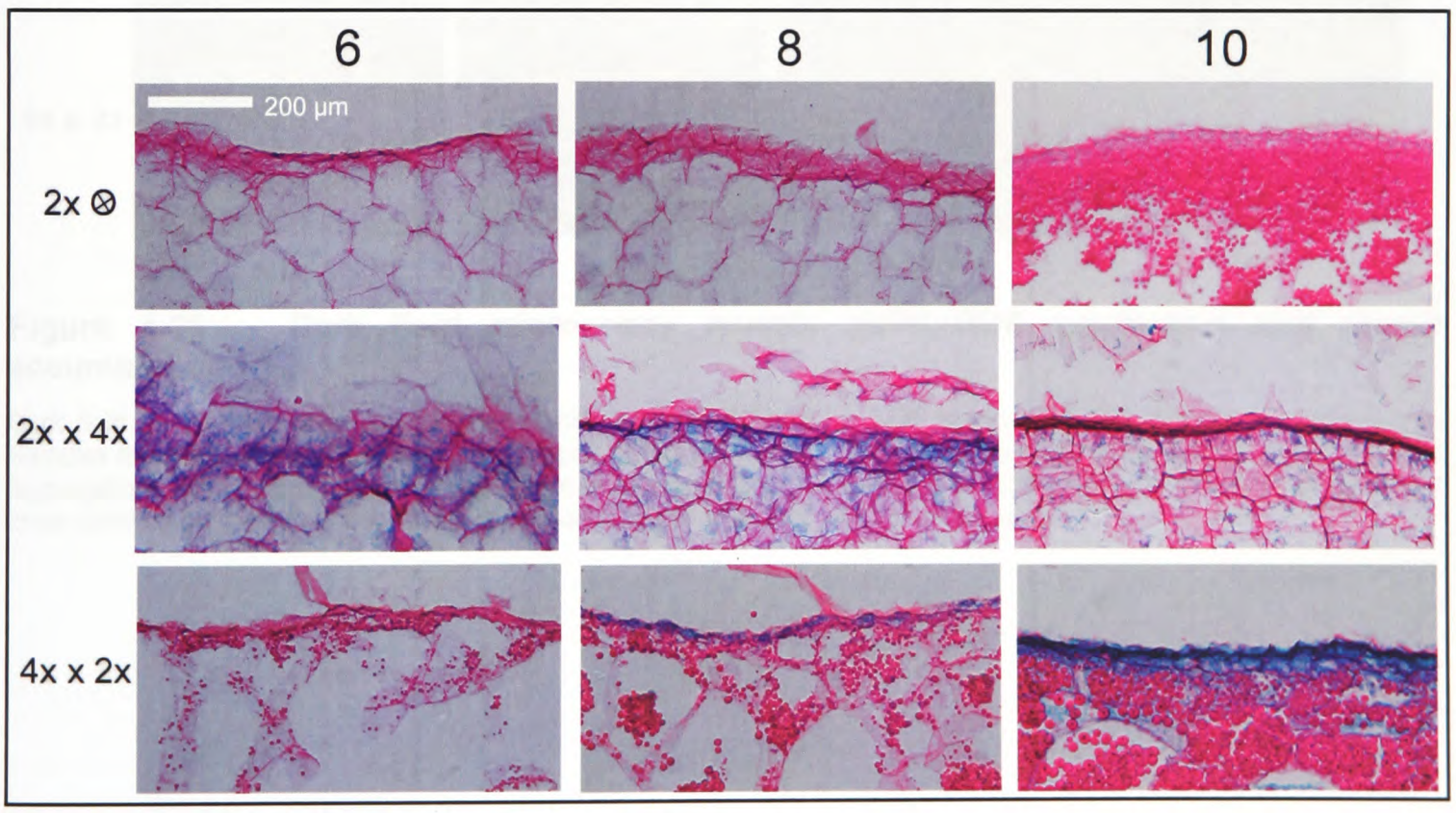


Figure 4-25 pVP1::GUS reporting profiles in comparison to aleurone identity

Samples were stained for GUS, and after sectioning were stained with PAS, to allow a comparison between the GUS localization (blue), and cell morphology (PAS stains polysaccharides pink).

4.2.11 Commencement of starch accumulation is altered in interploidy crosses

Maternal excess endosperms commenced accumulation of starch early, from 6 DAP compared to the 8-10 DAP in the balanced cross controls as demonstrated by white clusters in dark field images of the endosperm (Figure 4-26). In contrast the paternal excess endosperms showed delayed initiation of starch accumulation, beginning from around 14 DAP, and often it was restricted to the apical half of the endosperm, whereas the lower half either remained glassy, or developed into an unusual tumour-like tissue.

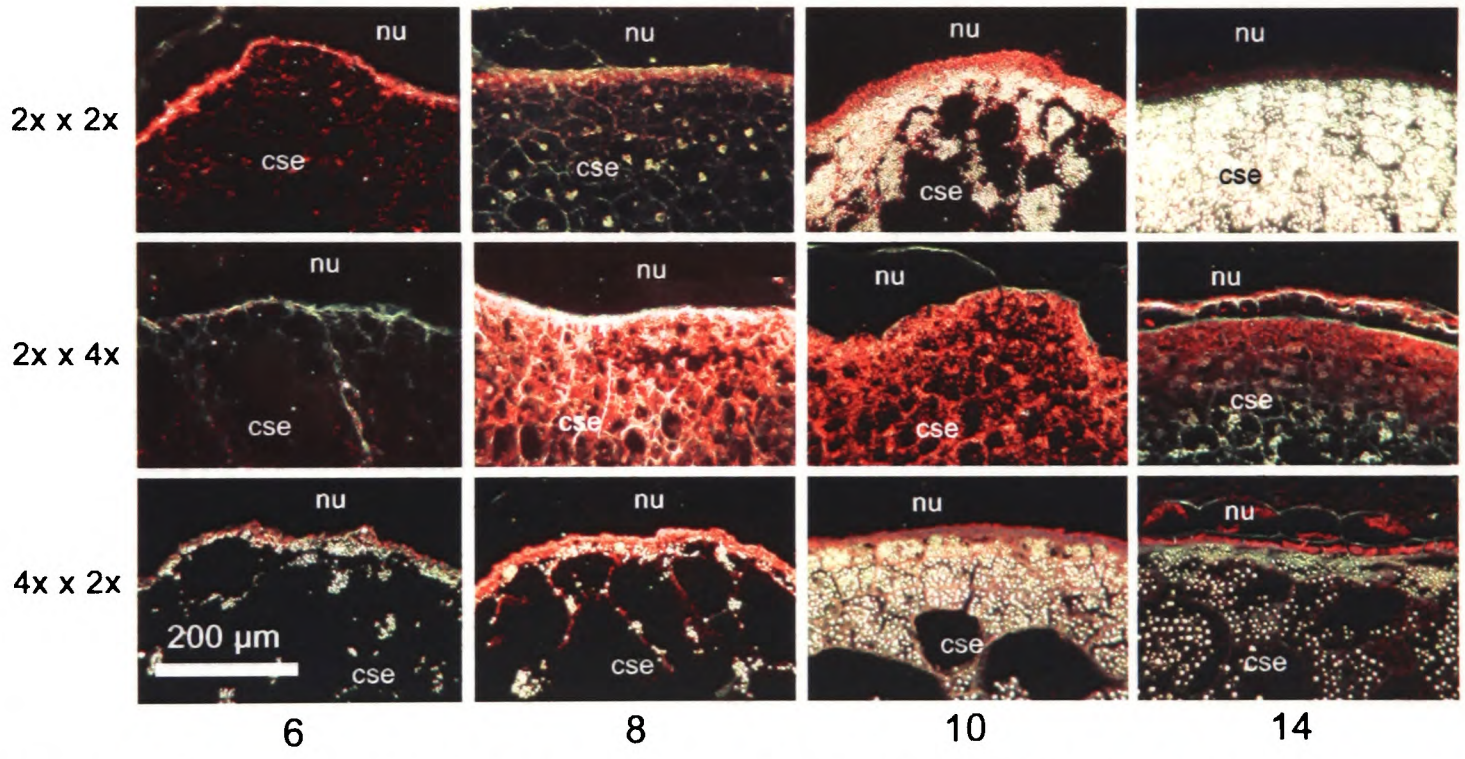


Figure 4-26 Dark field microscopy reveals pVP1::GUS production and starch accumulation

Dark field optics were used on samples which had been stained for GUS activity, the GUS precipitate appears red. Samples are from germinal sections, at the apex of the endosperm. Accumulation of starch is also observed by white aggregations within cells. Samples were taken at the number of DAP listed under the graph, for the three directions of cross detailed on the left of the figure. nu = nucellus, cse = central starchy endosperm.

4.2.12 Aleurone formation is distorted in both directions of interploidy cross

In order to assess whether the observations made with the pVP1::GUS marker also accurately reflected the acquisition of aleurone phenotype, late stage kernels were stained with toluidine blue, and visualized under bright field optics (Figure 4-27). This clearly demonstrates a difference in the differentiation of the aleurone in these crosses. In both the 14 and 30 DAP control samples (Figure 4-27A and B), the aleurone comprises regularly sized cells forming a single cell layer, and in the later stage endosperm ,these are cuboidal in profile. In the paternal excess cross (Figure 4-27C), the cells are not cuboidal, have a variable shape, and there are multiple and variable numbers of cell layers. In the maternal excess cross there is only a single cell layer as in the control, and the cells are more regularly sized and shaped than in the paternal excess. However, the cells in the maternal excess aleurone vary are flatter than in the 30 DAP control, and are more rounded.

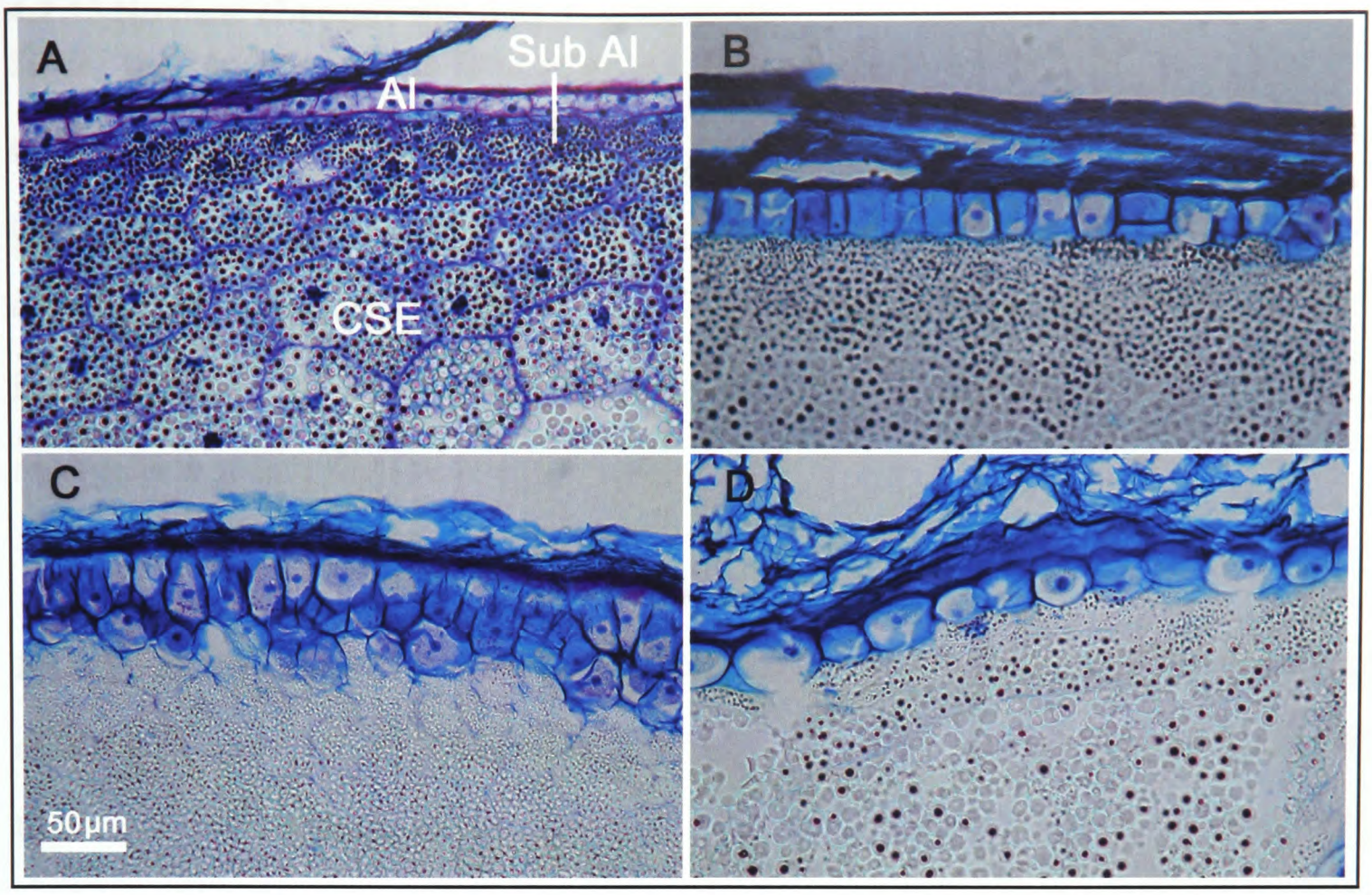


Figure 4-27 Toluidine blue staining of peripheral layer

Sectioned kernels were stained in toluidine blue, samples were taken from near the crown in a sagittal section. **A** and **B** represent balanced diploid control crosses; **C,** paternal excess; and **D,** maternal excess. **A**) is a 14 DAP sample, and **B-D** are around 30 DAP samples. Al = aleurone, Sub Al = Sub aleurone, and CSE = starchy central cells. A and B display a single cell layer aleurone, and in B the sub-aleurone is also clearly distinct. In the paternal excess (**C**), there are multiple and variable numbers of cell layers, and are disorganized, cells vary in size and shape. **D**) Although only one cell layer is present in the maternal excess cross, the cells are flatter than in the control cross, and have less cuboidal profiles.

4.2.13 Nuclear changes through development

Since the endosperm is characterized by changes in DNA content due to endoreduplication, DAPI was used to observe the appearance of nuclei at various stages though development (Figure 4-29). Clear differences are revealed in the interploidy crosses. In the control, numbers of nuclei remain fairly constant, but their size increases between 6 and 14 DAP, but afterwards the nuclei become smaller, and cell enlargement means they become more dispersed.

Paternal excess endosperm show considerably more nuclei in the same area, suggesting cells are smaller. Nuclei are considerably smaller than the control at 10 DAP, despite the larger DNA content in the tetraploid genome, presumably this is because these nuclei are highly condensed due to active mitosis, which by 10 DAP would normally be ceasing in a balanced cross. By 30 DAP the central area of the endosperm is appears to have lost DAPI fluorescence, possibly through cell death, or components of the cells may be interfering with the signal from the DAPI. However, the basal part of the endosperm is often frequently malformed, showing numerous tumour-like outgrowths (Figure 4-28), these appears to remain highly proliferative, and comprise many condensed nuclei Figure 4-30C.

Nuclei in the maternal excess display enlarged nuclei, in comparison with the control at both 6 and 10 DAP, suggesting they may be at a later stage in development reflecting the normal enlargement of the nuclei.

4.2.14 Embryo development in interploidy crosses

During staining for GUS it was often noticed that an embryo was not easily located in the paternal excess kernels (Figure 4-14 and Figure 4-17), and sectioning of samples

also failed to find even a reduced embryo, probably only a few percent of kernels actually had a distinguishable embryo, although often when they did they were sometimes very large and misshapen (Figure 4-31). In contrast embryos were consistently found in maternal excess kernels and were often larger in proportion to the size of the endosperm than in balanced crosses. Kernels were found in sections, and stained with DAPI (Figure 4-32), which demonstrates the rapid growth in the diploid, and the comparative delay in the maternal excess, where formation of an axis is retarded, and there appears to be very little organization, such as the files of cells present in the diploid control. Similarly the paternal excess embryo is delayed, particularly in the two samples at 6 and 10 DAP, although the 14 DAP sample appears to have undergone the first stages of axis formation.

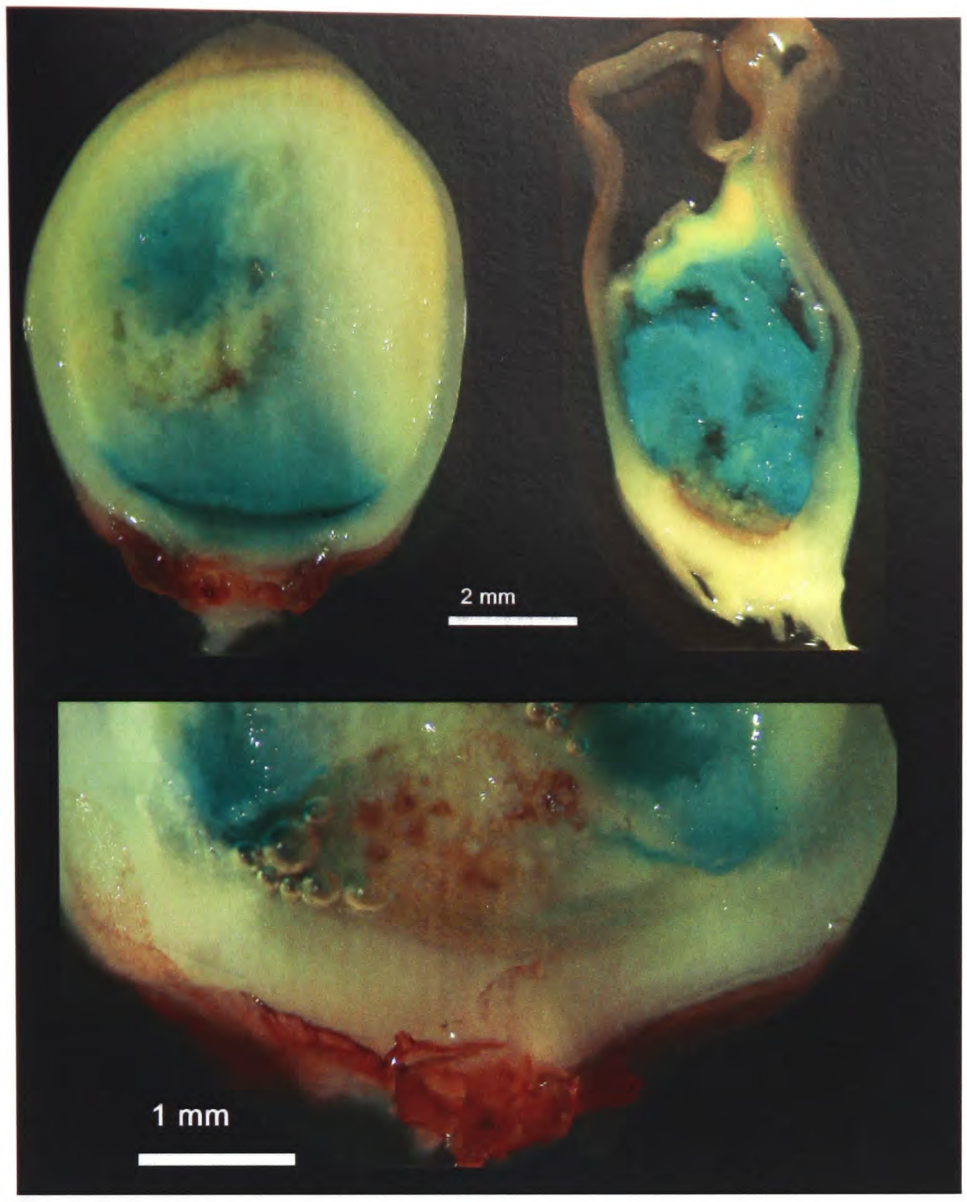


Figure 4-28 Tumour like outgrowths of the surface of paternal excess endosperm

Late stage paternal excess kernels, frequently displayed areas of high proliferative growth in the basal region, apparent through a highly irregular region ,comprising many of small spherical outgrowths of the tissue margins, as displayed in these hand-cut 30 DAP kernels, which are cut in the horizontal (top-left, and bottom), and transverse planes.

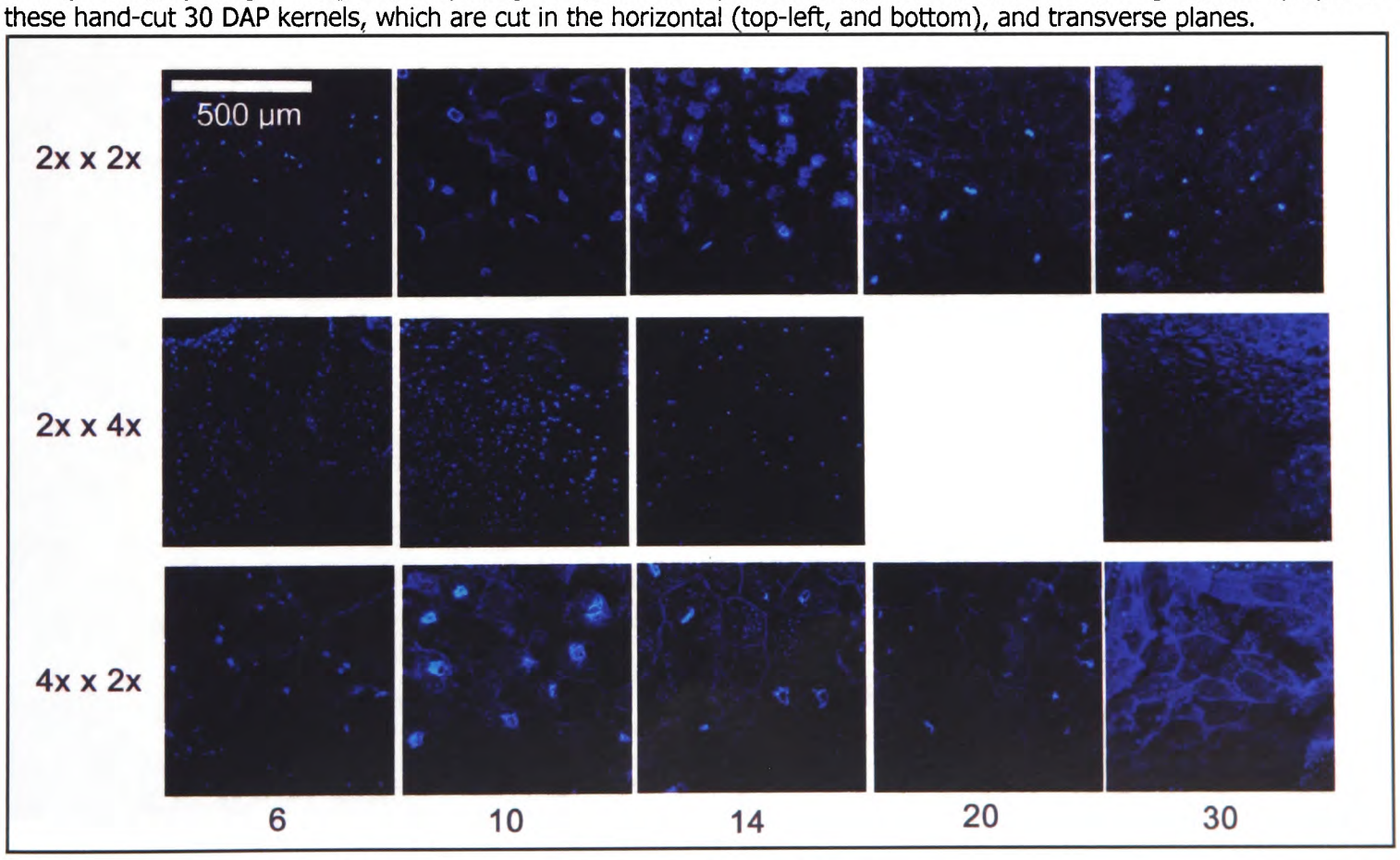


Figure 4-29 DAPI staining of centrally located nuclei

Endosperm sections were stained with DAPI, and centrally located regions from sagittal sections were imaged from samples at the time points noted in DAP below the panels.

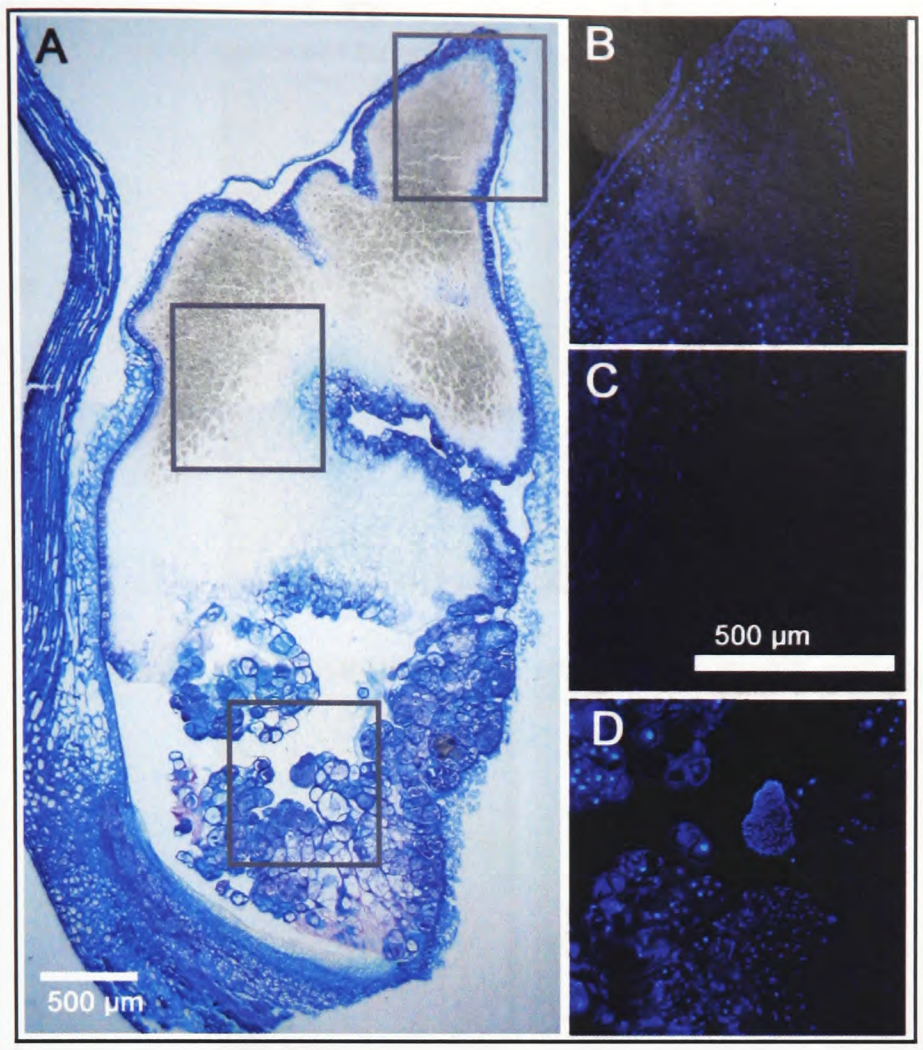


Figure 4-30 Areas of different mitotic activity in 26 DAP paternal excess endosperm

A) A late stage (30 DAP) paternal excess endosperm (stained in toluidine blue) reveals zones of high proliferative activity, and more typical central regions which lack DNA staining following staining for DAPI (**B**-**D**). **B**) Reveals the aleurone and a few underlying central starchy cells still stain for DAPI, as in control crosses. **C**) Centrally located cells show no fluorescence. **D**) Cells located in the basal area, where there are many irregular outgrowths, display a high mitotic activity.

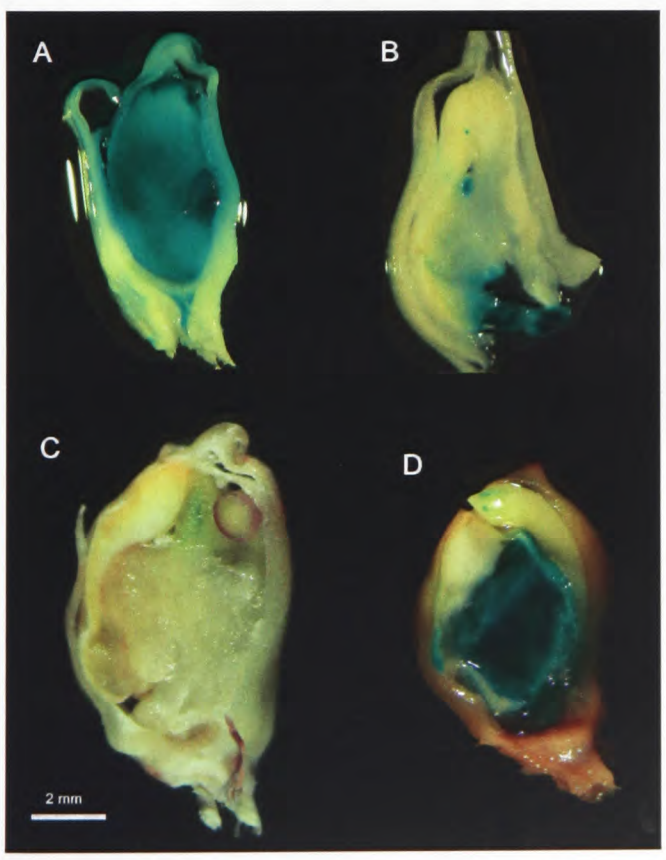


Figure 4-31 Appearance of embryo kernels in paternal excess crosses

Figure demonstrates the variability in embryo sizes from different paternal excess kernels (**A** and **B** 14 DAP kernels, **C** and **B** 30 day kernels), all kernels sectioned in a sagittal plane.

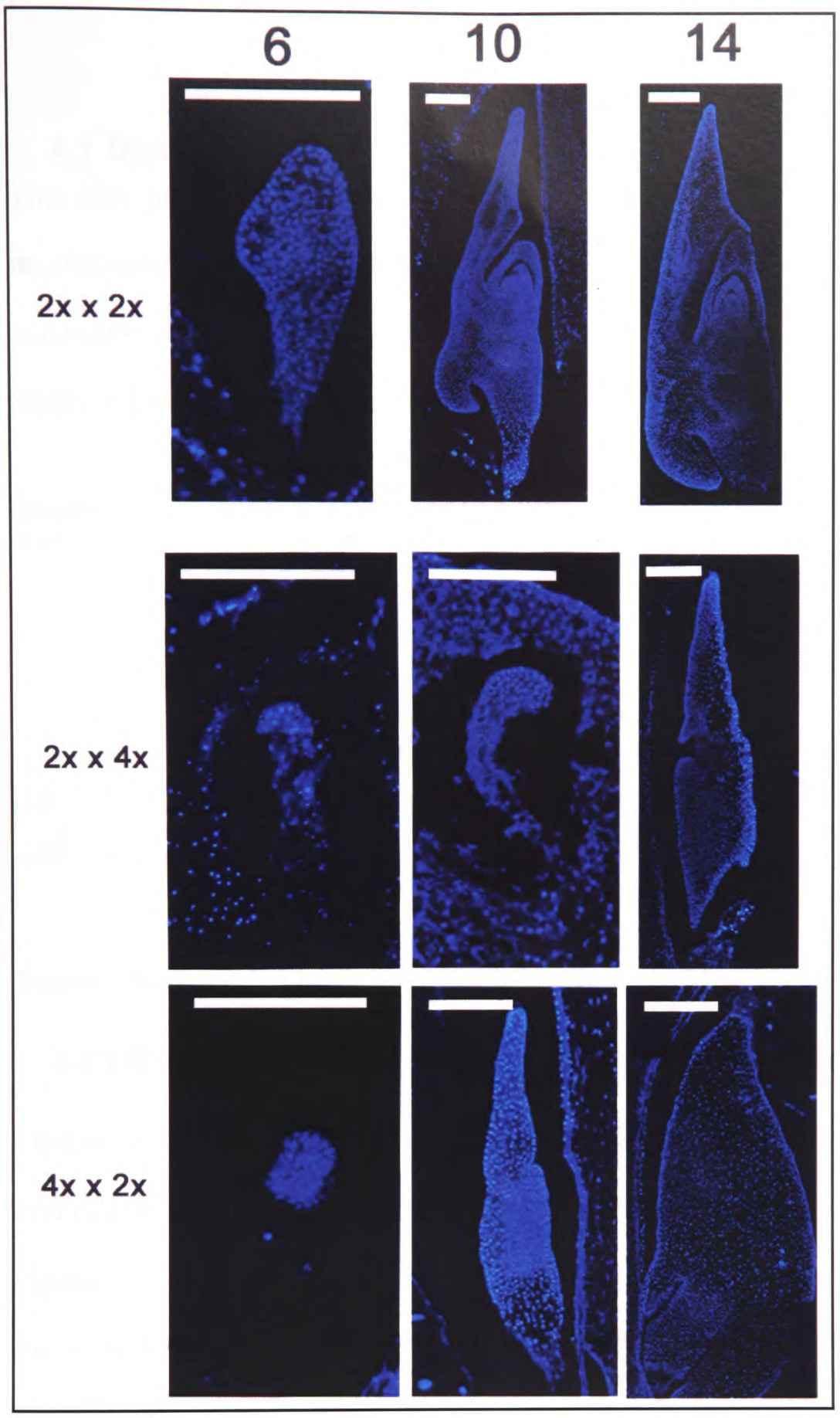


Figure 4-32 Appearance of embryos

Embryos were located in sectioned kernel samples at the time-points listed at the top in DAP (6, 10 ,and 14). DAPI, was used to stain nuclei, and show general embryo morphology. Rule = 500μm.

4.3 Discussion

The data presented in this chapter has provided some novel insights into how the mechanisms of cell fate specification are disrupted in interploidy crosses. This summary will detail and consider the importance of these observations. Table 7 Summarizes the main findings from this work.

	Paternal Excess	**Maternal Excess**
General observations	Delayed differentiation of aleurone and CSE. Aleurone is irregular and found on ectopic surfaces. BETL appears to form aberrantly. Delayed starch accumulation. Highly proliferative, condensed nuclei. Rarely produces an embryo. Central region disintegrates. Proliferative zone forms in adgerminal margin. Mis-expression of genes from 10 DAP onwards. High variability of seed size	Premature differentiation of aleurone and CSE. Aleurone cell morphology is slightly distorted, and layer invades BETL region. Highly reduced BETL domain. Accelerated starch accumulation. Early cessation of mitosis. Embryo usually present but often delayed. Seed size relatively regular. Endosperm gradually shrinks.
pVp1::GUS	Throughout periphery, even where misshapen.	Throughout periphery. Invades region of BETL
pBETL::GUS	Starts as irregular clusters of a few cells, predominately in BETL region. Later becomes more widespread throughout whole endosperm.	Consistently found in reduced domain at basal germinal corner, but low frequency.

Table 7 Summary of histological observations following interploidy crosses

4.3.1 Balanced crosses with reporter lines reveal markers for cell types but not cell identity

Central to this work was the use of reporter lines that marked certain cell types in the endosperm. Although these markers have previously been reported in the literature (Costa et al., 2003; Hueros et al., 1999a), their use in a developmental survey has not been carried out, nor has a clear correlation with the histological appearance of tissue-specific cells been described.

The pVp1::GUS reporter displays a pattern very close to the expected cell domain of the aleurone, in that it was localized consistently to the outer periphery of the endosperm margin throughout the period profiled (Figure 4-3 and Figure 4-4). GUS is first seen at the endosperm apex at around 6 DAP, and processes down the margins of the endosperm towards the base, extending to the base by 8 DAP. This is consistent with

reports of wave-like developmental processes for other aspects of endosperm development (Becraft and Asuncion-Crabb, 2000; Kowles and Phillips, 1988). It is also evident from the current work that the transcriptional controls which regulate the expression of the Vp1 gene in both the endosperm and embryo differ, since the reporter is active in the embryo several days earlier than it is in the endosperm. Closer inspection of the material through microscopy (Figure 4-5) revealed an interesting observation: that the reporter activity precedes the adoption of an aleurone identity. The first observation of GUS activity occurs when the peripherally located cells are histologically no different from the underlying cells which will become CSE cells. At this time, there is also some diffusion of the GUS product, presumably because at this stage the cells do not have mature cells walls. True aleurone identity, as assessed by histology, did not appear until around 10 DAP. Thus, although the pVp1::GUS reporter is an excellent general reporter of aleurone fate, it is not an exclusive reporter of aleurone identity, and as such it is important that any pattern of aleurone reporting is also verified by histology to see whether marked cells also have the expected morphology.

The pBETL1::GUS line also displays a reporting pattern similar to the expected BETL cell domain (Figure 4-6 and Figure 4-7). From around 10 DAP and throughout the period monitored, GUS activity was located at the basal portion of the endosperm at the region overlying the maternal pedicel, in a domain which expanded in relation to the increase in endosperm size. The simultaneous appearance of GUS activity along the entire base of the endosperm from 9 or 10 DAP suggests that all cells located in this region must begin to express the GUS reporter at the same time, rather than starting in an initial focus of a few cells that broadens across the whole base. Samples were investigated using light microscopy at the time point at which the BETL is reported to form (Kiesselbach, 1949a). This confirmed the initial appearance of typical BETL cells at 6 DAP, and the transformation of all basal cells by 7 DAP (Figure 4-8). Therefore,

BETL morphology precedes GUS activity by three to four days, revealing the pBETL1::GUS line to be a good reporter of later stage BETL cells, although not of BETL identity *per se*.

Balanced crosses were performed reciprocally between the reporter lines and H99-2x plants. This work suggested that the timing of GUS activity was very similar between reciprocal crosses, although there was a delay of around 24 hours for the first appearance of GUS activity for both reporters when the transgene was introduced through the male parent. This could reflect either silencing of the paternal transgene, but most probably reflects different developmental rates between the crosses. Most importantly the location of GUS activity was identical between the reciprocal crosses. Hence, by the stages at which the GUS transgenes are active, reporting patterns of the markers are not affected by the parent through which it was inherited.

For both reporter lines, the limits of the interpretation of the data must be appreciated: namely that the activity of the endogenous promoter in its true genomic environment is not being observed, and nor is GUS activity a true assessment of the protein or mRNA localization. The endogenous promoter may be modified by distant modifiers and enhancer elements which would not be included on the approximately 1kb of promoter included in these constructs. Additionally, the half-life of the GUS protein is most likely to differ from that of the endogenous Vp1 and BETL polypeptides. Furthermore, the blue GUS product may diffuse between cells - which may not be true of the mRNA or protein.

In spite of these caveats, GUS is commonly used to assay patterns of gene activity in plants, and remains one of the most straightforward methods by which this can be achieved. Generally GUS reporting is held to give a reasonably accurate representation

of true gene activity (Guivarch et al., 1996), and gross deviations from control patterns will usually be suggestive of changes in expression, whereas more subtle deviations might require techniques with a higher resolution such as *in-situ* RNA or immuno-labelling. Protein fusions with fluorescent GFP-derived proteins might offer an alternative approach although this approach has similar provisos associated with promoter fusions and fusion-proteins, but possibly fewer problems with substrate diffusion (Haseloff, 1999; Sheen et al., 1995).

4.3.2 Maternal excess interploidy effects on seed development

Following interploidy crosses in which the tetraploid plant is used as the female parent, both endosperm domains identified by the reporter lines are disturbed.

Although the aleurone reporter is observed at the periphery of the endosperm commencing around the time expected, based on observations from the control crosses (Figure 4-9), its location is altered. GUS activity persists in the periphery throughout the course of development, as anticipated, although its activity may cease sooner than normally expected. Although the periphery stains as in control crosses, this region is very different in shape and size to kernels from balanced crosses, the endosperms of which are almost spherical with a smooth outer margin expanding throughout most of development. The endosperm from maternal excess crosses show a smooth periphery only in the first 12 days, after which the endosperm shows signs of shrinkage with the resulting deformation of the peripheral surface into an uneven surface. In spite of this the margin continues to stain for GUS.

Most significantly, the aleurone marker is also seen to be expressed ectopically, as it extends into the region which would normally be occupied by BETL cells. The extent to

which this ectopic aleurone invades this would-be BETL region varies between kernels, although it is seen to some degree in all. Usually, there is a small circular region which does not stain for GUS, and this is consistently located in the same location, on the germinal side of the endosperm.

Histological analysis of GUS-staining cells in these crosses confirms that they have an aleurone-like morphology - including most of those cells located in the would-be BETL region (Figure 4-25). The cells appear to show aleurone cell morphology sooner than in control crosses, from around 6 DAP. These aleurone cells are not characteristic of those from balanced crosses however, being flatter and rounder. Perhaps this morphology represents an altered developmental profile conferred by the additional maternal genome on these cells, or it might simply be a mechanical effect. Unlike in balanced endosperms, where high rates of cell growth of the underlying cells demands high rates of cell division of the aleurone layer to keep pace with the expanding surface area needs. Possibly this growth generates forces between adjacent aleurone cells that might contribute to their rigid morphology, whereas the maternal excess endosperm do not show such expansion, and are apparently in a process of shrinkage by 30 DAP, so such potential forces would be absent (Figure 3-4).

It has been previously reported that cells that are located on the edge of the endosperm will form aleurone cells, even when the edge is not typical of normal kernels, such as those formed by invaginations, or enclosures of the cell surface in certain mutants (Costa et al., 2003, Becraft, 2000). This has led to the proposal of a "surface" rule for the formation of the aleurone (Olsen, 2004a). In normal development, however, the BETL defies this rule, by being a surface cell that does not form aleurone. These interploidy crosses reveal that even cells located in this region will, under some circumstances form an aleurone. It is not clear what prevents these cells from forming

an aleurone in normal development, but it might be simply because these cells, evidently capable of either aleurone or BETL identity, normally receive a BETL-forming signal first and it has been reported that such cells might be specified as early as before cellularization (Gomez et al., 2002). It may therefore be that the maternal excess endosperm is somehow incapable of responding to the BETL-inducing signal.

The BETL marker, in contrast to that for the aleurone, is highly disrupted in maternal excess crosses. Only a low proportion of kernels display GUS activity, and those that do have a weak signal, which occupies about a third of the normal BETL domain. The location of GUS activity in these crosses is consistently in the basal region closest to the embryo - the region that was observed not to appear aleurone-like for the pVp1::GUS marker. Histological analysis of these endosperms suggested that the cells staining blue did possess some of the typical signs of BETL morphology, such as feather-like invaginations, although the region was always highly reduced, and never extended more than one cell deep into the endosperm. Those cells located within the gap of the ectopic aleurone domain, and that did not show BETL characteristics were neither aleurone, nor CSE, nor BETL-like. This suggests that a distinct transition between the reduced BETL and aleurone fails to be established in these endosperms (Figure 4-12). This suggests that these cells may have an intermediate identity between aleurone and BETL cells, which would be a very interesting scenario. *In situ* RNA detection for markers for both might clarify this issue.

General indicators of development suggest that the normal developmental programme is accelerated by maternal genomic bias. Maternal excess endosperms are also characterized by considerably fewer numbers of larger cells, the cellular dynamics of which appear to be accelerated compared to control crosses (Figure 4-29). These cells accumulate starch prematurely (Figure 4-26). Interestingly, embryos are found in all

maternal excess kernels, although their development frequently halts at some point during development (Figure 4-32).

4.3.3 Paternal excess interploidy effects on seed development

The pVp1::GUS reporter reveals that a partial aleurone fate is adopted at approximately the same time as in balanced kernels, since signal appears in the peripheral cells around 6 DAP (Figure 4-14), although formation of a distinct layer does not appear until later in development (Figure 4-25), and even then it is highly aberrant (Figure 4-27). The endosperms in these crosses have a highly convoluted surface, perhaps due to aberrant mitotic activity, particularly in the adgeminal face of the kernel (Figure 4-28). However, in spite of the invaginations and crevasses which form on the endosperm surface, GUS activity is seen to appear at all these locations, giving further support to the 'surface' rule for aleurone identity (Olsen, 2004a). At later stages of development, small patches of GUS activity are observed at random locations throughout the endosperm. In this direction of interploidy cross the basal region does not form an extension of the aleurone layer as occurs in the reciprocal cross.

The BETL marker shows a completely different response to the maternal excess cross. At the time the BETL is first specified in control crosses, cells located in the equivalent region in the paternal excess endosperm stain intensely for PAS, but lack the expected cell wall modifications Figure 4-23. This is in agreement with a study by Charlton (1995). Although contradictory indications were reported in a separate study (Hueros et al., 1999b), where cells in this region, which also expressed the *BETL2* marker gene in a patchy discontinuous fashion, showed typical BETL cell wall modifications. The discrepancy might be accounted for by either genotypic differences conferring different tolerances to genome distortion, or because different stages of development were studied in the two investigations (10 DAP in Charlton *et al* versus 15 DAP in Hueros *et*

al). The current work demonstrates that BETL cells are more likely to form at later stages (Figure 4-23). The patchy expression of *BETL2* and reported inter-kernel variability (Hueros et al., 1999b) is, however, strikingly similar to the GUS observations made in this study Figure 4-22, which justifies the use of the pBETL::GUS reporter. It also suggests that because the cells located in this region are expressing mid-term BETL cell genes, yet appear atypical histologically, these cells have undergone some genetic program towards BETL cell identity - but it has gone awry.

From as early as 10 DAP in crosses involving H99-4x, but from a slightly later stage for those involving HybT, ectopic GUS expression occurs outside the BETL region. Small clusters of ectopic activity can be seen both around the periphery, and also at internal locations. Intriguingly, this is also seen for the pVp1::GUS marker in paternal excess endosperms, and also for an endogenous gene involved in anthocyanin synthesis. Since the BETL and aleurone have been proposed to be specified by different means (Olsen, 2001), the fact they both show a very similar effect in late development suggests it is not because these cells have adopted the cell fate marked by the reporter they are expressing. Instead, it seems more likely that this aberrant activity is due to a more general mechanism that is causing sporadic gene expression generally.

In contrast to the maternal excess, endosperm from paternal excess crosses display a high rate of proliferation during early to mid stages of development, which persists in the base of the endosperm, where the tissue has become a highly proliferative and tumour-like, confirming the study by (Cooper, 1951). It seems that this higher rate of proliferation causes a resultant delay of normal developmental progress, as evidenced by late starch accumulation (Figure 4-26), and late aleurone layer formation (Figure 4-25).

The difficulty of finding embryos in both hand-cut or sectioned samples of paternal excess interploidy kernels, suggests that they often failed to form, and when they did they were often highly retarded in growth. In spite of this, a small proportion of triploid seedlings could be recovered from these kernels; interestingly these were not necessarily from the heaviest kernels.

5 Effect of Interploidy crosses on gene expression

The biological consequences of interploidy crosses are profound, and result in abortion of the seed. The results presented in this thesis demonstrate that interploidy crosses result in a failure to define correctly the domains of the endosperm, in spite of its relatively simple organization. Such crosses also result in a change in the normal rates of endoreduplication (Leblanc et al., 2002). Interploidy crosses result in not only distortion of the total genome number and hence dosage, but also results in a change in the ratio between maternally-inherited and paternally-inherited genomes. Nevertheless, how these changes come to result in such gross morphological and biological changes is unknown, although it has been proposed for both maize and arabidopsis that it might result from effects of different sets of paternally and maternally imprinted genes (Birchler, 1993; Lin, 1984; Scott et al., 1998).

The advent of microarray technology and the production of EST libraries from maize, provides a timely opportunity to survey a near-global representation of the maize transcriptome, as presented in this chapter. Additionally custom-produced arrays were produced to assess a subset of genes of specific interest, largely comprising chromatin remodelling factors, an overview of the activity of these genes being given below. Independent verification of some genes suggested as being interesting from the microarrays will be verified using an independent approach, using semi-quantitative RT-PCR.

The chapter documents the rationale for the selection of the arrays and tissue used, in addition to providing details of the experimental design and labelling methods. Data from both of these experiments will be discussed in the context of our current understanding of seed biology.

5.1 Introduction

5.1.1 Microarray Production

The ability to measure the amount of a specific RNA transcript (mRNA) that is present in a cell, tissue, or organism of interest is one of the most fundamental techniques in the biologist's arsenal, and has served molecular biology well for many decades (Hames and Higgins, 1985). Initially the process of gel-separated RNA transfer to a membrane support of a whole RNA population allowed the investigation of the presence of a single RNA species or, following stripping and re-probing, several genes serially. The concept of a microarray turned this methodology on its head by immobilising the known genes (target) to a solid glass support, and hybridising fluorescently labelled populations of cDNA to this (Schena et al., 1995). The rapid expansion of genome sequencing projects has made it desirable to use this platform to assess large numbers of genes in an organism in such transcriptional experiments. Microarray experiments are now seen as 'fishing excursions' in which a general assessment of a transcriptional response to a given treatment or mutant is required, with the potential identification of many candidate genes of interest rather than the use of a clearly formulated assessment of the change of a specific gene under differing conditions.

For the current study, a very broad biological theme is being investigated with no *a priori* assumptions: namely what is the effect of interploidy crosses on endosperm gene expression in maize? As such, the recently developed microarrays produced by the maize community [5] offer an appropriate means to assess these effects. Additionally, custom-made arrays were produced prior to the availability of these larger-scale arrays, which did make *a priori* assumptions, as they contained a limited subset of gene classes.

[5] http://www.maizearray.org/

Although a brief review is provided below to justify the selection of protocols used in this thesis, the rapidly expanding field of microarrays is constantly evolving, and as such an exhaustive review seems inappropriate. However several reviews in the field are recommended: general protocols (Hegde et al., 2000; Holloway et al., 2002); experimental design (Churchill, 2002; Yang and Speed, 2002); data processing (Causton et al., 2003), and data analysis (Leung and Cavalieri, 2003).

5.1.1.1 Target

The nature of the spotted probe is an important consideration, and various options exist. The first microarrays employed amplified cDNA fragments (Schena et al., 1995), but more recently oligomeric nucleotides (oligomers) of varying length are used (Call et al., 2001). cDNA fragments offer the advantage of being cheaper for small-scale slide prints, and being double stranded, can hybridize with antisense and sense probes. However, cDNA fragments should generally be greater than 150 bp in length, else processing becomes difficult. Length is an important consideration because, although a positive correlation between length and signal intensity exists (Chou et al., 2004), longer targets have lower specificity, which is a particularly important consideration for organisms such as maize which contains so many duplicate genes (Fernandes et al., 2002). Studies into the optimal length of oligomers, in terms of balance between signal and specificity have suggested lengths of 60 and 150 bases are best (Chou et al., 2004; Hughes et al., 2001). Comparisons between the performance of cDNAs and oligos have generally yielded reasonable correlations (Hughes et al., 2001; Lee et al., 2004; Wang et al., 2003b).

The target DNA can be delivered in a range of solutions, but for the custom array a solution of 1.5M betaine and 3x SSC was chosen for its reported low volatility, the

production of spots with even size and low noise: signal ratios and good performance in hybridisations (Diehl et al., 2001; Hessner et al., 2003).

5.1.1.2 Labelling Method

Although platforms for hybridisation of individually labelled RNAs exist (Affymetrix, Santa Clara, CA)(Lipshutz et al., 1999), most cDNA and oligomer protocols utilize dual labelling of two RNA populations. The most commonly employed flurophores are cyanine 3 and 5, and these were used for the custom microarrays (Hegde et al., 2000), although some interesting novel alternatives are emerging (gold and silver labelling (Gupta et al., 2003; Huber et al., 2004)). Labelling can be performed by the direct inclusion of the fluorophore in a reverse transcription reaction, or by including a modified base followed by a secondary coupling step to the fluorophore (indirect labelling) (Hasseman, 2001). Both techniques are similarly comparable, although indirect labelling is sometimes considered less prone to gene-specific biases. Both techniques are however limited by their need for a large starting amount of RNA (up to 50 μg (Schena et al., 1995)). For experiments where material is limiting, alternative methods are required to obtain a reasonable signal with lower starting amounts of RNA.

5.1.1.3 Signal amplification

One of the most commonly used methods of signal amplification is by using a T7 RNA polymerase to perform *in vitro* transcription on reverse transcribed RNA (van Gelder et al., 1990), which can amplify RNA by up to several hundred-fold, and was employed for hybridisations to the custom arrays in this thesis. Although some concerns have been raised about whether the system amplifies in a totally linear fashion depending on the starting concentration (Baggerly et al., 2004), it is generally accepted that although a

bias may exist, it is usually consistent, and good correlations have been shown between amplified and non-amplified samples (Feldman et al., 2002; Zhao et al., 2002).

An alternative approach to increasing signal is by the use of dendrimer molecules (Shchepinov et al., 1997; Stears et al., 2000). These are three dimensional branching macromolecular structures of self-annealed DNA sub-units, which aggregate to form a mesh-like structure, which not only carries a large number of fluorescent molecules, but also many copies of a DNA sequence which is complementary to a sequence tagged onto the end of the oligo-dT primer used in the first strand cDNA synthesis reaction. The labelled cDNA can hybridise to the highly fluorescent capture reagent. Dendrimer technology has been shown to show the low dye-specific effects in comparison with other methods (Baggerly et al., 2004), and a low level of background signal and utilizes sub-microgram amountx of starting RNA (Stears et al., 2000).

5.1.2 Gene regulation by chromatin remodelling

Due to the growing awareness of the importance of epigenetic modifications in gene regulation, as explained in chapter 1.3, their potential importance in manifesting transcriptional changes observed in interploidy crosses was considered, particularly in light of the fact that many of these factors are known to have dosage-dependent activities. This introductory chapter provides a brief overview of the roles of these genes. The field is an area of considerable research, at both a genetic and biochemical level, and as such details of comprehensive reviews are provided.

The DNA present in a cell is associated with a large number of different proteins, collectively referred to as chromatin. The bulk of chromatin comprises the DNA itself and repeating multimeric subunits called histomers which comprise two subunits, each

of which is composed of four different proteins called histones (2A, 2B, 3 and 4). The histomer provides a surface over which the DNA can coil two times by neutralization of its negative charge, which facilitates DNA compaction, each histomer being associated with 146 bp of DNA, a complex called a nucleosome (Aalfs and Kingston, 2000). Neighbouring nucleosomes are separated by a linker region of DNA of 20-35 bp, which can be associated with a linker histone, H1. Linker histones can then facilitate the further compaction of the chromatin by coiling of nucleosomes into a fibre of 20nm diameter (Horn and Peterson, 2002), and still higher order packaging can be achieved by supercoiling of the chromatin. Native chromatin occurs in two histologically distinguishable forms. Heterochromatin is highly compacted, and is readily stained and visible in the interphase nucleus, and is primarily associated with inactive non-expressed regions of the genome. The euchromatic fraction of the DNA chromatin is not as easily stained, and represents the actively transcribed regions of the genome. Chromatin associated proteins are important in regulating the level of condensation of the chromatin. Generally speaking the more condensed the chromatin is, the more transcriptionally silent it is, as the compaction reduces the accessibility by transcription factors to gene promoters. Some chromatin associated factors modulate gene activity of the underlying DNA. Polycomb proteins, are generally associated with gene repression, whereas Trithorax group proteins are associated with gene activity.SWI2/SNF2 ATP-dependent remodelling enzymes are thought to be involved in facilitating the mobility of histone octamers along the DNA backbone, and as such may be implicated in both gene activation and repression.

5.1.2.1 DNA methylation

The most studied epigenetic mark is DNA methylation., This modification occurs on cytosine residues, usually within a symmetrical CpG or CpNpG context. The modification is commonly associated with heterochromatic DNA, and inactive genes,

whereas euchromatic DNA is mostly undermethylated, and particularly at the so-called CG islands in the vicinity of gene promoters. Methylation of DNA is carried out by specific enzymes, which catalyse the transfer of a methyl group from a carrier to cytosine residues. Three classes of methyltransferases are recognized in plants (Finnegan and Kovac, 2000). The MET1 class is a homologue of the mammalian methyltransferases DNMT1 (Kankel et al., 2003), and acts as a maintenance methylase which acts to maintain symmetrical methylation of DNA following DNA replication on a hemi-methylated template. *met1* mutants show decreased levels of methylation through the genome, although some genes show an increased level (Jacobsen et al., 2000) The chromomethylase class (CMT) represent a plant specific class of methyltransferases (Henikoff and Comai, 1998), and a mutant in a maize CMT gene, *Zea methyltransferase2* shows a 13% global reduction in this modification (Papa et al., 2001). The third class of methyltransferases is the DNA Reduced Methylation (DRM) class (Cao and Jacobsen, 2002).

It seems that DNA methylation is usually an irreversible event, which can only be lost through passive demethylation by DNA replication, however, an example of a specific demethylation activity has been shown to be necessary for activation of a gene, MEDEA, and is carried out by a repair-excision mechanism by a DNA gylcosylase enzyme (Choi et al., 2002).

In the maize endosperm, there is an apparent significant undermethylation of the female genome in comparison to the paternal genome and leaf tissue following fertilization and into mid-stage endosperm development (Lauria et al., 2004).

Certain proteins can bind to methylated DNA, methyl binding domains (MBDs), and facilitate interactions between it and other proteins, such as chromatin remodelling

factors, which seem to enable the silencing effects of methylated DNA (Springer and Kaeppler, 2005; Zemach and Grafi, 2003).

5.1.2.2 Histone modifications

The histones are globular proteins, but have flexible N-terminal tails which protrude from the surface of the nucleosome. These tails are the target of enzymatic modifications which modulate affinity for other chromatin associated factors and transcriptional regulators. Methylation of histone tails can have both positive and negative effects on gene transcription, whereas acetylation of tails is generally associated with gene activation, although an antagonistic de-acetylation activity is also found in cells.

Histone methylation occurs on specific residues of histone H3 (Lysines 4, 9, 27, and 36) and H4 (lysine 20) (Lachner and Jenuwein, 2002; Tariq and Paszkowski, 2004), and methylation of different target lysine residues leads to different consequences. It is generally held that the methylation of K9 of histone 3 leads to gene repression, or heterochromatic silencing, whereas methylation of K4 leads to gene activation.

The enzymes which catalyse the methylation of amino acid residues are characterized by the presence of a domain of around 130 amino acids called the SET domain; however these proteins do show some considerable diversity in their domain composition, and four main families have been proposed (Kouzarides, 2002). These families are: the SUV39 family, which methylates H3-K9; the SET1 family which methylates H3-K4, the SET2 family; which methylates H3-K36; and the Retinoblastoma-interacting zinc finger (RIZ), which lacks a critical residue in the

functional domain, suggesting it may not function as an histone methyltransferase, but might have antagonistic activity.

A critical protein thought to be important in the manifestation of both silencing in heterochromatin and at repressed gene promoters is HP1, a known transcriptional repressor, which co-localizes with SUV39 around heterochromatin in mammalian cells (Lachner and Jenuwein, 2002). HP1 can bind methylated histone K9 residues via its chromodomain (Jones et al., 2000). Although the mechanisms of silencing by HP1 is unclear, it is postulated to act by linking multiple nucleosomes to reduce chromatin accessibility. Alternatively Hp1 might function through recruitment of other enzymatic activities.

Although it is considered that histone methylation is a static modification, recent evidence of a histone demethylase activity suggests it is in fact reversible (Shi et al., 2004).

One of the most flexible ways in which chromatin conformation can be altered, is through the addition or removal of acetyl moitities to histones (Lusser et al., 2001). Changes in histone acetylation are strongly correlated with changes in the expression of an mRNA transcript. Most organisms have several different classes of both histone acetylases and deacetylases, which means competition between these two enzymatic activities can determine the expression state of a gene (Graessle et al., 2001). Generally speaking, acetylation confers increased expression on genes, whereas removal of acetyl groups leads to silencing at those loci. However, some genes show the opposite effects (Munshi et al., 1998), and ultimately the consequence depends on the target gene in question, the composition of the enzyme complex, and the cell type in which it occurs.

Analysis of acetylation activity and gene structure between organisms suggests functional differences exist between genera (Pandey et al., 2002). For example, the major target of acetylation: in animals and fungi it is histone H4 , but it is histone H3 in plants (Waterborg et al., 1990).

The mechanism by which acetylation modulates gene expression is not certain, but could be due to decreased stability of the histone/DNA interaction, facilitating access by transcriptional enzymes. Acetylation may also have a more general effect changing chromatin packaging. Finally, the acetylation mark may change the affinity for other regulatory proteins.

There are three main families of histone de-acetylase, and they generally act as components of large multimeric complexes (Ahringer, 2000). The RPD3/HDA1 family is the most studied. A knockout in the arabidopsis *RPD3* gene displayed delayed flowering, although it is not lethal (Wu et al., 2000). The HD2 family is unique to plants (Dangl et al., 2001; Lusser et al., 1997). The SIR2 family is distributed widely throughout both eukaryotes and prokaryotes, and uniquely requires NAD^+ as a cofactor (Imai et al., 2000).

Histone acetylases comprise four families, and also act as part of large multimeric complexes (Ogryzko, 2001). The GNAT/MYST (GCN5-related N-terminal acetyltransferase/MYST) superfamily: the GNAT domain family genes are named after the *Saccharomyces cerevisiae GCN5* gene which was initially identified as a transcriptional activator (Neuwald and Landsman, 1997). The CREB-binding family is a large, diverse protein family, consisting of a number of protein domains, and is

unrelated to the GNAT/MYST family (Bannister and Kouzarides, 1996) and are more highly represented in plants than in animals A structural difference exists between plants and animals, in that plants lack a bromodomain: a domain involved in acetyl-lysine recognition. The $TAF_{II}250$ family contains a HAT domain which is unrelated to domains found in the other histone de-acetylase groups (Mizzen et al., 1996). The domain is around 260 amino acids long and is similar in humans, fruit fly and arabidopsis, and contains bromodomain motifs.

5.1.2.3 Polycomb Group Proteins

Polycomb group proteins are a diverse set of proteins which were first identified in drosophila and are involved in transcriptional repression (Otte and Kwaks, 2003). They were identified through mutations which caused homeotic transformations: that is mutations which cause the production of an organ or structure in an inappropriate (ectopic) location. This class of mutant encompass a wide array of proteins, which interact in functional complexes. Two complexes have been recognized, termed Polycomb Repressive Complex (PRC) 1 and 2, which repress genes not active in a certain cell type. PRC1 comprises several proteins, homologues of which are widely distributed through metazoans, but have not been identified in plants (Saurin et al., 2001). PRC2, however, is found in metazoans and plants, and comprises four main proteins in drosophila (Enhancer of zeste (E(Z)), Extra Sex Combs (ESC), and Suppressor of Zeste 12 (SuZ(12)) (Muller et al., 2002).

The PRC2 complex is thought to be directed to the DNA, possibly by the gene Pleiohomeotic, in a sequence-specific manner, where it can facilitate access to DNA by histone modifying enzymes such as deacetylases making the DNA sequence inaccessible to transcription factors. The PRC2 complex might also associate with

histone methyltransferases which methylate histone H3 at lysines 9 and 27 to promote recruitment of silencing factors.

The first PRC component in plants was identified in arabidopsis when the *Curly Leaf Gene* (*CLF*) (Goodrich et al., 1997), was shown to be highly similar in structure to the *E(Z)* gene, which is characterized by the presence of several domains including the SET, SANT and Enhancer of Zeste domain. Mutations in the gene resulted in misshapen leaves, and premature flowering. It was shown to be involved in the repression of genes required for the transition to flowering. Another homologue of *E(Z)* was identified through screens for genes involved in reproductive development which revealed a gene involved ion the repression of post-fertilization events, *MEDEA* (Grossniklaus et al., 1998).

Maize also has homologues of the PRC2 complex components. The three maize *MEZ* genes are similar to the *E(Z)* class of genes (Springer et al., 2002), which are similar to the arabidopsis *Curly Leaf* gene. These three genes are expressed throughout maize plant, although splice variants are differentially located. The homologues of ESC proteins in maize are *ZmFIE1* and *ZmFIE2* (Danilevskaya et al., 2003; Springer et al., 2002), which show different organ expression patterns. *ZmFIE2* is expressed in most plant tissues, whereas *ZmFIE1* is restricted to expression in the ovule and early endosperm, and shows parent of origin expression patterns (Danilevskaya et al., 2003).

5.1.2.4 ATP-dependent remodelling enzymes

The ATP-dependent remodelling enzyme class of genes, are implicated in both gene activation and repression, and they function to modulate the accessibility of DNA by facilitating the translocation of chromatin nucleosomes and chromatin proteins along

its length. These proteins are characterized by a DNA or RNA helicase-type domain, and three subfamilies can be derived on other sequence features (Havas and Whitehouse). The Snf2-like family contain a DNA-dependent ATP-ase domain and a C-terminal bromodomain, this class includes the arabidopsis protein DDM1 (Jeddeloh et al., 1999), mutants in which show de-repression of genes and retrotransposons. The ISWI-like family comprises an ATP-ase domain and a SANT domain. The CHD family is characterized by the presence of chromodomains and DNA binding domains.

5.1.3 Summary of large scale experiments in maize

In spite of the increase in the microarray literature, only a limited number of experiments have been performed in maize using microarrays, and at the time of writing, none at all for the slide set used in this thesis. The first published maize microarray experiment was a comparison between material from 10-14 DAP endosperms and immature ears, when hybridised to arrays comprising EST libraries of the either organ.. The study was predominantly an assessment of the technology, and a comparison with EST libraries but it also revealed that, whereas the majority of ear genes are also expressed in 14 DAP endosperm, only 57% of endosperm genes were expressed in ear tissue (Fernandes et al., 2002). A broad comparison of gene expression between different tissues used suggested that 7% of genes were specific to one organ, whereas 56% of genes were not significantly differentially expressed between any organ (Cho et al., 2002). The same study also gave an insight into the evolutionary origin of the husk tissue of the ear, which displayed an expressional profile more similar to the floral organs than to leaf tissue. Other studies have provided insights into the stress response of maize plants following UV irradiation, water deficit and salt stress (Casati and Walbot, 2003; Wang et al., 2003a; Yu and Setter, 2003).

5.2 Results

5.2.1 Oligomer array design

Microarray hybridisations were performed to monitor gene expression on as close to a global scale as is currently possible. Arrays produced by the Maize Oligonucleotide Project[6] were used to assess global effects of interploidy crosses on gene expression. These arrays comprised 58,000 spots of 70mer oligonucleotides, printed between two glass slides (A and B). These spots constitute around 30,000 individual maize genes, which represents a substantial fraction (50.8%) of the predicted total gene number of 59,000 (Messing et al., 2004). The arrays also contained some other non-genic features, for example printing control spots, blank spots, and positive and negative controls, as listed in Table 8.

Feature	Slide A	Slide B
Total Number of features	32448	32448
'xxxx' (*sic*)	0	4
SSC (blanks)	960	960
Blank	1022	2949
12 Print Controls (10 µm)	480	480
12 Print controls (20 µm)	440	440
10 Alien Controls	92	92
12 Human Negative Controls	144	144
10 Maize Positive Controls	268	295
Other spots (Genes, transposons etc)	29042	27084

Table 8 Features of maize oligomer microarray

The time of isolation of samples after pollination was an important consideration, and 4 DAP was selected as it represents the completion of cellularization (1.2.4 and (Olsen, 2001)). As interploidy endosperms develop, the severity of the phenotype increases, so in order to identify the initial consequences of the interploidy cross, rather than cumulative downstream effects, the decision was taken to use the earliest material practical.

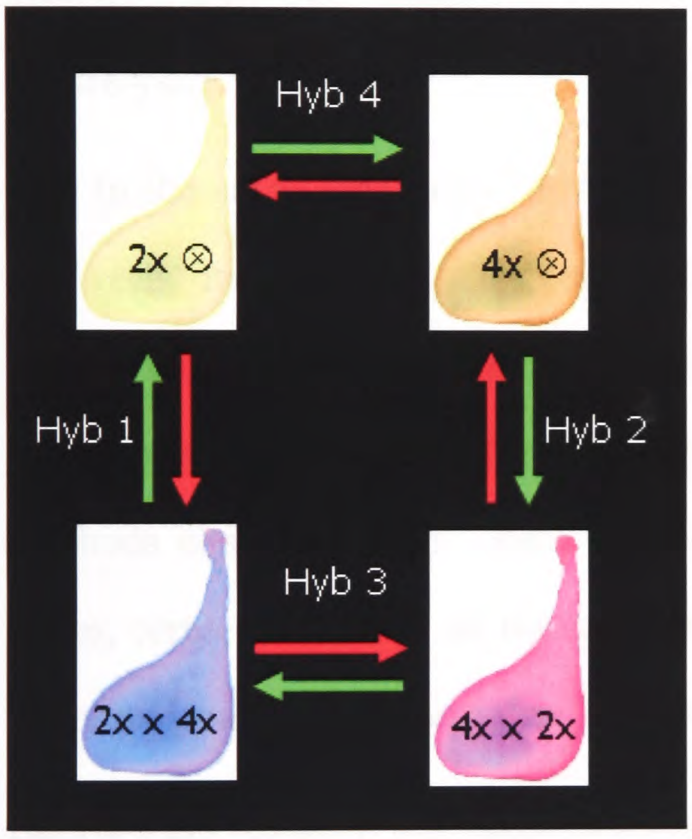
The RNA labelling method used in this study, fluorescent dendrimers (Sub-Chapter 2.7.2), requires a minimum of 2 μg total RNA to be labelled. It was estimated that each 4 DAP endosperm yielded between 50 and 240 ng of RNA, hence it was necessary to isolate numerous endosperms. Mechanical isolation can be performed at most stages of endosperm development, and even unfertilized egg cells can be isolated, but in a technically complex and time-consuming manner (Kranz et al., 1998). Between 1 and 4 DAP the syncitial development phase means that the endosperm is a very fragile structure, which will be easily damaged by isolation. However, by 4 DAP the fully cellular endosperm could be isolated with relative ease, by using a sharp scalpel blade and tweezers to cut through the pericarp and prize the developing endosperm from the nucellus (Figure 5-1). From a biological perspective 4 DAP is also an interesting developmental stage, as having just completed cellularization, it is held that both the ESR and BETL layers have already been specified (Olsen, 2001).

[6] http://www.maizearray.org/



Figure 5-1 Mechanical isolation of 4 DAP endosperm

Demonstration of endosperm isolation. A scalpel is used to initially make an incision through the adgerminal pericarp, and some of the nucellus. Tweezers are used to prize open the two halves of the kernel, to reveal the now partially exposed endosperm, and its impression in one half of the kernel. Tweezers are then used to prize the endosperm at the base, releasing it from the maternal tissue.



	Hyb 1	Hyb 2	Hyb 3	Hyb 4
Slide(s)	30A, 39B	31A, 41B	33A, 43B	32A, 42B
Cy3	2n x 4n *(#31)*	4n x 4n *(#23.1)*	4n x 2n (#65.2)	2n x 2n *(#61.1)*
Cy5	2n x 2n *(#60.1)*	4n x 2n *(#28.1)*	2n x 4n (#30)	4n x 4n *(#24.1)*

Figure 5-2 Design of oligomer microarray experiment.

Logic for the loop-design microarray experiment: arrows represent the flurophore used to label the sample from which it points (green: Cy3, red: Cy5), in the Hybridisation indicated. Lower panel tabulates the design in the top panel, detailing biological sample numbers.

5.2.2 Experimental strategy

Four separate hybridisations were performed in which two differentially labelled RNA samples were directly compared. A loop design was used that allowed the comparison of each direction of cross with two others. Hybridisations were performed to compare both directions of interploidy cross with a balanced cross sharing the same maternal parent (Hyb1: H99-2x × 4x versus H99-2x × -2x, and Hyb2: H99-4x × 2x versus H99-4x × 4x). Another hybridisation compared both interploidy directions with each other (Hyb3: H99-2x × 4x versus H99-4x × 2x) and finally a direct comparison between diploid and tetraploid balanced controls was carried out (Hyb4: H99-2x × -2x versus H99-4x × 4x). The design also incorporated a dye-swap for each sample, as shown in Figure 5-2.

Due to the small amounts of RNA available, it was essential to quantify it accurately and verify its integrity. A Nanodrop spectrophotometer, was used to ensure the OD ratio (OD_{260}/OD_{280}) was close to 2.0, and the RNA quality was assessed with an Agilent Bioanalyzer Nanochip (Figure 5-3). Hybridizations were carried out as detailed in the Methods chapter (2.7.2). Each genome array is physically split between two glass slides, termed A and B, so for each experiment two arrays were processed; the data generated was then processed independently and the results jointly interpreted.

Arrays were scanned, and channel intensities adjusted to ensure (a) a minimal number of saturated spots, and (b) similar intensities in both channels. This was achieved by varying the signal gain, and by using the histogram function in the GenePix 5.1 software to obtain similar channel signal profiles; by ensuring the total signal ratio of 1.0 between channels; and by ensuring individual spots that had a high signal had a ratio close to 1.0 (Leung and Cavalieri, 2003) One set of arrays (A33/B43) in which channels were poorly matched was re-scanned the following day, with no apparent loss of signal.

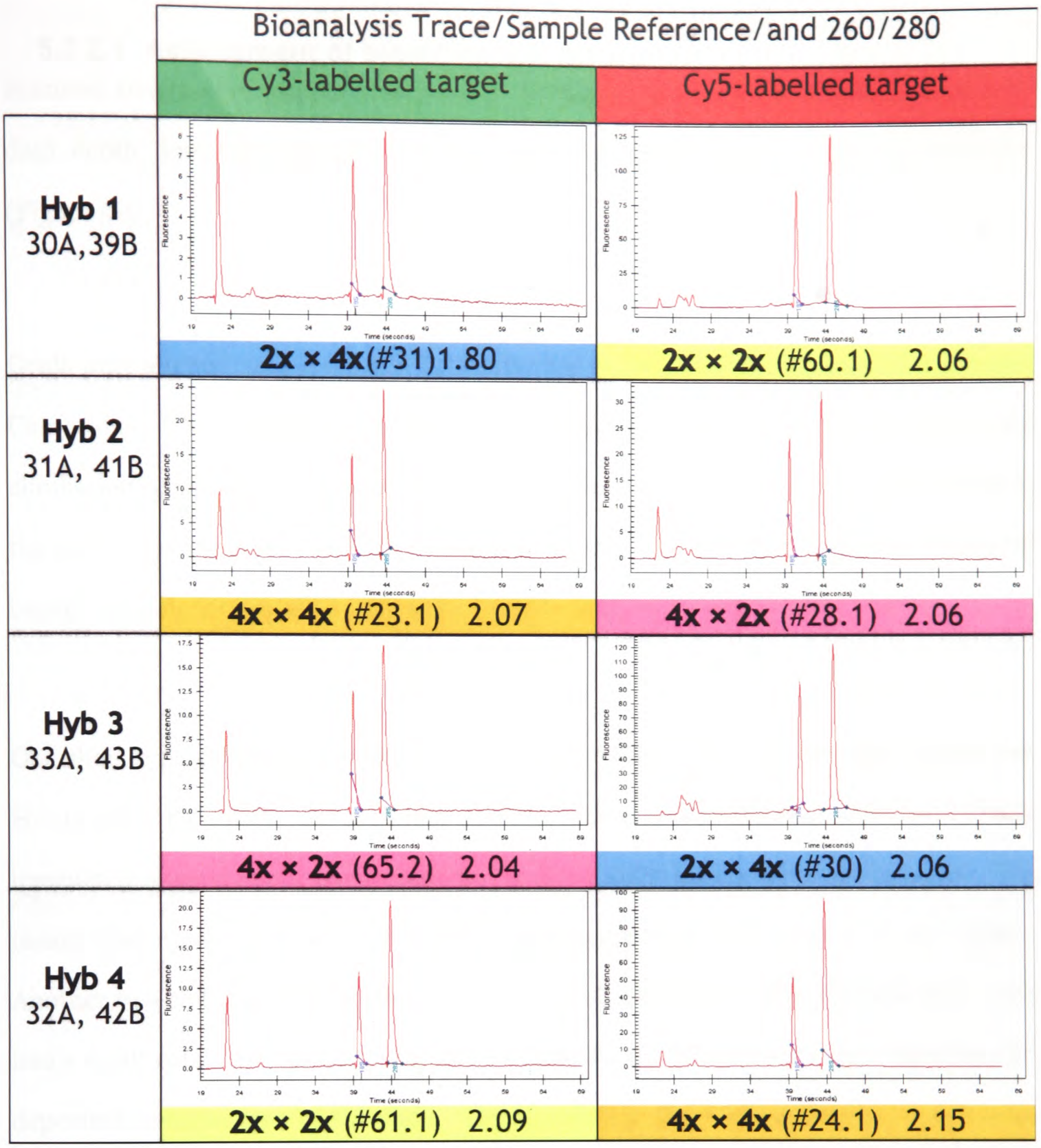


Figure 5-3 Details of RNA samples used in the oligo array experiment

Panels display the identity of the biological sample, and graphs are bioanalyzer trace, confirming the quality of the RNA used in labelling experiments. Numerical values below each trace detail the sample reference number (in parentheses), and the OD260/280 reading for each sample as measured using a nanodrop spectrophotomer.

5.2.2.1 Assessment of hybridisation quality

Scanned arrays were evaluated to be of to be of high quality as assessed by their large data depth, low background intensity, and negligible smears, streaks or scratches (Figure 5-5).

Grids were aligned onto spots and data extracted in the BlueFuse package (BlueGnome, Cambridge, UK). Spots were assessed in detail by eye, and dubious spots were eliminated, generally these were caused by speckles of dust or other artefacts biasing the spot intensity in one or both channels (Figure 5-4A), or by an unusually high local background signal (Figure 5-4B), or by irregularly shaped spots (Figure 5-4C).

One slide in particular was noticed to be of lower quality than the rest: B39 (Figure 5-6, Hyb1, slide B) which, on scanning prior to DNA immobilization, displayed many irregularly-shaped spots, particularly in the lower right-hand corner of the array. This meant that more spots were eliminated in flagging from this array than the others. Another two slides were also affected by a separate problem: slides A33 and B42 each had a small volume of the coupling solution used by the SlideBooster hybridization unit deposited onto an area of the slide. This resulted in these areas having, in one case, very low signal, and in the other instance, high signal in both channels. Affected subgrids were flagged and eliminated from further analysis.

Following visual assessment of the spots, data was 'Fused': a data integration feature of the BlueFuse package, in which the data is filtered (Haan, 2004); in this process single scores are obtained for replicate spots, flagged spots are eliminated, and confidence values are given to all spots. The confidence value, or pON score, is a figure for the likelihood that a spot is expressed which is derived from several measures of confidence

involving Bayesian statistics, including measures such as signal intensity, spot circularity, and offset spot from expected location. Following fusion, intensity plots were plotted as shown in Figure 5-6. Ratio intensity (RI) plots were also produced, which is a modified version of the direct plot of $\log_2$ (signal intensity) from each channel often seen in microarray literature. Instead, the RI plot is a transformation of the data through 45°, and therefore emphasises differences between channels, rather than absolute signal intensity correlation which is very strong and detracts from differentially expressed spots (Yang et al., 2002). These data reveal that the dendrimer system is very effective, as the plots are essentially parallel to the x-axis, with no significant bending or 'banana' effects, as is often the case with direct incorporation labelling methods (Quackenbush, 2002). Hybridization 1 has the greatest spread of data at the lower intensity levels (highlighted in green box in Figure 5-6, Hyb1 slide B), and this seems to be biased slightly towards one channel (C1).

Table 9 summarises the data relating to spots on the arrays before and after fusion, providing an indication of the quality of the arrays, and a relative assessment of statistics between arrays. The total number of spots listed represents the number of unique non-control spots, and the percentage represents the fraction these represent of the total unique non-control probe set. The value retained ranges from between 82.7% (A33) in the worst case (the slide in which several subgrids had to be eliminated for technical reasons), to 98.5% (B42) of the spots.

Plots of spot intensity frequencies are shown in Figure 5-7. This shows that similar intensity distributions are observed between the two channels on each array and between arrays, with greatest differences occurring towards the lower end of expression values. The graphs have a distribution that tails to the right, with a sharp decline at the lower end of intensities, so the underlying distribution clearly does not conform to a

normal distribution, so only non-parametric analysis should be performed on the data. The shape of this distribution means that the mean and median are quite far apart from each other, as detailed in Table 9.

Table 9 also details the ratio of total summed spot intensities which is close to 1.0 on most arrays, but deviates by almost 30% between channels of slide B41, and 11.0% in slide A31. Although summed intensities should be around 1.0 on a well hybridized and channel-matched array, this higher ratio could be due to several different factors, including a higher background on one slide, but it can be overcome by normalization methods so does not present too significant a problem for downstream analysis.

The pON value employed by Bluefuse to assign each spot in each channel a value for the 'likelihood being real' and is an additional valuable control by which to assess the data. In the absence of replicate spots on these arrays, it is essential to use other methods to assess spot quality since statistics such as standard deviation cannot be employed. The data was thus filtered to remove those spots with a pON of less than 0.5 in both channels, (regulated genes may be assessed as having a low quality in one channel, but a high signal and high quality in another). Table 9 shows the absolute number of spots with pON values greater than 0.5 in either or both channels, and as a percentage of spots remaining after manual flagging. Slide A generally performs better under these criteria than slide B, as the former has over 80% of spots retained in all slides except one, whereas, slide B has between 60 and 73% of its spots retained. This value is well within the fraction of genes normally remaining after filtering techniques (Rinn et al., 2003). It is not clear why slide A33 performed so poorly when compared with other, identical slides.

	Slide Number							
	Hyb 1		**Hyb 2**		**Hyb 3**		**Hyb 4**	
	A30		A31		A33		A32	
Number after fusion (% Unique spots)	25681 (88.4%)		27159 (93.5%)		24022 (82.7%)		28266 (97.3%)	
Sample	2x x 4x	2x x 2x	4x x 4x	4x x 2x	2x x 2x	4x x 4x		
Channel	Cy3	Cy5	Cy3	Cy5	Cy3	Cy5	Cy3	Cy5
PMT Gain	640	650	560	630	620	660	610	680
Total Intensity	$4.4x10^7$	$4.25x10^7$	$3.65x10^7$	$3.28x10^7$	$4.42x10^7$	$4.56x10^7$	$6.22x10^7$	$6.05x10^7$
Intensity Ratio	1.04		1.11		0.97		1.03	
Mean	1709	1650	1340	1206	1835	1896	2197	2137
Median	320	353	269	265	325	331	385	404
Spots pON>=0.5	18353	21061	21215	21421	15911	15868	21779	22373
Spots pON >=0.5 both Channels	21287 (82.7%)		21993 (80.8%)		16721 (69.5%)		22975 (81.1%)	
Threshold	570	558	498	465	359	422	427	407
Genes Greater than threshold	9207 (35%)	9257 (36%)	9789 (36%)	9876 (36%)	11307 (47%)	10422 (43%)	13528 (48%)	13644 (48%)

	Slide Number							
	Hyb 1		**Hyb 2**		**Hyb 3**		**Hyb 4**	
	B39		B41		B43		B42	
Sample	2x x 4x	2x x 2x	4x x 4x	4x x 2x	2x x 2x	4x x 4x	4x x 2x	2x x 4x
Number after fusion (% Unique spots)	23086 (85.2%)		26543 (98.0%)		25751 (95.0%)		26677 (98.5%)	
Channel	Cy3	Cy5	Cy3	Cy5	Cy3	Cy5	Cy3	Cy5
PMT Gain	640	650	560	630	600	640	610	680
Total Intensity	$3.53 x10^7$	$3.54 x10^7$	$5.44 x10^7$	$4.24 x10^7$	$7.07 x10^7$	$7.30 x10^7$	$6.13 x10^7$	$6.41 x10^7$
Intensity Ratio	0.99		1.28		0.97		0.96	
Mean	1524	1532	2048	1596	2741	2829	2295	2398
Median	302	359	509	445	745	814	495	548
Spots pON>=0.5	14119	16557	15084	15800	17265	17760	15306	15834
Spots pON >=0.5 both Channels	16826 (72.7%)		16132 (60.7%)		18247 (70.7%)		16297 (61.0%)	
Threshold	309	233	542	457	702	745	711	805
Genes Greater than threshold	11341 (49%)	13949 (60%)	12724 (48%)	12943 (49%)	13269 (52%)	13468 (52%)	10678 (40%)	10585 (40%)

Table 9 Summary Statistics of microarray hybridisations

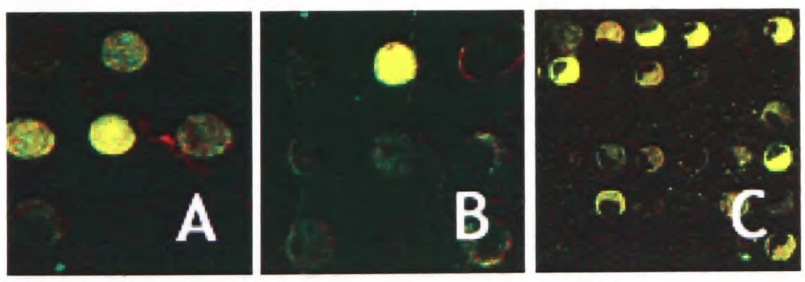


Figure 5-4 Examples of poor quality spots eliminated from the array

A) Example of spots which might have been biased by local speckles, in this case a red speckle on the right hand spot. B) Spots which are within an area of high local background, as in this green-smeared region, were eliminated. C) Some spots were irrgularly shaped which, if severely non-circular, were eliminated.

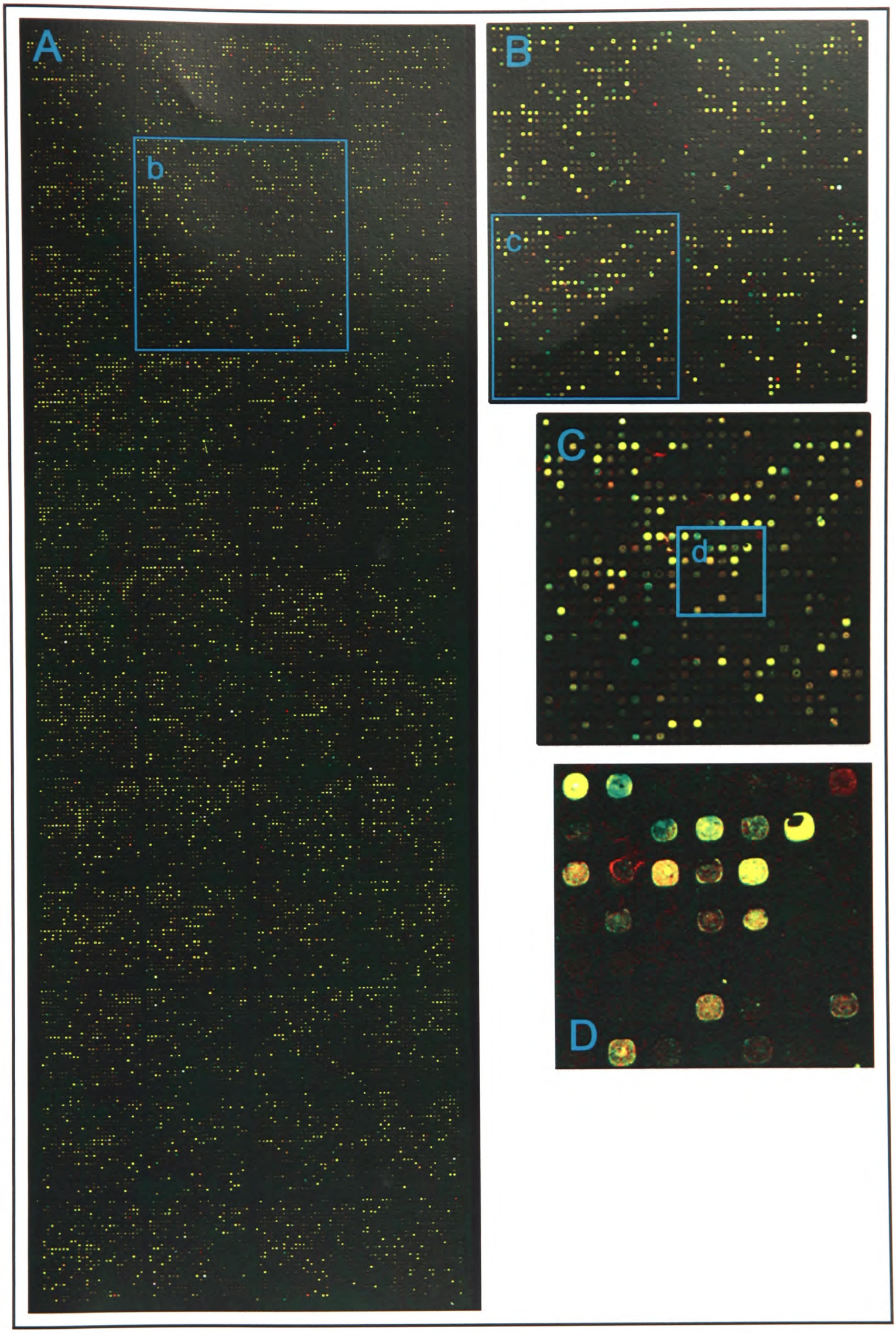


Figure 5-5 Scan of one microarray (Hybridisation 1, slide A39)

One full array (Slide A31), showing both channel 1 (Cy3) and Channel 2 (Cy5) data. Panels show increasing magnifications of the slide, revealing the high clarity, low background and broad data depth of the slide.

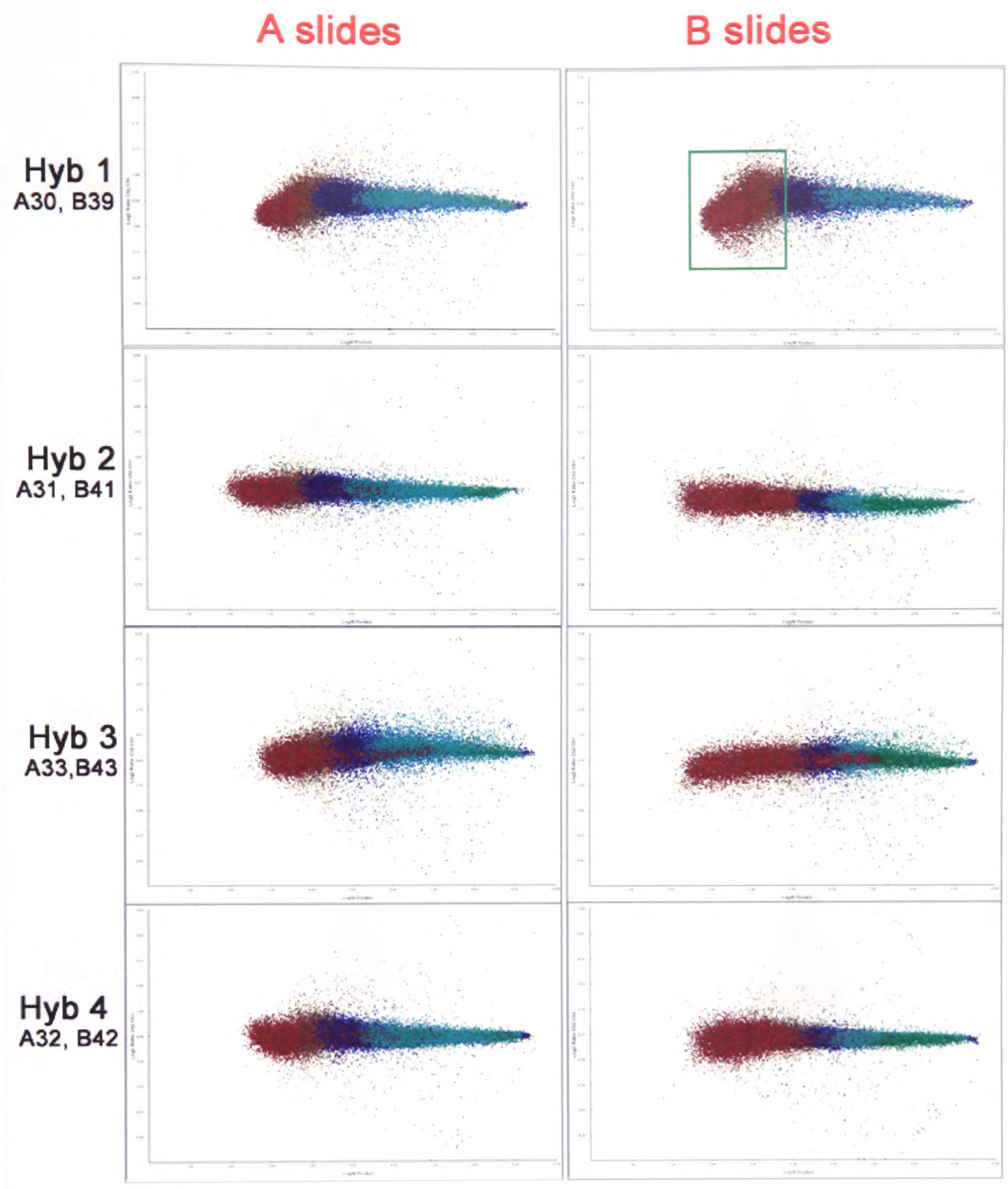


Figure 5-6 Table of RI plots for oligomer array hybridizations

Graphs show the RI transformation of fused data from oligo arrays. As each hybridization was carried out over two slides (A and B), two images are shown for each hybridisation. The x-axis is the $\log_{10}$ product of the signal intensity of both channels (Cy3 and Cy5), and the y-axis is the $\log_2$ ratio of Cy5/Cy3.

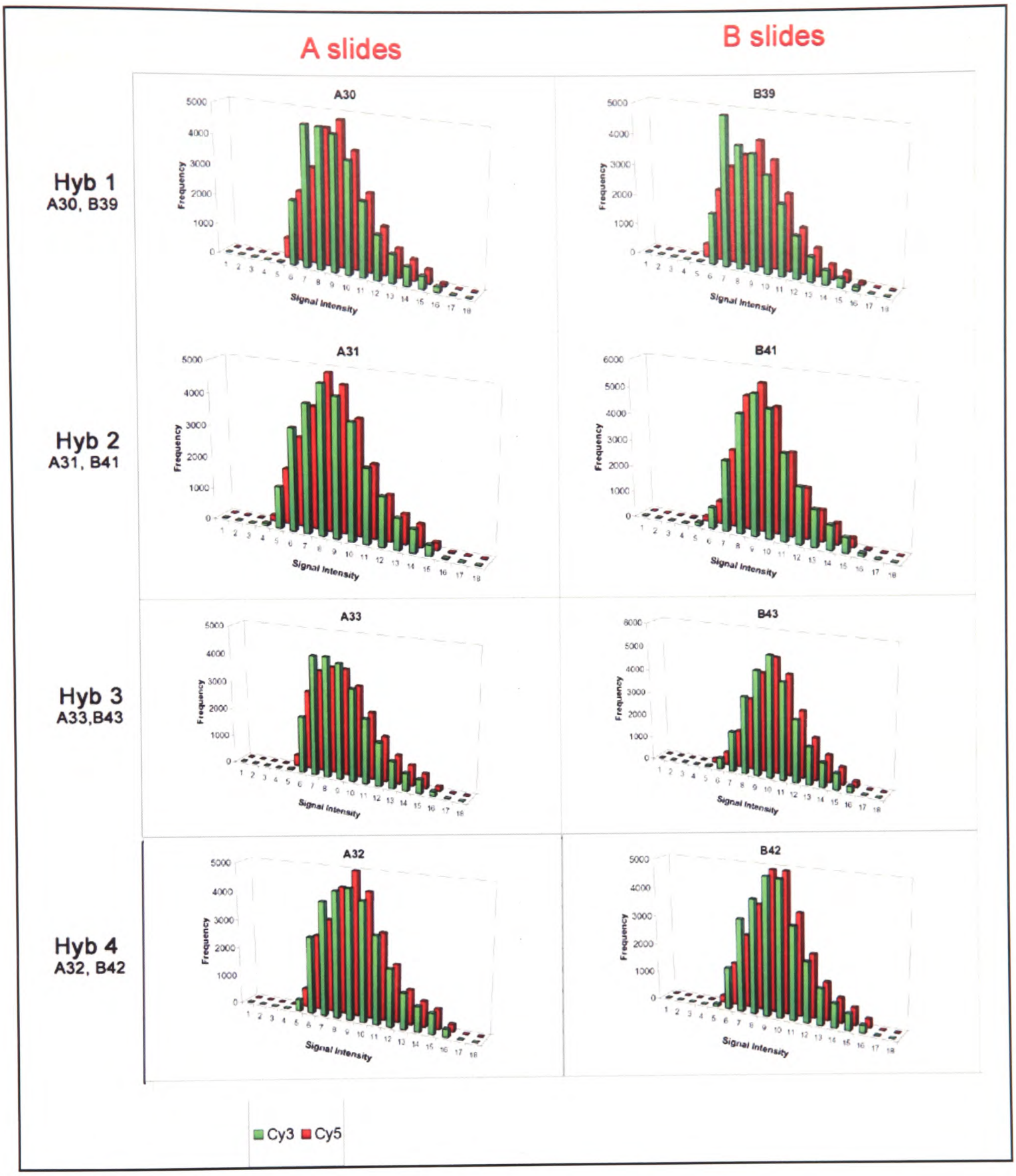


Figure 5-7 Frequencies of spot intensity values

Frequency distributions of spots at different fluorescent intensities (x-axis, $\log_2$ Intensity) are plotted for both channels (Cy3 in green, and Cy5 in red) for both slides A and B in each hybridisation. Frequency of spots in each interval is plotted on the y-axis in a linear scale. The data clearly does not conform to a normal distribution.

5.2.2.2 Dendrimer technology shows sequence specific biases

It was noticed during preliminary investigation of the data that some of the most highly up- and down-regulated sequences in same-slide comparisons were common between distinct arrays, in spite of the experimental design which ensures that sample replicates are labelled with both dyes (Figure 5-2). This pointed to some sort of channel specific effect. Such effects had never previously been noticed in the laboratory where the work was carried out (Dr N. J. Saunders, University of Oxford, personal communication) in spite of many hybridisations having been carried out using this labelling technology. However, the current experiments were the first on such a complex organism as maize, and the array size contained one or more orders of magnitude more features than the ones used during the development of the system. The literature and labelling kit manual suggested that this effect had been reported previously, and had not been satisfactorily eliminated (Revollo et al., 2004). One study suggested that the dendrimer approach produced the fewest signal-specific effects between several experiments compared (Baggerly et al., 2004).

Inspection of the data, and further experiments performed in the laboratory, have shown that the effect appears to be specific to a limited number of genes, and seems to be highly polar – i.e. specific to one channel and showing very high intensity values. Hence, these genes initially appeared to be the most highly regulated genes. Although this is a clear disadvantage of this technology, only a tiny fraction of spots are affected (0.37%) and the advantages of this approach still greatly outweigh this one small disadvantage. Also, the dendrimer system is superior to other signal-enhancing technologies, such as RNA amplification, as it is straight-forward to assess which spots are affected, and eliminate them from further analysis. Since the identification of this effect, pre-hybridization of the dendrimer labels to each new array type is now used as an initial quality control test.

Analysis of sequences showing signal-specific effects did not show any significant similarity either to each other or to the dendrimer capture sequence (Dr N. J. Saunders, personal communication). If, as it is assumed, these effects are highly polar, then it is straight-forward to remove affected spots from downstream analysis, by virtue of the experimental design implemented. In theory, the product of signal intensity ratios of genes across arrays should approximate to 1.0. However, for sequences showing these highly polar effects their products would be very high or very low. Noise in the system means that a range of values would be obtained and a graph of this data is plotted in Figure 5-8. It shows that there is an inflection at the higher and lower margins, at a point of around 4.0 or 0.25 ($\log_2$[product of the cross-hyb ratios]) However, data from samples using the same hybridization system, with slightly more stringent conditions, and on a similarly sized array (arabidopsis, which has a smaller genome), also showed channel specific effects. This was for a direct dye-swap experiment using identical RNA samples, so represents an estimation of the minimum artefactual signals in the system. For this 'control' of the system the $\log_2$ of the maximum cross array product was 13.2 and the minimum was -11.6, and an inflection was seen at 2.5 and -1.5 respectively. 57 and 148 spots exceeded these respective values, so were considered technical artefacts, which represents 0.71% of the total 28,800 spots affected in either channel.

Extrapolation to the maize data-set (across four arrays) predicts maximum channel ratio products of 26.4 and -23.2. The observed values are close, at 24.2 and -25.1 for the A slides, and 25.4 and -24.4 for the B slides. Additionally, the points of inflection are close to the expected values of -3.0, and 5.0, being -4.0 and +4.0 for A slides and +2.5 and -2.5 for B slides. Elimination of spots with values outside this range removes 62 genes with a high C1/C2 cross slide product, and 28 with low C1/C2 ratios in A slides and 63 and 89 for B slides. In total 242 of the total 64,896 spots show this

phenomenon (0.37%) in this data-set. All further downstream analyses have had these spots removed. Seven of these eliminated sequences were selected at random and tested by RT-PCR to confirm they were not differentially expressed between samples. The results suggested that none of them showed differential expression, validating the approach employed to eliminate these artefacts.

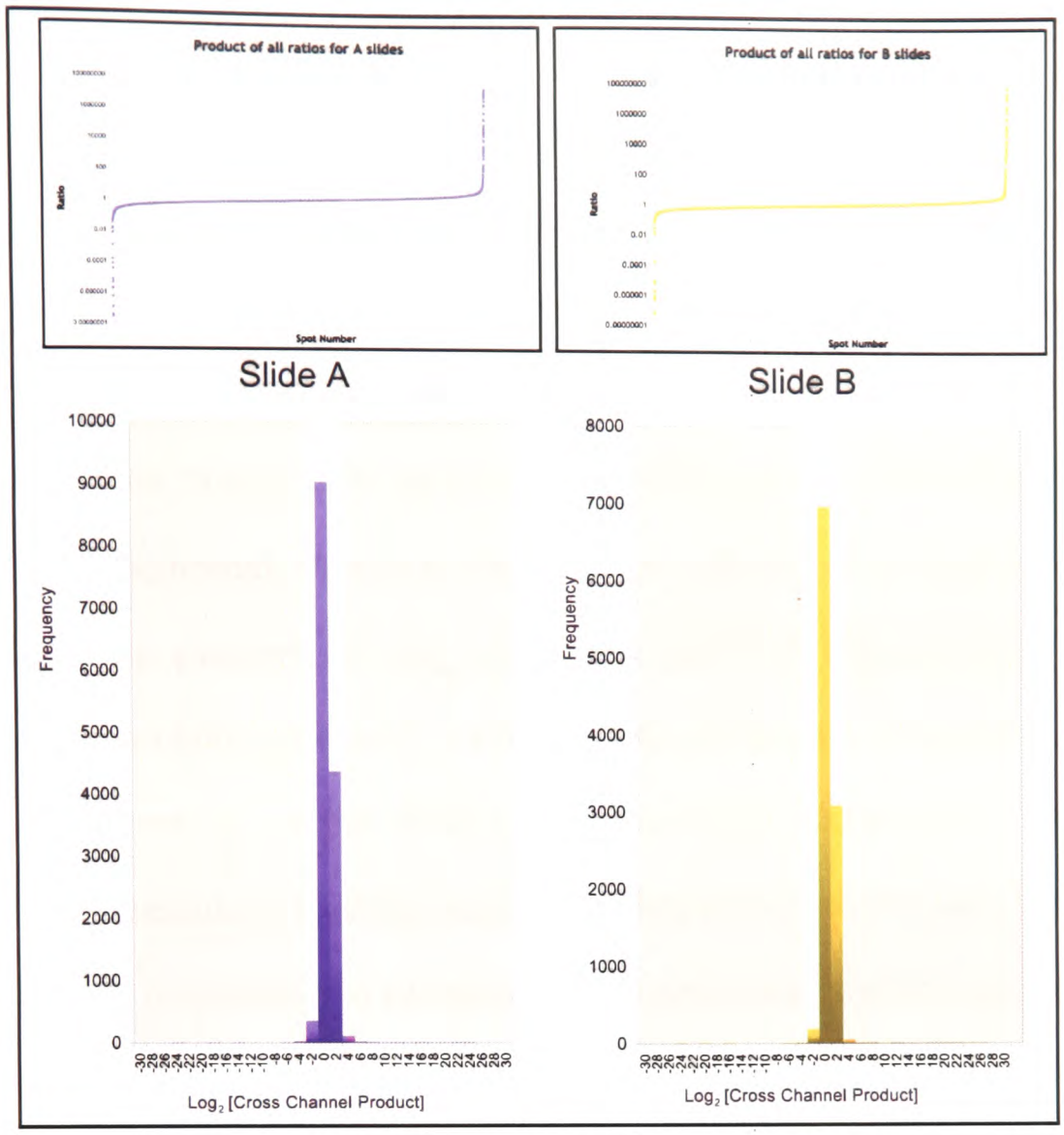


Figure 5-8 Cross-hyb product of gene channel ratios.

For each gene the normalized Cy3:Cy5 ratio was multiplied between all four hybridizations. Cross hyb-products should approximate to 1, but here sequences with polar effects would be very high or very low. All $\log_2$ products were sorted in increasing value, and plotted along the x-axis (top graphs). The majority of genes multiply close to 1, but tend to increase gently at the higher and lower values. Data was plotted as frequency chart, and values eliminated are shaded red.

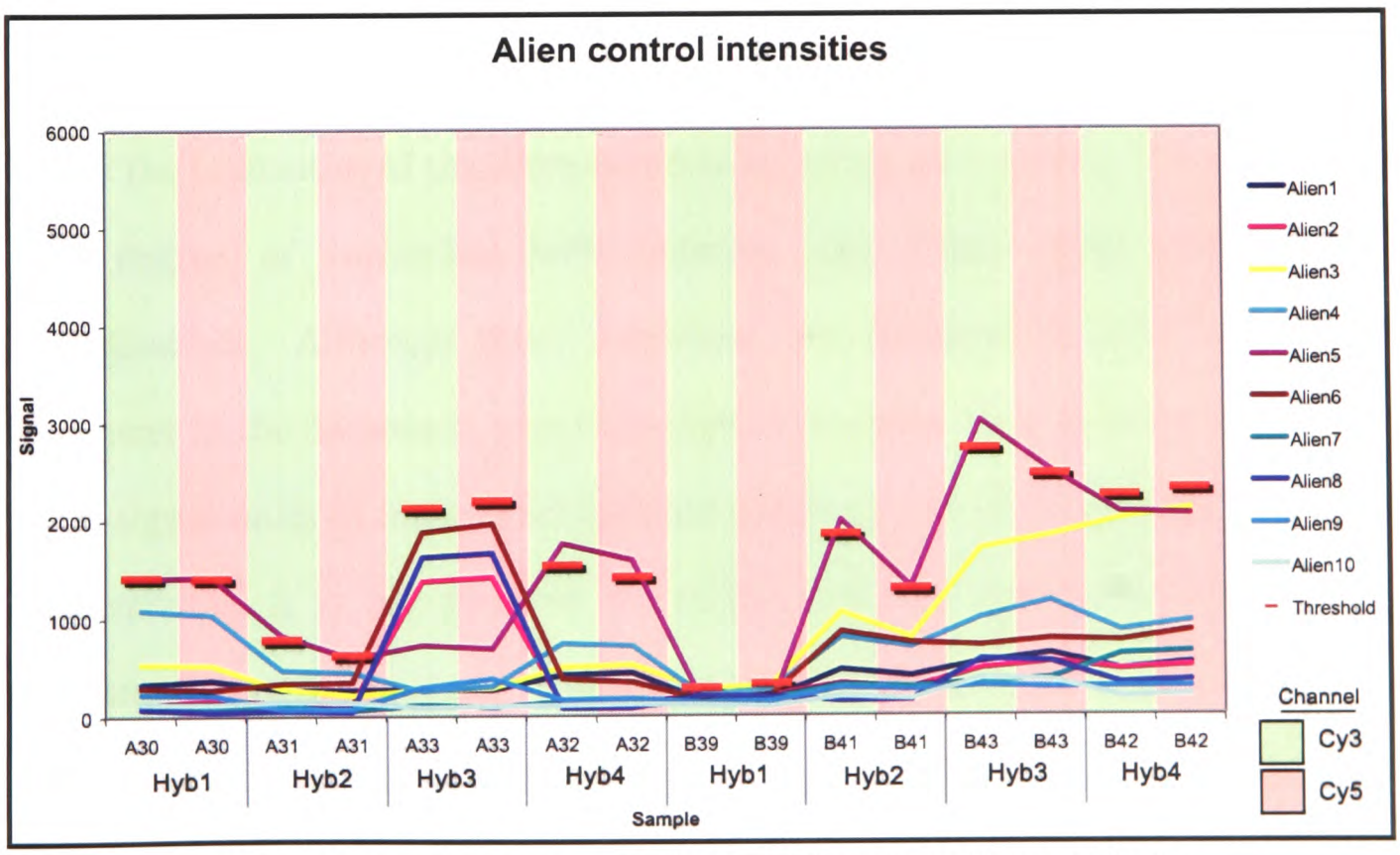


Figure 5-9 Estimation of background thresholds using Alien controls.

Green and red bars of background represent channels 1 and 2 respectively. Fluorescence intensity is displayed on the y-axis, and the median value for each Alien in each array channel is plotted, as is the calculated background threshold value (red bars).

5.2.2.3 Estimation of the number of genes expressed in endosperms

Microarrays always show some degree of noise, which must be accounted for before inferring any trends. Normally a background value is calculated, below which any apparent signal is not considered real. The extent of gene expression will vary from being off (not detectably expressed) to reaching the upper limit of the scanner's detection range. It is important to decide upon a lower value which is considered to be truly expressed, however, noise in the system means this is not a simple task. Spike controls present on the array are useful in this regard. Such spikes represent commercially produced artificial DNA sequences designed to show no specificity to sequences in current databases (Alien DNA, Stratagene). By 'spiking' RNA samples complementary to these controls at known concentrations into a labelling reaction, it should be possible to estimate mRNA gene concentration of other spots and investigate dynamics of the system. These exogenous spikes were not available for these experiments, but they were useful in estimating a background threshold.

Since nothing should hybridize to the Alien DNA features, their measured signal should estimate a threshold intensity below which signal is assumed to be noise. Figure 5-9 shows the intensities of the Alien controls following data fusion. Even these suggest a high degree of variability both between and within alien spots in different hybridization. Although these sequences are designed to have no similarity to sequences in the databases, very high signal is seen for Alien 5: being between 100 and 3000 signal units in most hybridisations channels. As the sequences of the spikes is not available, it is not possible to confirm whether they were homologous to an endogenous maize sequence. The Alien controls showing the least variation between sample hybridizations were selected to determine a theoretical minimum background level for each hybridization channel combination. These were Alien controls 1, 9 and 10, which all show a cross channel standard deviation no greater than 150 units. The

fused value (approximately the median) of these were used as a baseline. For each array-channel pair, all replicates for each of these spots were then assembled and used to calculate the standard deviation. The baseline background figure was then taken to be two standard deviations above the fused value (Hu et al., 2002).

From this, an estimation could be made of the total genes considered to be expressed above background levels (Table 9). These results suggest that between 51% and 56% of maize genes are expressed to detectable levels in the endosperm. Due to the limited number of replicates, it is not possible to say with confidence whether any of the directions of cross result in a greater fraction of expressed genes.

Data was globally normalized using a statistic obtained from a subset of genes. Since data was to be compared as a group in clustering analysis, it was scaled to reduce both label-specific and slide-specific biases, evident from the differing sum total of intensities, and from differences in mean and median values in the data. The total sum intensity was only calculated for those genes which showed a pON value of greater than or equal to 0.5 for each channel in all sets of arrays. Additionally, the 90 and 152 spots which were deemed channel specific in slide sets A and B were removed. A summary of these normalization statistics is presented in Table 10.

Individual data-sets for each array was normalized using the same statistic. The ratio of channel intensities (Cy3/Cy5) was then calculated, and plotted as a histogram (Figure 5-10).

		Hyb 1		Hyb 2		Hyb 3		Hyb 4	
		A30		A31		A33		A32	
	Number of Spots	13899							
	Channel	Cy3	Cy5	Cy3	Cy5	Cy3	Cy5	Cy3	Cy5
Before Normalization	Ratio Intensity	1.01		1.10		0.94		1.03	
	Mean	2646	2615	2141	1935	2878	3048	3638	3527
	Median	686	736	633	595	738	810	995	956
	St Dev	5962	5413	4088	3687	6117	6188	7133	6916
	Normalization Factor	1.37	1.39	1.70	1.88	1.00	1.03	1.19	1.26
After Normalization	Ratio Intensity	1.00							
	Mean	3639							
	Median	943	1025	1076	1119	934	967	995	987
	St Dev	8199	7531	6947	6933	7733	7386	7133	7136

		Hyb 1		Hyb 2		Hyb 3		Hyb 4	
		B39		B41		B43		B42	
	Number of Spots	10215							
	Channel	Cy3	Cy5	Cy3	Cy5	Cy3	Cy5	Cy3	Cy5
Before Normalization	Ratio Intensity	0.97		1.29		0.96		0.94	
	Mean	2965	3031	4222	3268	4222	3268	2965	3031
	Median	965	1083	1593	1300	1593	1300	965	1083
	St Dev	6038	5647	6473	5056	6473	5056	6038	5647
	Normalization Factor	1.91	1.86	1.34	1.73	1.34	1.73	1.91	1.86
After Normalization	Ratio Intensity	1.00							
	Mean	5651							
	Median	1839	2020	2132	2248	2132	2248	1839	2020
	St Dev	11507	10528	8664	8742	8664	8742	11507	10528

Table 10 Normalization statistics of microarray

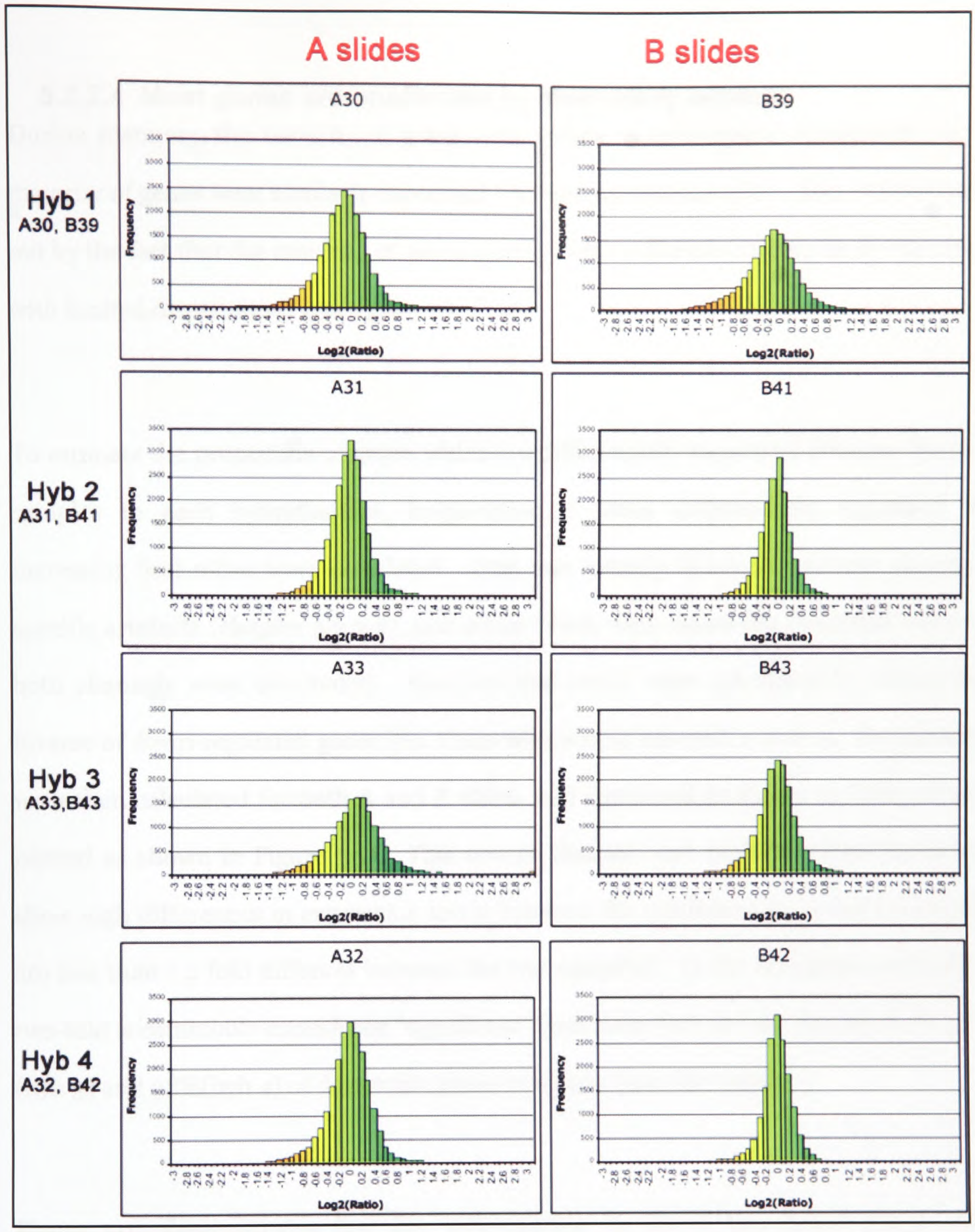


Figure 5-10 Histogram of ratio frequencies

Frequencies of spots within intervals of the $\log_2$ of the ratio signal intensity for the two channels (Cy3/Cy5) are plotted for both slides of each hybridisation.

5.2.2.4 Most genes are unaffected by interploidy crosses

During scanning, the majority of genes were yellow in appearance, suggesting that a majority of genes were similarly expressed between the two samples. This is also borne out by the fact that the majority of genes cluster close to the central axis in the RI plots, with limited dispersion (Figure 5-6).

To estimate the proportion of genes which are differentially expressed between the two samples in each hybridisation, frequencies of genes differentially expressed by increasing fold ratios were calculated. Data was initially filtered to remove channel-specific artefacts (chapter 5.2.2.2), and genes which were below the threshold value in both channels were eliminated. Absolute fold ratios were calculated by taking the inverse of down-regulated genes (i.e. those with a ratio between 0 and 1). Frequencies were then calculated for both A and B slides, and combined as shown in Table 11 and plotted as shown in Figure 5-11. This reveals that the vast majority of genes do not show high differences in expression levels between the conditions observed (over 50% are less than 1.2 fold different between the two samples). In the literature a cut off of two-fold is commonly considered "significant", and these data reveal, that between 93% (hyb 3) and 98%(hyb 4) of detectable genes to be less than this value.

Figure 5-11 demonstrates that the hybridisation in which the fewest genes were differentially regulated between the two samples is hyb 4, which represents the two within-ploidy samples (2x × 2x versus 4x × 4x), and the hybridisation showing the most extreme differences in expression ratios is hyb 3, the direct comparison of the two reciprocal interploidy conditions. Interestingly, of the two comparisons between an interploidy and balanced cross, the comparison of the paternal excess with diploid self cross (hyb 1), shows the greatest difference in expression ratio.

It is interesting to see that there is a reasonable number of genes which show very high differences (above six-fold) in expression between all conditions, although these are lowest in the within-ploidy comparison, and greatest in the direct comparison between interploidies.

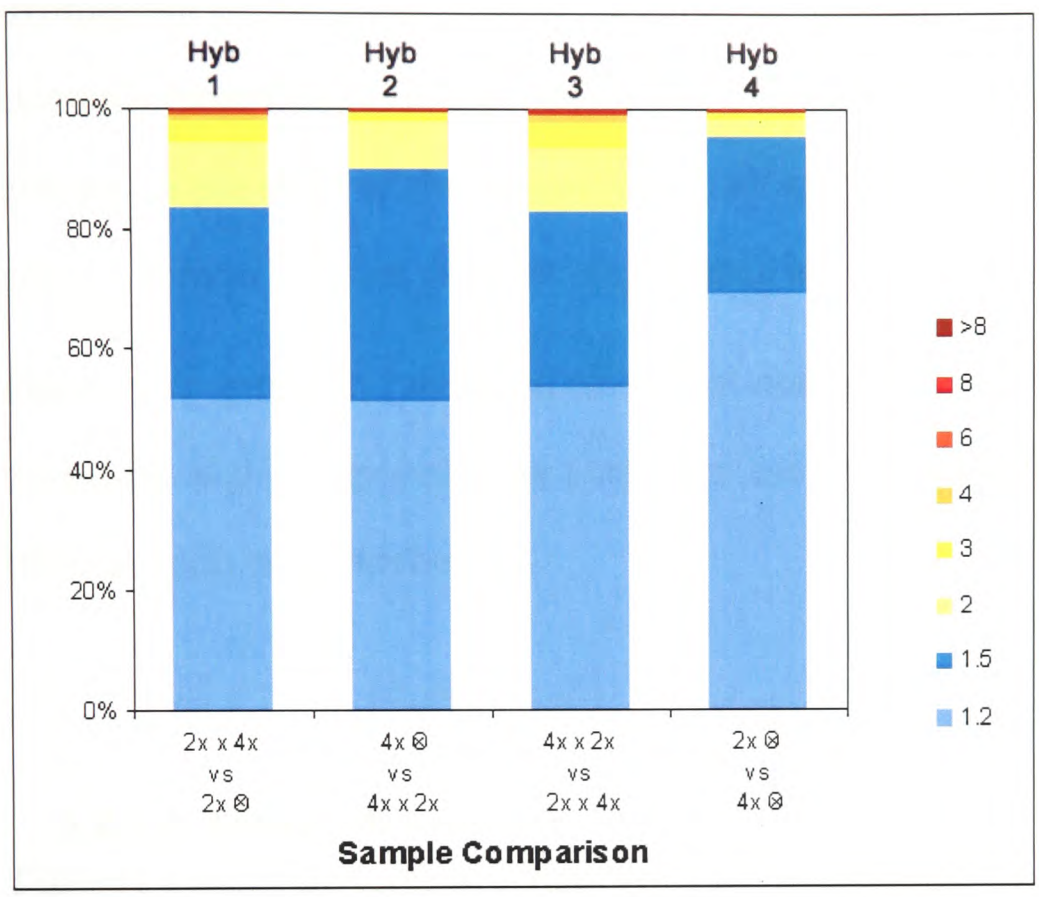


Figure 5-11 Histogram of fold ratio changes

The proportion of genes showing positive or negative fold regulation in the four hybridisations are plotted.

	Sample Comparison							
	Hyb 1		Hyb 2		Hyb 3		Hyb 4	
	A30/B39		A31/B41		A33/B43		A32/B42	
Fold Ratio	2x × 4x vs 2x × 2x		4x × 4x vs 4x × 2x		4x × 2x vs 2x × 4x		2x × 2x vs 4x × 4x	
1.2	12636	52%	12262	51%	13964	54%	17691	69%
1.5	7723	32%	9171	38%	7649	29%	6652	26%
2	2643	11%	1895	8%	2767	11%	853	3%
3	938	4%	284	1%	1030	4%	153	1%
4	204	1%	62	0%	263	1%	45	0%
6	110	0%	29	0%	165	1%	55	0%
8	52	0%	17	0%	63	0%	26	0%
>8	132	1%	115	0%	160	1%	110	0%
Total	**24438**		**23835**		**26061**		**25585**	

Table 11 Frequency table fold-ratio changes between sample

The number of genes expressed differentially between either sample in a dual-colour labelling are shown, alongside the percentage of all spots for that hybridisation.

5.2.2.5 Most highly expressed genes are similar between conditions

The most highly expressed genes were investigated to understand the expressional make-up of the early endosperm. Of the spots with the 10 percent highest signal intensities in each sample, 58% were common to all eight samples investigated. This result suggests a similar expression profile at the higher end of transcriptional activity, irrespective of the genome constitution of the samples. Of these genes 54.8% do not currently have Gene Ontology annotations, so could not be categorized in to functional groups. A pie chart of the proportion of annotated genes which are in the top 10% of genes common between samples is presented in

Figure 5-12 genes and included ribosomal and protein synthesis genes. Cytoskeleton genes are highly expressed, as shown by the large numbers of genes classed under tubulin, actin and cytoskeleton.

5.2.2.6 Genes of interest

Table 12 details the fold induction, up or down, for genes of particular interest. A clear effect on transfer layer development is shown, as the expression of BETL2 (which has several replicates on the array) is only slightly regulated in the paternal excess, but it is substantially down-regulated in the maternal excess, surprisingly Hybridisation 3 suggests a lesser difference in expression between the two interploidy crosses than might be predicted from hybridisation 1 and 2, although this is in part explained by the fact that hybridisation 4 suggests that the gene is expressed considerably more highly in the tetraploid self than in the maternal self. There is an apparent down-regulation of ESR genes in the paternal excess cross, and a slight up-regulation in the maternal excess cross. The third hybridisation does not fully support the expected observations, once again, as there is an apparently greater level of expression in the paternal than maternal cross. This might simply underlie the inherent variability between kernels. Sequences for BETL1 and MEG1 were not represented on the array.

	Fold Induction					
	Hyb 1	Hyb 2	Hyb3	Hyb 4		
	A30B39	A31B41	A33B43	A32B42	Gene Name	Oligo ref
Cy3	2x × 4x *(#31)*	4x × 4x *(#23.1)*	4x × 2x (#65.2)	2x × 2x *(#61.1)*		
Cy5	2x × 2x *(#60.1)*	4x × 2x *(#28.1)*	2x × 4x (#30)	4x × 4x *(#24.1)*		
Cell Domains						
ZmMRP1	1.7	1.5	#N/A	1.5	TC224705	MZ00008631
ZmMRP1	1.2	#N/A	1.8	2.5	TC224705	MZ00027320
BETL2	1.3	8.7	1.6	8.0	CF019574	MZ00010467
BETL2	1.3	5.6	1.2	7.3	CF007169	MZ00039524
BETL2	1.1	11.5	1.4	8.5	CF005419	MZ00010113
BETL2	1.1	7.5	1.2	7.6	TC219118	MZ00043025
BETL3	1.5	3.8	1.0	3.4	TC241503	MZ00028881
BETL4	7.8	1.0	1.7	1.3	TC238834	MZ00026988
ZmEBE-1	3.0	1.0	1.2	1.0	TC246937	MZ00032325
ZmEBE-2	1.9	1.1	1.2	1.1	TC227891	MZ00044461
INCW	#N/A	1.3	1.2	1.0	NP004184	MZ00040576
ESR1g1	13.6	1.3	2.3	1.0	TC240807	MZ00028154
ESR1g1	#N/A	1.4	#N/A	#N/A	CF019541	MZ00054198
ESR2g2	6.3	1.6	#N/A	1.0	TC236878	MZ00000024
ESR2g2	4.8	2.3	3.2	1.1	CF005644	MZ00010116
ESR3g2	6.4	1.2	2.3	1.1	TC226094	MZ00042317
ESR3g2	#N/A	2.8	#N/A	#N/A	CF019462	MZ00010464
Polycomb Group genes						
FIE1	1.8	1.4	1.6	1.2	NP663435	MZ00040727
FIE2	1.3	1.7	3.1	1.2	TC223966	MZ00027027
EZ1 (Enhancer of zeste protein 1)	1.3	1.6	1.2	1.1	TC238921	MZ00016749
EZ2 (Enhancer of zeste protein 2)	1.7	1.1	#N/A	1.3	TC232451	MZ00031682
EZ3	#N/A	#N/A	#N/A	#N/A	TC228512	MZ00020025

Table 12 Table of regulation of genes of interest

Fold changes are tabulated, and are displayed in green when higher in Channel 1 (Cy3-labelled), or red if higher in channel 2 (Cy5 labelled). #N/A represents values that were eliminated during fusion due to low quality spots, or were below background threshold in both channels.

5.2.2.7 Genes most significantly changed in different conditions

Table 13 to Table 16 list those genes which were most significantly regulated between the samples compared. In Hyb1, genes identified as being up-regulated in 2x × 4x crosses include calreticulin, a calcium-dependent kinase, and a putative calmodulin gene, suggesting the importance of calcium signalling in these endosperms. Intriguingly an apospory-associated protein C like gene is up-regulated. Genes which are down-regulated in this cross, include a marker gene for the BETL (BETL4), and a marker of the CSE, a zein gene.

Hyb2 suggests that the maternal excess results in a down-regulation of BETL marker genes BETL2 and BETL3, which are different from those down-regulated in the paternal excess. A number of genes involved in cell wall biosynthesis are up-regulated in these maternal excess crosses (ß-expansin, and β-1,3-glucanase).

The direct comparison of paternal and maternal excess crosses is carried out in Hyb3, and this reveals a large number of genes associated with the CSE as being more highly expressed in the 4x × 2x kernel, such as glutelin and zein genes. Additionally a number of heat shock proteins are expressed higher in maternal excess, as are genes involved in protein biosynthesis (EF5a and 60S ribosomal protein). The paternal excess endosperm displays higher expression of several photosynthesis-related proteins: Cytochrome P450, and rubisco-binding subunits. A number of histidine amino acid transport and histidine phosphotransferase proteins are also expressed greater than in the maternal excess cross. Like in Hyb1, an amino acid transporter is shown to be up-regulated in the paternal excess.

Hyb 4 might reveal those genes which are affected by a recent tetraploidization event. Clearly the formation of the BETL is affected, as BETL1, BETL2 transcripts are more highly expressed in comparison to the diploid cross, as is another basal layer antifungal peptide.

5.2.2.8 Analysis of functional categories of genes regulated between samples

The program ErmineJ[7] was used to group genes that were regulated between samples in each hybridisation into GO categories, in the three ontologies of cellular process, function and compartment, as shown in

Figure 5-12.

These data reveal that the paternal excess cross leads to an up-regulation of genes involved in protein biosynthesis and ribosomal biogenesis and assembly, but a down-regulation in genes involved in protein modification. Interestingly, the maternal excess cross leads to the opposite effect, as a large proportion of genes involved in protein biosynthesis and ribosome biogenesis are down-regulated. Transport genes are also up-regulated in the maternal excess crosses.

5.2.2.9 Analysis of genes over-represented in different conditions

The IGA package (Breitling et al., 2004) was used to look for groups of genes that were statistically over-represented in samples on the basis of GO categorization (Breitling et al., 2004). Several broad common themes are affected by interploidy crosses, such as photosynthesis, genes in this process seeming to be up-regulated in paternal excess and

[7] http://microarray.genomecenter.columbia.edu/ermineJ/

down-regulated in maternal excess (although this is not supported by the data for Hyb 3 which suggests the opposite).

The program BiNGO (Maere et al., 2005) was also used to investigate functional classes which were over-represented in any sample and works on a different principle to EASE, by considering the DAG relationships of GO terms. The GO classes assessed as significantly altered between conditions using this method are tabulated in Table 18.

Up-regulated in 2x × 4x Hyb 1 - A30/B39 Up in C1		
Gene description	**Oligo ID**	**Fold Change**
r40c1 protein	MZ00007057	11.4
fibrillin-like protein	MZ00024841	7.0
calreticulin precursor	MZ00012607	6.0
NA	MZ00045503	5.6
NA	MZ00055451	5.6
calcium-dependent protein kinase 3	MZ00020905	5.6
hypothetical protein	MZ00024769	5.6
NA	MZ00043765	5.5
Profilin 4 (ZmPRO4)	MZ00012721	5.4
NA	MZ00006524	5.3
amino acid selective channel protein	MZ00017395	5.3
unnamed protein product	MZ00017940	5.2
apospory-associated protein C-like protein	MZ00026019	5.1
NA	MZ00006591	5.1
nectarin 5	MZ00020469	4.9
thioredoxin h isoform 1	MZ00023641	4.9
NA	MZ00036975	4.9
cinnamic acid 4-hydroxylase	MZ00025513	4.9
putative calmodulin	MZ00031310	4.8
NA	MZ00037141	4.6
NA	MZ00008533	4.5
ASR2	MZ00041306	4.5
Chalcone synthase WHP1	MZ00040585	4.2
OJ991214_12.11	MZ00015921	4.2
OSJNBa0053K19.11	MZ00018077	4.1
NA	MZ00043634	4.1
NA	MZ00018954	4.1
hypothetical protein	MZ00012423	4.0
NA	MZ00033212	4.0

Down-regulated in 2x × 4x Hyb 1 - A30/B39 Up C2		
Gene description	**Oligo ID**	**Fold Change**
zein	MZ00037565	8.9
BETL4	MZ00026988	7.8
hypothetical protein	MZ00015160	7.7
serine protease	MZ00022574	7.7
unknown protein	MZ00031634	6.8
putative protein kinase	MZ00012285	6.7
naringenin 3-dioxygenase	MZ00044256	6.6
unknown protein	MZ00010416	6.5
OSJNBb0065L13.5	MZ00003122	6.4
NA	MZ00032540	6.4
putative serine protease	MZ00007338	6.1
unknown protein	MZ00017062	5.3
OSJNBa0019K04.16	MZ00016177	5.2
NA	MZ00006304	5.2
NA	MZ00042935	5.1
NA	MZ00039489	5.0
putative xyloglucan endo-transglycosylase	MZ00024062	4.9
ribosomal protein L12	MZ00036407	4.9
OSJNBa0013K16.5	MZ00031500	4.8
LIM transcription factor	MZ00036188	4.7
NA	MZ00039702	4.6
NA	MZ00029893	4.6
NA	MZ00037085	4.4
NA	MZ00007710	4.3
hypothetical protein	MZ00014890	4.1
NA	MZ00043963	4.0
NA	MZ00010156	4.0

Table 13 Most highly regulated genes in 2x × 4x crosses

Down-regulated in 4x × 2x Hyb 2 - A31/B41 C1 Up		
Gene description	**Oligo ID**	**Fold Change**
NA	MZ00039700	16.8
BETL2	MZ00010113	11.5
NA	MZ00039509	10.1
NA	MZ00016135	9.2
BETL2	MZ00010467	8.7
NA	MZ00033167	7.9
BETL2	MZ00043025	7.5
NA	MZ00044201	5.9
BETL2 protein	MZ00039524	5.6
NA	MZ00039726	5.0
hypothetical	MZ00020339	4.4
beta-expansin 7	MZ00024970	4.1
NA	MZ00056301	4.0
NA	MZ00024909	4.0
NA	MZ00008533	3.9
glucosyltransferase NTGT2	MZ00012836	3.9
putative gag-pol polyprotein	MZ00032547	3.9
NA	MZ00052655	3.8
BETL3 protein	MZ00028881	3.8
Nonspecific lipid-transfer protein precursor (LTP)	MZ00041613	3.6
hypothetical protein	MZ00023020	3.6
ae3 protein	MZ00033327	3.5
NA	MZ00000359	3.2
Pyruvate decarboxylase isozyme 1	MZ00026798	3.2
NA	MZ00031135	3.2
NA	MZ00000306	3.2
NA	MZ00039063	3.2
OSJNBa0088K19.19	MZ00031065	3.2
expressed protein	MZ00031096	3.1
putative stearoyl-acyl carrier protein desaturase	MZ00035674	3.0
NA	MZ00052825	3.0

Up-regulated in 4x × 2x Hyb 2 - A31/B41 C2 Up		
Gene description	**Oligo ID**	**Fold Change**
At1g76150/T23E18_38	MZ00015088	4.6
unknown protein	MZ00018036	4.2
seven transmembrane protein Mlo8	MZ00045219	4.0
putative alpha-galactosidase preproprotein	MZ00026491	3.6
NA	MZ00035398	3.4
beta-expansin 7	MZ00024972	3.4
ABA- and ripening-inducible-like protein	MZ00036717	3.3
unnamed protein product	MZ00009959	3.3
beta-1,3-glucanase	MZ00002271	3.1
uncleaved legumin-1	MZ00038091	3.1
hypothetical protein	MZ00015160	3.0
NA	MZ00010165	3.0
ESTs AU033170	MZ00026627	3.0
NA	MZ00039489	2.9
NA	MZ00021974	2.9
putative guanine nucleotide-binding protein	MZ00007821	2.9
allene oxide synthase	MZ00044190	2.8
NA	MZ00039644	2.8

Table 14 Most highly regulated genes in 4x × 2x crosses

Expressed Greater in 4x x2x than 2x x 4x Hyb 3 - A33/B43 C1 Up		
Gene description	Oligo ID	Fold Change
unknown protein	MZ00006898	15.6
NA	MZ00009613	15.5
Eukaryotic translation initiation factor 5A	MZ00023425	14.7
OSJNBa0072K14.4	MZ00041464	13.8
putative Myo-inositol-1-phosphate synthase	MZ00037687	13.6
Glutelin 2 precursor (Zein-gamma)	MZ00037963	12.7
putative GTP-binding protein	MZ00051033	12.4
NA	MZ00037616	11.9
Glutelin 2 precursor	MZ00023426	11.7
NA	MZ00013595	11.0
ESTs C98280	MZ00013596	10.5
17.5 kDa class II heat shock protein	MZ00044655	10.1
Glutelin 2 precursor	MZ00037969	9.9
unknown protein	MZ00029009	9.1
heat shock protein 18	MZ00032788	9.1
RTNLB32	MZ00014116	9.0
NA	MZ00008710	8.7
10K zein precursor	MZ00038140	8.7
ribosomal Pr 117	MZ00039697	8.6
NA	MZ00054075	8.6
putative snRNP splicing factor - related	MZ00010031	7.8
Glutelin 2 precursor	MZ00038046	7.7
At1g76150/T23E18_38	MZ00015088	7.7
putative steroid membrane binding protein	MZ00014094	7.6
60S ribosomal protein L18a	MZ00011459	7.5
small heat-shock protein class I	MZ00017449	6.3
NA	MZ00001209	6.2
NA	MZ00040100	6.2
hypothetical protein	MZ00019583	6.2
low-molecular-weight heat-shock protein	MZ00013234	6.1
NA	MZ00021974	6.0
contains EST AU07067	MZ00031906	6.0
OSJNBa0013K16.6	MZ00014872	6.0
putative GDSL-motif lipase/hydrolase protein	MZ00017058	5.9
putative HGA1	MZ00013397	5.9
hypothetical protein	MZ00040770	5.7
NA	MZ00011108	5.7
putative cysteine proteinase	MZ00034244	5.7
putative ids-4 protein	MZ00004129	5.5
ESTs C98280	MZ00041098	5.4
OSJNBa0053K19.8	MZ00037592	5.4
NA	MZ00021949	5.1

Expressed Greater in 2x x4x than 4x x 2x Hyb 3 - A33/B43 C2 Up		
Gene description	Oligo ID	Fold Change
hypothetical protein	MZ00020339	11.3
putative Cytochrome P450	MZ00030855	10.3
OSJNBb0020O11.1	MZ00050723	9.6
putative gamma-lyase	MZ00036774	9.2
histidine amino acid transporter	MZ00047030	8.8
NA	MZ00044718	8.7
small heat shock protein Hsp23.6	MZ00029223	8.2
NA	MZ00019462	8.1
putative gamma-adaptin 1	MZ00032390	8.0
NA	MZ00013501	7.9
subtilisin-like protease	MZ00003635	7.6
putative rubisco subunit binding-protein beta subunit	MZ00031970	7.6
NA	MZ00032046	7.5
VrrB	MZ00010135	7.2
putative cytochrome P450	MZ00021198	6.9
NA	MZ00005609	6.9
histidine-containing phosphotransfer protein	MZ00042108	6.8
OSJNBa0033G16.4	MZ00026108	6.6
AU088590	MZ00014873	6.6
NA	MZ00038495	6.6
NA	MZ00008533	6.5
NA	MZ00037128	6.4
NA	MZ00053333	6.4
G-box binding factor 1	MZ00007000	6.2
putative cell death associated protein	MZ00021780	6.0
Thiazole biosynthetic enzyme 1-2, chloroplast precursor.	MZ00035698	5.9
NA	MZ00037141	5.9
hypothetical protein	MZ00012605	5.9
putative glutathione S-transferase	MZ00015071	5.9
putative MADS-box transcription factor CDM51	MZ00032384	5.8
putative transposase	MZ00045365	5.8
histidine-containing phosphotransfer	MZ00016546	5.8
putative chloroplast nucleoid DNA-binding protein	MZ00001706	5.6
putative 40S ribosomal protein	MZ00017067	5.6
NA	MZ00006976	5.6
type IIB calcium ATPase	MZ00006975	5.5
hypothetical protein	MZ00050995	5.5
NA	MZ00037558	5.4
Cytochrome P450	MZ00032324	5.4
NA	MZ00033167	5.4
NA	MZ00005044	5.4

Table 15 Most differentially expressed between 2x × 4x and 4x × 2x crosses

Expressed Higher in 2x × 2x than 4x × 4x Hyb 4 - A332B42 C1 Up		
Gene description	**Oligo ID**	**Fold Change**
hypothetical protein	MZ00055351	10.1
putative cdc21 protein	MZ00043604	6.1
dnaJ-related protein T22F8.50	MZ00026928	5.9
putative speckle-type POZ protein	MZ00027793	5.7
NA	MZ00031280	5.7
hypothetical protein	MZ00008076	4.8
NA	MZ00000098	4.4
NA	MZ00008082	4.1
putative pinhead	MZ00013070	3.9
unknown protein	MZ00015282	3.9
putative cinnamyl alcohol dehydrogenase	MZ00001274	3.5
NA	MZ00037594	3.4
putative potassium transporter	MZ00028347	3.4
acetoacetyl CoA thiolase	MZ00011562	3.3
putative potasium transporter	MZ00040085	3.2
protein phosphatase 2C-like	MZ00041219	3.1
NA	MZ00008168	3.1
hypothetical protein	MZ00016048	3.1
Nitrate reductase	MZ00035929	3.1
OSJNBa0076N16.25	MZ00011668	3.0

Expressed Higher in 4x × 4x than 2x × 2x Hyb 4 - A332B42 C2 Up		
Gene description	**Oligo ID**	**Fold Change**
NA	MZ00039063	9.0
NA	MZ00013595	8.6
BETL2 protein	MZ00010113	8.5
BETL2 protein	MZ00010467	8.0
BETL2 protein	MZ00043025	7.6
gamma-glutamylcysteine synthetase	MZ00048017	7.5
BETL2 protein	MZ00039524	7.3
NA	MZ00039509	7.0
elongation factor EF-2	MZ00007343	6.8
basal layer antifungal peptide	MZ00039970	6.6
ESTs C98280	MZ00013596	5.9
NA	MZ00024909	5.4
NA	MZ00039700	5.3
helicase-like protein	MZ00006394	5.2
NA	MZ00033167	5.0
hypothetical protein BET1	MZ00040763	5.0
OSJNBa0039C07.7	MZ00047935	5.0
auxin-induced protein	MZ00000559	4.5
NA	MZ00044201	4.4
NA	MZ00009810	4.2
NA	MZ00005192	4.2
NA	MZ00016135	4.0
xyloglucan endo-1,4-beta-D-glucanase	MZ00034412	4.0
NA	MZ00001244	4.0
AU088719	MZ00014942	4.0
unknown protein	MZ00056021	3.9

Table 16 Most up and down regulated genes in 2x × 2x and 4x × 4x crosses

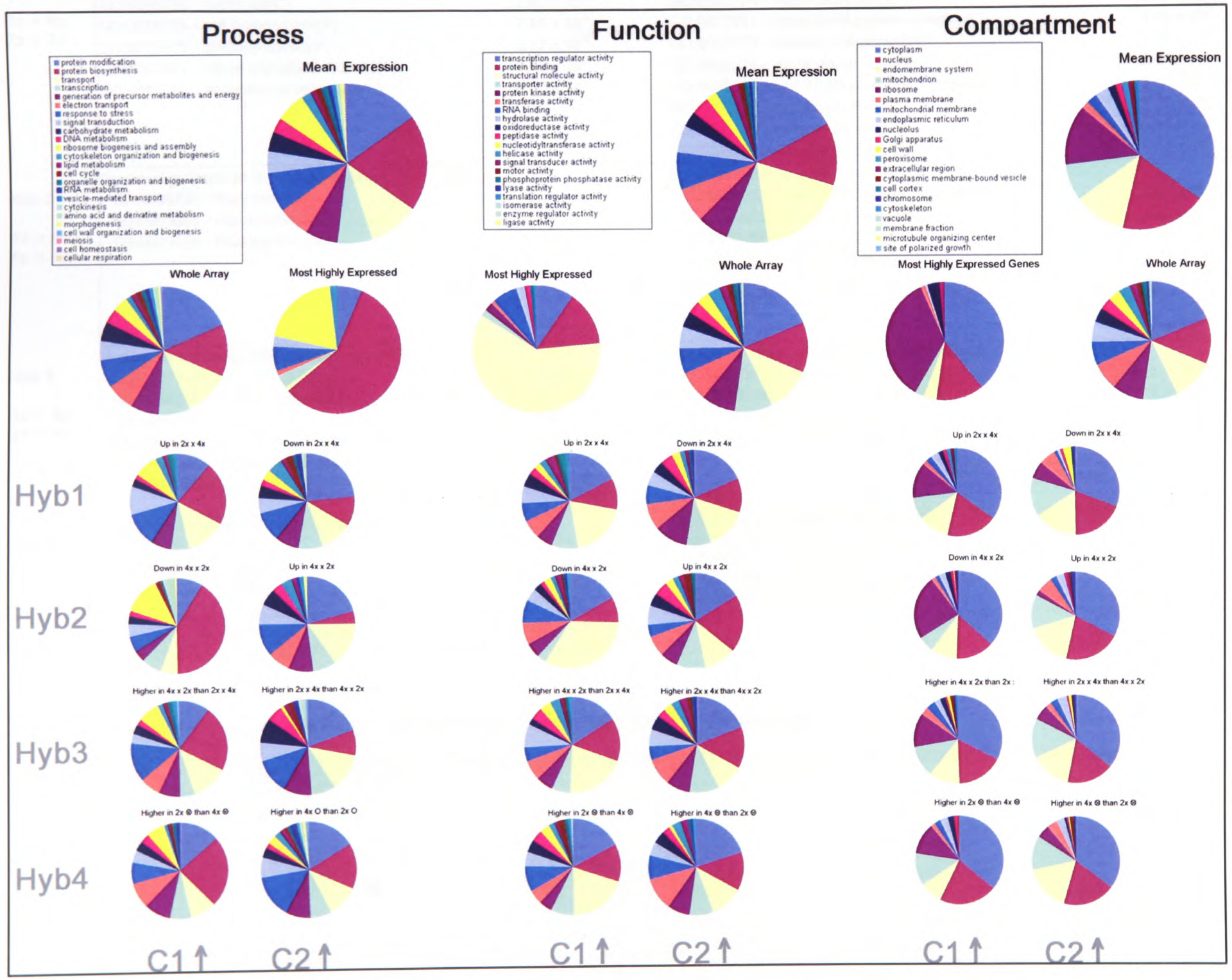


Figure 5-12 Gene Ontology groupings of regulated genes

For each hybridisation, the number of genes which can be categorized in to Gene Ontology categories (based on the three classes of Process, Function, and Compartment) are plotted as pie charts, as a fraction of all genes (top), as a fraction of the whole array, as a fraction of the 10% most highly expressed genes, and as a fraction of genes up- or down-regulated by more than 2 standard deviations of the mean ratio of expression.

Hyb 1	**Up-regulated in paternal excess**		**Down-regulated in paternal excess**	
	GO:0005623 - cell(C)	1.73×10^{-3}	GO:0040007 - growth(P)	2.21×10^{-3}
	GO:0005730 - nucleolus(C)	2.00×10^{-3}	GO:0003774 - motor activity(F)	2.85×10^{-3}
2x × 4x	GO:0019725 - cell homeostasis(P)	2.14×10^{-3}	GO:0005941 - unlocalized protein complex(C)	3.06×10^{-3}
2x × 2x	GO:0015979 - photosynthesis(P)	3.17×10^{-3}	GO:0005575 - cellular_component(C)	3.30×10^{-3}
	GO:0004871 - signal transducer activity(F)	3.46×10^{-3}	GO:0005576 - extracellular region(C)	3.46×10^{-3}
	GO:0005783 - endoplasmic reticulum(C)	4.45×10^{-3}	GO:0030528 - transcription regulator activity(F)	3.49×10^{-3}
			GO:0009579 - thylakoid(C)	3.97×10^{-3}
			GO:0005794 - Golgi apparatus(C)	5.11×10^{-3}
			GO:0005654 - nucleoplasm(C)	5.30×10^{-3}
Hyb 2	**Down-regulated in maternal excess**		**Up-regulated in maternal excess**	
	GO:0005730 - nucleolus(C)	2.21×10^{-3}	GO:0005941 - unlocalized protein complex(C)	2.27×10^{-3}
	GO:0000003 - reproduction(P)	2.38×10^{-3}	GO:0009908 - flower development(P)	2.50×10^{-3}
4x × 4x	GO:0015979 - photosynthesis(P)	3.40×10^{-3}	GO:0005768 - endosome(C)	2.90×10^{-3}
4x × 2x			GO:0003774 - motor activity(F)	3.39×10^{-3}
			GO:0030528 - transcription regulator activity(F)	3.42×10^{-3}
Hyb 3	**Higher in maternal than paternal excess**		**Higher in paternal than maternal excess**	
	GO:0007267 - cell-cell signaling(P)	2.99×10^{-4}	GO:0004871 - signal transducer activity(F)	6.32×10^{-4}
	GO:0030246 - carbohydrate binding(F)	1.04×10^{-3}	GO:0005578 - extracellular matrix (C)	1.96×10^{-3}
4x × 2x	GO:0005730 - nucleolus(C)	1.27×10^{-3}	GO:0009856 - pollination(P)	6.95×10^{-3}
2x × 4x	GO:0019725 - cell homeostasis(P)	2.36×10^{-3}	GO:0009653 - morphogenesis(P)	7.31×10^{-3}
	GO:0005635 - nuclear membrane(C)	2.72×10^{-3}		
	GO:0015979 - photosynthesis(P)	5.51×10^{-3}		
Hyb 4	**Higher in diploids than tetraploids**		**Higher in tetraploids than diploids**	
	GO:0005623 - cell(C)	1.21×10^{-3}	GO:0005623 - cell(C)	1.81×10^{-3}
	GO:0005730 - nucleolus(C)	2.14×10^{-3}	GO:0005730 - nucleolus(C)	1.87×10^{-3}
2x × 2x	GO:0007267 - cell-cell signaling(P)	2.71×10^{-3}	GO:0007267 - cell-cell signaling(P)	2.79×10^{-3}
4x × 4x	GO:0008289 - lipid binding(F)	4.35×10^{-3}	GO:0008289 - lipid binding(F)	4.51×10^{-3}
	GO:0008150 - biological_process(P)	4.64×10^{-3}		
	GO:0003774 - motor activity(F)	$6.88 \times 10^{-}$		

3

Table 17 Summary of GO classes classed as significant from IGA package.

GO categories are displayed which were assessed as significantly overrepresented using the IGA program (p value is displayed). GO numbers are listed, and categorized as P = Cellular process, C = cellular compartment, and F = molecular function.

Hyb 1	**Upregulated in paternal excess**			**Down-regulated in paternal excess**	
	GO:0009059(P)	macromolecule biosynthesis	6.3 E-07	GO:0007017(P) microtubule-based process	5.2E-05
	GO:0006412(P)	protein biosynthesis	6.3 E-07	GO:0007010(P) cytoskeleton organization and biogenesis	0.0004
2x × 4x	GO:0007028(P)	cytoplasm organization and biogenesis	3.9 E-05	GO:0040007(P) growth	0.0006
2x × 2x	GO:0030529(C)	ribonucleoprotein complex	2.2E-08	GO:0045298(C) tubulin	2.7E-11
	GO:0005840(C)	ribosome	6.9E-07	GO:0005874(C) microtubule	4.3 E-10
	GO:0005622(C)	intracellular	1.6E-06	GO:0015630(C) microtubule cytoskeleton	3.7E-09
	GO:0003735(F)	structural constituent of ribosome	6.4E-08	GO:0005200(F) structural constituent of cytoskeleton	6.4E-08
	GO:0005198(F)	structural molecule activity	5.8E-07	GO:0008236(F) serine-type peptidase activity	4.4E-05
Hyb 2	**Downregulated in maternal excess**			**Upregulated in maternal excess**	
	GO:0007046(P)	ribosome biogenesis	<2.2E-016	GO:0009644(P) response to high light intensity	4.0E-06
	GO:0009058(P)	biosynthesis	<2.2E-016	GO:0006511(P) ubiquitin-dependent protein catabolism	7.0E-06
4x × 4x	GO:0009059(P)	macromolecule biosynthesis	<2.2E-016	GO:0005874(C) microtubule	1.7E-04
4x × 2x	GO:0006412(P)	protein biosynthesis	<2.2E-016	GO:0045298(C) tubulin	3.6E-04
	GO:0019538(P)	protein metabolism	<2.2E-016	GO:0005200(F) structural constituent of cytoskeleton	1.5E-08
	GO:0030529(C)	ribonucleoprotein complex	< 2.2E-016	GO:0005515(F) protein binding	6.6E-06
	GO:0005840(C)	ribosome	< 2.2E-016		
Hyb 3	**Higher in maternal than paternal excess**			**Higher in paternal than maternal excess**	
	GO:0009059(P)	macromolecule biosynthesis	9.7E-07	GO:0006519(P) amino acid and derivative metabolism	6.8E-05
	GO:0006412(P)	protein biosynthesis	9.7E-07	GO:0042221(P) response to chemical substance	7.7E-04
4x × 2x	GO:0030529(C)	ribonucleoprotein complex	1.3E-07	GO:0005578(C) extracellular matrix	3.0E-02
2x × 4x	GO:0005622(C)	intracellular	7.1E-07	GO:0009574(C) preprophase band	3.9E-02
	GO:0005198(F)	structural molecule activity	5.9E-08	GO:0003824(F) catalytic activity	1.1E-05
	GO:0003735(F)	structural constituent of ribosome	2.5E-07	GO:0016829(F) lyase activity	1.9E-05
Hyb 4	**Higher in diploids than tetraploids**			**Higher in tetraploids than diploids**	
	GO:0006412(P)	protein biosynthesis	1.7E-09	GO:0008150(P) biological_process	9.8E-05
	GO:0009059(P)	macromolecule biosynthesis	1.7E-09	GO:0009605(P) response to external stimulus	1.2E-04
2x × 2x	GO:0005622(C)	intracellular	4.6E-11	GO:0006950(P) response to stress	2.1E-04
4x × 4x	GO:0005737(C)	cytoplasm	2.8E-09	GO:0005635(C) nuclear membrane	6.1E-05
	GO:0030529(C)	ribonucleoprotein complex	4.3E-07	GO:0005643(C) nuclear pore	1.4E-04
	GO:0003735(F)	structural constituent of ribosome	1.4E-07	GO:0046930(F) pore complex	1.4E-04
	GO:0005198(F)	structural molecule activity	1.1E-06	GO:0003674(F) molecular_function	2.9E-09
	GO:0003676(F)	nucleic acid binding	1.7E-06	GO:0003773(F) heat shock protein activity	2.6E-06

Table 18 Significantly different gene classes from BINGO analysis

GO categories are displayed which were assessed as significantly overrepresented using the IGA program (p value is displayed). GO numbers are listed, and categorized as P = Cellular process, C = cellular compartment, and F = molecular function.

5.2.2.10 Correlations of expression patterns between samples

In order to assess the similarity of expression profiles on a general scale, it is useful to look at the correlations between datapoints within and between slides. One approach to this is to find the R correlation statistic between slides, which is an appropriate method for data with a normal distribution. However, as discussed previously microarray data does not have such a distribution, and for that reason non-parametric means are more suited for looking at correlations. A non-parametric measure was chosen to assess levels of correlation: the Spearman's Rank correlation co-efficient (Kuo et al., 2002) (Figure 5-13).

These statistics reveal that all sample slides have very strong correlation co-efficients within slides, between the two flurophores, irrespective of the biological material being compared, being greater than 0.97 in all cases for the Spearman's Rank statistic. The highest correlation is that between the two within ploidy comparison (hyb 3, 2x × 2x versus 4x × 4x), and least between the two interploidies (hyb4, 2x × 4x versus 4x × 2x).

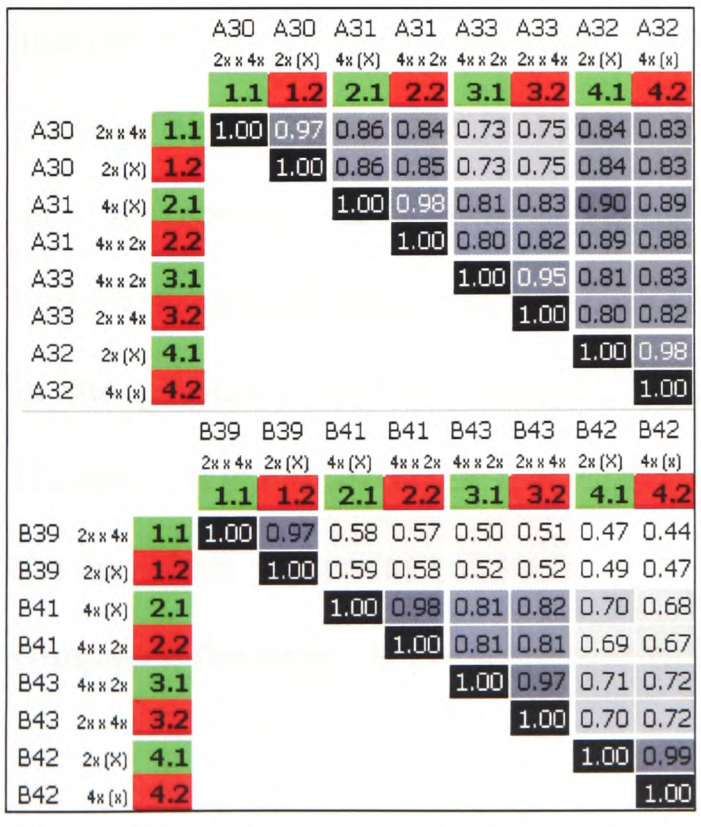


Figure 5-13 Spearman's Rank Correlations between slides

5.2.2.11 The number of genes regulated in the same channel between different comparisons

Due to the biological variability, slight differences in the samples used, and intrinsic uncertainty in microarray techniques, (particularly in the case of limited replication), there must be degree of uncertainty in the data obtained. It would be instructive from both a biological point of view, but also to provide more confidence in the data obtained, to see which genes that are up- or down-regulated in any comparison with those in others. For this reason, a program was written in PERL to compare the genes up-regulated between samples. Its development is detailed in section 2.9.3. Comparisons were made for genes regulated more than 2σ from the local mean, and those changed more than 1σ. In both comparisons, data for those spots which had a high cross- slide product were eliminated.

Generally the number of genes up- or down-regulated is similar between all slides, ranging from around 2,300 to 2,700. However, only a relatively small proportion of genes show common direction of regulation in any sample comparison. At best, in the analysis of all genes greater than 1σ, 22% of genes were in common: between those showing higher expression in 4x × 2x than 4x × 4x and those more highly expressed in 4x × 2x over 2x × 4x . It is reassuring that those genes being up-regulated are in the maternal excess in both cases, as they may represent genes expressed specifically highly in the presence of more maternal genomes or a higher maternal to paternal ratio. However, it is not too surprising that higher numbers of common genes are not seen, as the numbers represent ratios between two biological samples, and no hybridisation compared the same two biological samples.

	Hyb1 C1 Up A30/B39	Hyb2 C1 Up A31/B41	Hyb3 C1 Up A33/B43	Hyb4 C1 Up A32/B42	Hyb1 C1 Down A30/B39	Hyb2 C1 Down A31/B41	Hyb3 C1 Down A33/B43	Hyb4 C1 Down A32/B42	Greater	Lesser
Hyb1 C1 Up A30/B39	2541	509	280	413	0	224	418	279	2x x 4x	2x x 2x
Hyb2 C1 Up A31/B41		2599	297	401	212	0	560	338	4x x 4x	4x x 2x
Hyb3 C1 Up A33/B43			2298	312	281	356	0	330	4x x 2x	2x x 4x
Hyb4 C1 Up A32/B42				2774	250	296	364	0	2x x 2x	4x x 4x
Hyb1 C1 Down A30/B39					2600	408	319	403	2x x 2x	2x x 4x
Hyb2 C1 Down A31/B41						2535	300	364	4x x 2x	4x x 4x
Hyb3 C1 Down A33/B43							2658	331	2x x 4x	4x x 2x
Hyb4 C1 Down A32/B42								2731	4x x 4x	2x x 2x
Greater	2x x 4x	4x x 4x	4x x 2x	2x x 2x	2x x 2x	4x x 2x	2x x 4x	4x x 4x		
Lower	2x x 2x	4x x 2x	2x x 4x	4x x 4x	2x x 4x	4x x 4x	4x x 2x	2x x 2x		

Figure 5-14 Number of genes regulated more than 1σ in common between samples

Table displays the number of genes shared in common between lists of genes expressed more than 1 standard deviation greater than the local mean. Both channels of each hybridization were compared. Squares coloured black represent the number of genes regulated greater than 1sd for that sample. Green shaded squares represent genes common showing higher expression in channel 1 than channel 2, and red squares are higher in channel two than three. Yellow squares show genes in common between different channels.

	Hyb1 C1 Up A30/B39	Hyb2 C1 Up A31/B41	Hyb3 C1 Up A33/B43	Hyb4 C1 Up A32/B42	Hyb1 C1 Down A30/B39	Hyb2 C1 Down A31/B41	Hyb3 C1 Down A33/B43	Hyb4 C1 Down A32/B42	Greater	Lesser
Hyb1 C1 Up A30/B39	509	48	19	16	0	15	45	28	2x x 4x	2x x 2x
Hyb2 C1 Up A31/B41		536	25	16	18	0	82	64	4x x 4x	4x x 2x
Hyb3 C1 Up A33/B43			448	16	28	39	0	27	4x x 2x	2x x 4x
Hyb4 C1 Up A32/B42				371	15	16	19	0	2x x 2x	4x x 4x
Hyb1 C1 Down A30/B39					396	42	29	21	2x x 2x	2x x 4x
Hyb2 C1 Down A31/B41						406	32	13	4x x 2x	4x x 4x
Hyb3 C1 Down A33/B43							624	22	2x x 4x	4x x 2x
Hyb4 C1 Down A32/B42								485	4x x 4x	2x x 2x
Greater	2x x 4x	4x x 4x	4x x 2x	2x x 2x	2x x 2x	4x x 2x	2x x 4x	4x x 4x		
Lower	2x x 2x	4x x 2x	2x x 4x	4x x 4x	2x x 4x	4x x 4x	4x x 2x	2x x 2x		

Figure 5-15 Number of genes regulated more than 2σ in common between samples.

As table 5-14 above, but for genes regulated greater than 2 standard deviations.

		Regulated greater than 1 SD								Regulated greater than 2 SD							
		Hyb1 C1 Up A30/B39	Hyb2 C1 Up A31/B41	Hyb3 C1 Up A33/B43	Hyb4 C1 Up A32/B42	Hyb1 C1 Down A30/B39	Hyb2 C1 Down A31/B41	Hyb3 C1 Down A33/B43	Hyb4 C1 Down A32/B42	Hyb1 C1 Up A30/B39	Hyb2 C1 Up A31/B41	Hyb3 C1 Up A33/B43	Hyb4 C1 Up A32/B42	Hyb1 C1 Down A30/B39	Hyb2 C1 Down A31/B41	Hyb3 C1 Down A33/B43	Hyb4 C1 Down A32/B42
A30/B39	Hyb1 C1 Up	100	20	11	16	0	9	16	11	100	9	4	3	0	3	9	6
A31/B41	Hyb2 C1 Up		100	11	15	8	0	22	13		100	5	3	3	0	15	12
A33/B43	Hyb3 C1 Up			100	14	12	15	0	14			100	4	6	9	0	6
A32/B42	Hyb4 C1 Up				100	9	11	13	0				100	4	4	5	0
A30/B39	Hyb1 C1 Down					100	16	12	16					100	11	7	5
A31/B41	Hyb2 C1 Down						100	12	14						100	8	3
A33/B43	Hyb3 C1 Down							100	12							100	4
A32/B42	Hyb4 C1 Down								100								100

Figure 5-16 Summary of percentage of genes in common between different samples.

5.2.3 Development of a custom micro-array

The community produced oligo arrays allowed a survey of a very large representation of the mRNA population of the cell, however, a more focused survey of gene expression was desirable for a subset of genes hypothesised to be important in endosperm development and development in general. For this reason a custom micro-array was developed, detailed in section 2.6.

The genes placed on the array are summarized in Table 19. These include chromatin remodelling factors of different subclasses (described in chapter 5.1.2). Such genes are important higher-order regulators of transcriptional dynamics, and it is assumed that significant changes occur in the general transcriptional program following interploidy crosses. Additionally, a range of retrotransposons are included, as the hybridisation effects of interploidy crosses may act as a genomic shock, which could potentially activate silenced retrotransposons and transposons (Bennetzen, 1996; Madlung and Comai, 2004). Additionally, a number of endosperm marker genes are included to determine whether the crosses distort the normal specification and development of endosperm marker genes.

Several comparisons were carried out, as listed in Table 20, initially using leaf and tissue total RNA to validate the results, and method. Hybridisation 2 included total RNA from 10 DAP endosperm, and all other hybridisations (numbers 3-10, hybridisation 4 failed), used amplified RNA from 4 DAP endosperms. Amplified RNA was run on a denaturing agarose gel, and through a bioanalyzer to ensure quality (Figure 5-17). Arrays were scanned using an Affymetrix 427 dual laser scanner, and then data was extracted using BlueFuse.

Due to the small number of spots on the array, and the bias towards genes which might be regulated between samples, most normalization protocols were inappropriate as they assume a generally balanced level of total transcription between samples. Instead, after removal of spots with pON values lower than 0.5 in both channels, signals from the interquartile range of gene ratios were obtained and summed to give a normalization statistic which was then applied to all spots. The median value between the three spot replicates on each array was then obtained.

5.2.3.1 Levels of expression

Although the ability to measure gene expression on dual hybridisations is best expressed as a ratio, between channels (since inter-spot variability makes absolute quantification using this platform unreliable), a very broad indication of the gene expression levels can be inferred from absolute signal intensities, which is informative as to the general processes occurring in the endosperm.

The positive controls are all generally highly expressed, particularly Histone 2A, Histone 4, tubulins and actin and EF1. Surprisingly some of the spike genes and exogenous negative controls showed a fairly high level of expression: SP2, a myosin heavy chain and SP3, a myosin light chain both showed a moderate to high amount of expression. Controls provided through a collaborator (Dr V. Chandler, University of Arizona) were mostly lowly expressed, although ubiquitin (VC9 and VC 18) were moderately expressed in most samples.

In general the transposons and retrotransposons investigated here have expression patterns at the lower end of the observed values, with the main exception of Prem1

which shows a high level of expression in all samples tested, and Tekay is also relatively high.

Chromatin genes identified on the chromatin arrays as showing low expression levels were the bromodomain proteins, SWI3 homologues, Global Transcription Factors B (SPT6), histone acetyltransferases (CBP Family), and nucleosome/chromatin assembly factor (CAC1). Moderately expressed genes included Enhancer of Polycomb Protein, Global Transcription Factor A (SPT5), histone acetyltransferases (TBP), histone deacetylases (SIR2 Family), methyl-binding domain proteins, nucleosome/chromatin assembly factor (MSI1), high mobility group, and Silencing Group B. Highly expressed genes included nucleosome/chromatin assembly factor (NAP1) (the most highly expressed), SWP73 Homologues, Class 1 DNA methyltransferases, polycomb group (Esc), histone acetyltransferases (GNAT/MYST) Family), histone deacetylase (RPD3/HDA1 Family), and histone deacetylases (HD2 Family). Mixed expression levels were observed for Mixed SWI2/SNF2 homologues, and Global Transcription Factor E (BDF1-FSH-RING3). The SET domain proteins showed a broad range of expression levels from very low (SDG103, SDG104, and SDG120) to highly expressed (SDG105).

5.2.3.2 Regulated Genes

Table 21 shows the most differentially regulated genes between the two channels, in descending order, and Figure 5-18 charts the data of gene regulation in tabular format, grouped by gene class. The leaf/root comparison (Custom 1) shows higher expression of several photosynthesis related genes (photosystem I, photosytem II, and Rubisco components), which would be expected in a photosynthetic organ. The higher expression of histones, DNA binding proteins (HMG) and an elongation factor might reflect the high rate of growth in root tissue.

The results for the different endosperm and kernel comparisons show more substantial gene regulation than in the comparison of two different organs (leaf and root), and some general themes are apparent. However, inconsistencies between hybridisations and samples might indicate variability in the effects of interploidy crosses, or demonstrate limits of this particularl technology. Some general themes are nevertheless apparent.

The high levels of detected transposon and retrotransposon transcripts in certain endosperms is the most conspicuous phenomenon of the array. Most apparent is up-regulation of these elements in the paternal excess endosperms. OPIE, PREM2, PREM1, XILON, CINFUL, TEKAY are up-regulated more than 3-fold in hybridisation Custom 2, some support for this is provided in Hybridisation Custom 7, for an independent biological sample. Comparison in a different genotype (Custom 10), also suggests up-regulation of SPM, CIN4, GYPSY and AC elements. Regulation of transposable elements in maternal excess endosperms is inconclusive, as two comparisons with endosperm from balanced tetraploid crosses (Hyb 6), showed opposing effects (Custom 5 and Custom 7). A direct comparison of diploid and tetraploid controls suggested transposable elements were more highly expressed in the diploid.

The data shows inconsistent effects on the regulation of the transfer layer between comparisons, as *BETL1* is considerably down-regulated in the maternal excess (Custom 5), and supports the 14.1 fold decrease in oligo Hyb2 (Table 12). However, maternal down-regulation of *BETL1* was not supported by a biological dye-swap replicate (Custom 6). The higher expression of *BETL1* in the paternal excess than maternal excess was seen (Custom 9), although in the comparison of 10 DAP endosperm the

direction of regulation was reversed (Custom 2). Down-regulation of *BETL1* in the paternal excess compared to the diploid control was seen in two 4 DAP endosperm genotypes (Custom 7 and Custom 10). The regulation of expression of two other BETL markers on the array, *MEG1* and *ZmMRP*, is generally consistent with the trends just outlined for *BETL*, although this is not the case for the 10 DAP endosperm, in which *MRP* appears to be more highly expressed in the paternal excess, whereas *BETl1* appears to be more highly expressed in the maternal excess.

Another endosperm marker on the array , *ESR1*, appears to be down-regulated in the paternal excess (Custom 7 and Custom 10), which supports oligo array hybridisation 1, and less expressed in the paternal excess than the maternal excess in 10 DAP endosperms (Custom 2). However, down-regulation of *ESR1* in the paternal excess endosperms is not supported by Custom array 3.

Effects on chromatin remodelling factors are even less certain, as there is mixed information about regulation of chromatin remodelling factors subunits (CHR119, 120, 125) in the maternal excess (Custom 5 and 6). There is some indication of the down-regulation of some histone deacetylases in maternal excess (HDA106, HDA110), and the up-regulation of some histone acetyltransferases (HAC108, HAC112) (Custom 5 and 6). Down-regulation of nucleosome assembly factors in the paternal excess are also evident (NFA101, NFD101, NFD103, NFC101).

Protein Family		
Chromatin Genes	**ChromDB Code**	**Number of Genes**
Bromodomain proteins	BRD	3
SWI3 Homologues	CHB	1
Swp73 Homologue	CHC	1
SWI2/SNF2 Homologues	CHR	19
Class 1 DNA Methyltransferase	DMT	4
Enhancer of Polycomb Homolgoue	EPL	1
Polycomb Group Esc	FIE	2
Global Transcription Factor A SPT5	GTA	4
Global Transcription Factor B SPT6	GTB	1
Global Transcription Factor C SPT16	GTC	1
Global Transcription Factor E BDF1-FSH-	GTE	2
Histone Acetyltransferase TBP	GTF	2
Histone Acetyltransferases CBP Family	HAC	2
Histone Acetyltransferases GNAT/MYST	HAC	6
Histone Acetyltransferases TafII-250 Family	GTD	2
Histone acetyltransferase complex ADA2	HXA	1
Histone Deacetylase RPD3/HDA1 Family	HDA	5
Histone Deacetylase HD2 Family	HDA	4
Histone Deacetylase SIR2 Family	HDA	1
Histones	HON	3
Methyl-Binding Domain Proteins	MBD	7
Nucleosome/chromatin assembly factor	NFA	3
Nucleosome/chromatin assembly factor	NFB	1
Nucleosome/chromatin assembly factor	NFC	5
High Mobility Group Family B	NFD	10
SET domain proteins	SGD	25
Silencing Group B	SGB	3
Transposable Elements		16
Endsperm Specific Genes		15
Controls		51
Total		**201**

Table 19 Genes represented on custom array

Further details are presented in Appendix 1

	Sample	
Hyb Number	Cy3	Cy5
Custom 1	Leaf	Root
Custom 2	2x x 4x Whole Kernels 10 DAP	4x x 2x Full kernels 10 DAP
Custom 3	H99-2x ⊗ 4 DAP (#52.3)	H99-2x x 4x 4 DAP (#6.1)
Custom 5	H99-4x ⊗ 4 DAP (23.1)	H99-4x x 2x 4 DAP (28.1)
Custom 6	H99-4x x 2x 4 DAP (#8.1)	H99-4x ⊗ 4 DAP (15.1)
Custom 7	H99-2x x 4x 4 DAP (#32.1)	H99-2x ⊗ 4 DAP(#60.1)
Custom 8	H99-2x ⊗ 4 DAP (#60.1)	H99-4x ⊗ 4 DAP (#23.1)
Custom 9	H99-4x x 2x 4 DAP (#28.1)	H99-2x x 4x 4 DAP (#30.1)
Custom 10	OH43-2x x 4x 4 DAP (#9.2)	OH43-2x x 2x 4 DAP (#1.1)

Table 20 Hybridisations carried out using custom array

Table details which samples were used in the custom arrays, and with which fluorophore they were labelled.

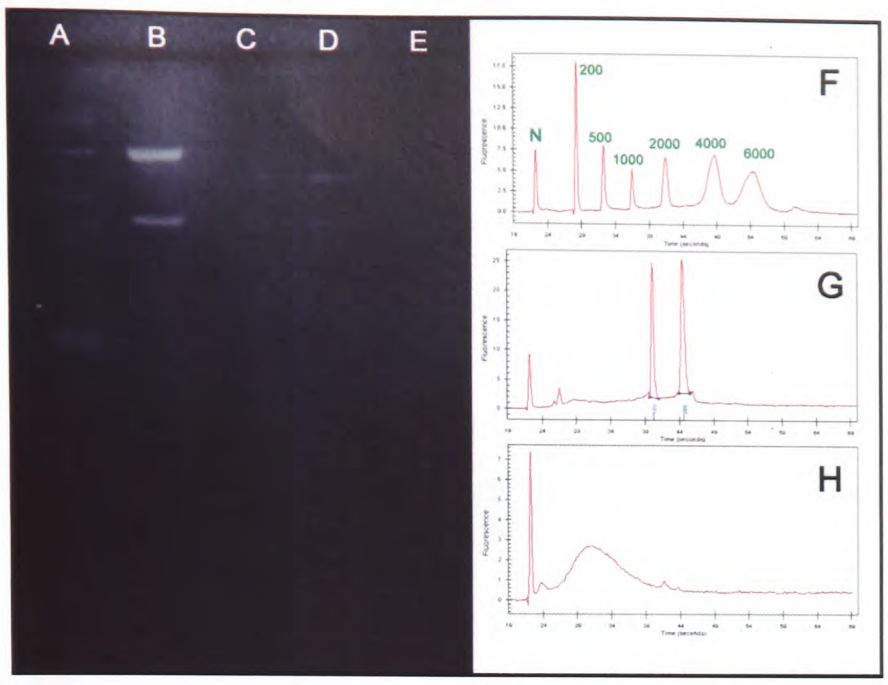


Figure 5-17 Visualization of RNA and aRNA on denaturing gel and Bioanalyzer

Left hand side is an image of a formaldehyde gel of RNA to verify integrity. Lane **A**) RNA ladder, **B**) 1.2 µg control HeLa RNA, **C**) 0.7 µg amplified Hela RNA, **D**) 0.34 µg 4 DAP diploid-selfed endosperm RNA, **E**) 0.37µg amplified RNA generated from same source as lane (**D**). **F-H** are Bioanalyzer traces of total and amplified RNA. **F**) RNA nanomarker, marker sizes labelled in green in bases, **G**) 4 DAP diploid-selfed endosperm RNA, **H**) amplified RNA generated from same sample as (**G**).

Hyb1 Up in C1 Leaf		Hyb1 Up in C2 Root	
PDP7_PS1	6.1	PDP6_H2A	2.8
PP6_Photosystem II	4.6	EA1	2.4
PDP11_	4.5	PP1_EF1A	2.1
PDP9_Rubisco SBP	2.0	PP4_HMG1	2.1
PDP8_Rubisco SBP	1.8	VC18_Ubiquitin	2.0
PP8_Rubisco SSU	1.7	PP3_Histone H4	2.0
TG7_Luciferase	1.6	CHR119	2.0
TG6_Hygromycin BPT	1.5		
GTD101	1.3		
NFB101	1.3		
NFD109	1.3		
MRP	1.3		
CHR120	1.3		
HAC109	1.3		
MRP	1.3		
HAC110	1.3		
MBD108	1.3		
FIE102	1.2		
HDA102	1.2		
GTB101	1.2		
TG3_GFP	1.2		
CRD109	1.2		
SDG106	1.2		
END1	1.2		
CHR124	1.2		
SDG125	1.2		
ZEON	1.2		

Hyb2 Up in C1 2x x 4x > 4x x 2x		Hyb2 Up in C2 4x x 2x > 2x x 4x	
HAC109	4.4	ESR	6.1
CHR119	3.3	PP5_MNBB	4.6
TG6_Hygromycin BPT	3.0	BETL	4.6
MRP	2.9	PP1_EF1A	4.3
CHR124	2.8	NFD103	4.1
REINA	2.8	VC18_Ubiquitin	4.0
HDA114	2.8	VC3_anthocyaninless1	3.9
CHR110	2.7	NFD103	3.6
CHR120	2.7	PDP7_PS1	3.3
CHR114	2.7	VC2_anthocyaninless1	2.9
PP7_Rubisco LSU	2.7	PP6_Photosystem II	2.8
GTC101	2.6	PP9_Tubulin ?1/2	2.7
HAC110	2.6	NFD101	2.7
PP7_Rubisco LSU	2.5	NFC101	2.6
VC15_MudrA	2.5	PDP1_Actin	2.6
GTA107	2.5	HDA101	2.6
TG4_Globin	2.5	NFB101	2.6
SDG119	2.4	PREM2	2.6
HAC112	2.4	PP4_HMG1	2.6
CINFUL	2.4	GTA106	2.5
MBD108	2.4	HAC109	2.5
HAC112	2.4	HDA105	2.5
CHR102	2.4	VC9_Ubiquitin	2.3
SDG106	2.3	HC3	2.3
AC900	2.3	PP10_Tubulin ?	2.3
HC4	2.3	PP3_Histone H4	2.3
CHR106	2.3	PDP11_	2.3

Hyb3 Up in C1 2x x 2x > 2x x 4x		Hyb3 Up in C2 2x x 4x > 2x x 2x	
NFD103	8.2	OPIE	20.0
NFD103	6.2	PREM2	16.0
NFD101	4.9	PREM1	12.6
NFD101	4.2	XILON	7.6
NFC102	3.9	MYBst	4.4
HDA114	3.7	CINFUL	4.2
SDG105	3.7	FIE102	3.8
MBD105	3.5	TEKAY	3.6
NFA104	3.4	MBD101	3.4
SDG107	3.4	SGB102	2.8
PP1_EF1A	3.3	DMT106	2.8
MBD105	3.2	SDG115	2.8
HAC108	3.2	SDG102	2.7
PDP1_Actin	3.0	FIE2	2.5
SDG117	3.0	PP11_Tubulin ?3	2.5
PP3_Histone H4	2.9	SP2_Myosin HC	2.4
DD1	2.9	ZEON	2.4
PP10_Tubulin ?	2.8	SDG103	2.4
PDP12_	2.7	NFA101	2.4
NFC101	2.6	SPM	2.3
SGB103	2.6	HDA109	2.3
PP5_MNBB	2.6	BRD111	2.3
NFA101	2.5	MEG1	2.3
CHR111	2.4	NFD105	2.2
SGB101	2.4	MRP	2.2
PP9_Tubulin ?1/2	2.4	HAC111	2.2
HON105	2.4	GTE101	2.2

Hyb5 Up in C1 4x x 4x > 4x x 2x		Hyb5 Up in C2 4x x 2x > 4x x 4x	
BETL	14.1	FIE102	2.5
MRP	8.3	FIE101	2.2
HAC112	6.7	HDA101	1.9
MRP	6.6	SDG105	1.7
CHR120	6.5	HDA101	1.7
MRP	6.5	PREM1	1.7
CHR119	6.4	HDA101	1.6
GTC101	6.3	HAC108	1.6
GTA107	5.8	PDP12_	1.6
GYPSY	5.6	PP9_Tubulin ?1/2	1.6
HC4	5.6	NFC101	1.5
FIE2	5.2	ESR	1.5
END1	5.2	PDP1_Actin	1.5
REINA	5.1	HAC108	1.5
HDA113	5.0	FIE2	1.5
DMT106	5.0	PP10_Tubulin ?	1.5
CHR124	4.9	CHR110	1.4
GTF102	4.9	DD1	1.4
CIN4	4.8	GTE101	1.4
TG3_GFP	4.7	NFC102	1.4
AC700	4.6	SDG118	1.4
OPIE	4.6	NFC101	1.4
GTE101	4.6	GTF101	1.4
TG7_Luciferase	4.5	CHC101	1.4
EA3	4.4		
END1	4.3		
HC1	4.3		

Hyb6 Up in C1 4x x 2x > 4x x 4x		Hyb6 Up in C2 4x x 4x > 4x x 2x	
CHR125	7.5	NFA103	2.3
AC900	7.0	HDA106	2.1
GTC101	5.9	HDA106	2.0
FIE2	5.8	NFA104	2.0
MRP	5.5	HDA114	2.0
PHP	5.2	HDA110	2.0
MRP	4.8	SDG118	2.0
CINFUL	4.8	NFD110	1.9
GYPSY	4.7	DMT101	1.9
HC4	4.7	NFA103	1.9
GTA105	4.6	HDA104	1.8
CHR120	4.6	NFD102	1.7
OPIE	4.5	PDP8_Rubisco SBP	1.7
STONOR	4.5	EA3	1.6
HAC112	4.3	SDG105	1.6
AC700	4.3	PDP7_PS1	1.5
DMT106	4.3	CDC2	1.5
END1	4.2	PDP1_Actin	1.5
END1	4.2	CHR122	1.5
CENTA	4.1	NFC104	1.4
VC14_Mu DR	4.1		
REINA	4.1		
VC16_MuRA	4.1		
GTA107	4.1		
VC3_anthocyaninless1	4.0		
TG7_Luciferase	4.0		
CIN4	4.0		

Hyb7 Up in C1 2x x 4x > 2x x 2x		Hyb7 Up in C2 2x x 2x > 2x x 4x	
MRP	6.4	ESR	5.9
BETL	6.3	FIE101	3.4
AC900	6.0	PP10_Tubulin ?	3.1
GYPSY	5.9	HDA104	3.0
GTC101	5.7	NFD101	2.9
REINA	5.5	NFA101	2.8
CIN4	5.1	NFD101	2.8
GTA105	4.8	PP9_Tubulin ?1/2	2.7
SPM	4.7	HDA105	2.7
NRP	4.5	CHR111	2.7
CHR125	4.4	PP1_EF1A	2.7
VC16_MuRA	4.4	NFA101	2.7
CHR124	4.3	PP4_HMG1	2.6
AC700	4.2	HAC108	2.5
HC1	4.2	SGB102	2.5
CINFUL	4.2	EA3	2.5
CINFUL	4.1	NFD101	2.4
END1	4.1	PDP12_	2.4
OPIE	4.1	NFA103	2.4
CENTA	4.0	HAC109	2.3
CHR120	4.0	CHR116	2.3
HDA109	3.9	SDG105	2.3
GTA107	3.8	PP5_MNBB	2.3
TG7_Luciferase	3.5	NFD103	2.3
GTA106	3.4	HDA106	2.3
HC4	3.4	NFC101	2.3
ZEON	3.3	NFA101	2.3

Hyb8 Up in C1 2x x 2x > 4x x 4x		Hyb8 Up in C2 4x x 4x > 2x x 2x	
CHR120	6.8	PP3_Histone H4	3.3
GTB101	6.4	PDP6_H2A	2.9
AC700	6.4	NFD103	2.3
GTC101	5.2	NFC101	2.3
CHR120	5.1	NFC101	2.3
HAC112	5.0	NFD101	2.3
CHR124	5.0	NFC102	2.3
XILON	4.9	DD1	2.3
GTA107	4.8	NFA103	2.2
GYPSY	4.6	SDG105	2.2
HDA110	4.6	SDG118	2.2
BRD113	4.4	NFD101	2.2
TG6_Hygromycin BPT	4.3	PP4_HMG1	2.2
MRP	4.3	HON105	2.2
PREM2	4.3	NFA101	2.2
SPM	4.2	SDG109	2.1
DMT106	4.2	NFA104	2.1
HC1	4.2	PP5_MNBB	2.1
CENTA	4.0	NFA103	2.0
REINA	4.0	NFD101	2.0
NRP	4.0	NFA101	2.0
TG7_Luciferase	4.0	HDA108	2.0
FIE2	3.9	HAC108	1.9
VC14_Mu DR	3.9	PEG1	1.9
HXA102	3.9	NFD105	1.9
CINFUL	3.9	NFD108	1.9
VC16_MuRA	3.8	HDA105	1.9

Hyb9 Up in C1 4x x 2x > 2x x 4x		Hyb9 Up in C2 2x x 4x > 4x x 2x	
CHR124	6.8	BETL	8.7
TG6_Hygromycin BPT	6.0	MEG1	3.6
AC900	5.0	PHP	3.4
AC700	4.4	NFD110	2.8
TG3_GFP	4.1	PDP7_PS1	2.8
TG4_Globin	4.0	END1	2.5
TG5_GUS	4.0	SDG101	2.5
DD1	3.8	NFA101	2.4
SDG115	3.7	NFA101	2.3
SDG112	3.6	MBD101	2.3
SPM	3.5	CHC101	2.3
MRP	3.3	DMT105	2.2
TG1_CRY1aC	3.2	HAC108	2.2
CHR120	3.2	NFA101	2.2
REINA	3.2	CDC2	2.2
NFC104	3.1	NFA101	2.2
SGB102	3.1	DMT102	2.2
VC3_anthocyaninless1	3.1	HDA105	2.1
HC3	3.1	GTE101	2.0
HDA110	3.1	NFA101	2.0
HAC112	3.1	SP3_Myosin RLC2	1.9
TG7_Luciferase	3.0	CHR122	1.8
VC14_Mu DR	3.0	PREM2	1.8
DMT106	2.9	CHR111	1.8
SDG116	2.9	SDG123	1.8
SDG103	2.8	SDG108	1.7
SP1_B-cell receptor	2.7	SDG123	1.7

Hyb10 Up in C1 2x x 4x > 2x x 2x		Hyb10 Up in C2 2x x 2 x > 2x x 4x	
SPM	7.7	PP3_Histone H4	16.4
CHR120	7.6	MEG1	10.6
TG6_Hygromycin BPT	7.3	BETL	10.3
TG7_Luciferase	7.3	NFD101	8.6
TG3_GFP	7.1	HON102	8.3
SDG115	6.7	PP9_Tubulin ?1/2	7.7
TG7_Luciferase	6.6	PDP6_H2A	7.0
CHR124	5.9	PP5_MNBB	6.8
SDG103	5.6	NFD101	6.8
GTB101	5.6	HON105	6.5
SDG103	5.4	HDA114	5.8
HC4	5.4	NFD103	5.8
TG5_GUS	5.3	VC18_Ubiquitin	5.0
VC3_anthocyaninless1	5.3	NFD103	4.9
AC900	5.2	PHP	4.7
TG1_CRY1aC	5.1	FIE101	4.5
GYPSY	5.1	NFC102	4.4
GTC101	5.0	SGB101	4.1
AC700	5.0	PP10_Tubulin ?	4.0
SDG103	4.9	NFA103	3.9
SDG122	4.9	DD1	3.9
CIN4	4.8	SP3_Myosin LC	3.8
CENTA	4.7	ESR	3.8
KN1_Knotted1	4.5	SGB101	3.7
EPL101	4.5	HAC109	3.6
VC1_anthocyaninless1	4.4	SGB103	3.6
HC1	4.3	NFA103	3.6

Legend

- 2x ⊗ Diploid Control
- 4x ⊗ Tetraploid Control
- 2x x 4x - Paternal Excess
- 4x x 2x - Maternal Excess

Figure 5-18 Genes most substantially regulated on custom array

Data which had been normalized using an inter-quartile range method was sorted to find the most highly regulated genes between the two conditions on each array. Nine separate hybridisation are shown, the left hand column is the Cy3 labelled channel, and the right hand side Cy5 labelled. Genes are listed in descending order of regulation in that channel. The sample labelled by either channel is indicated by colour at the top of each column.

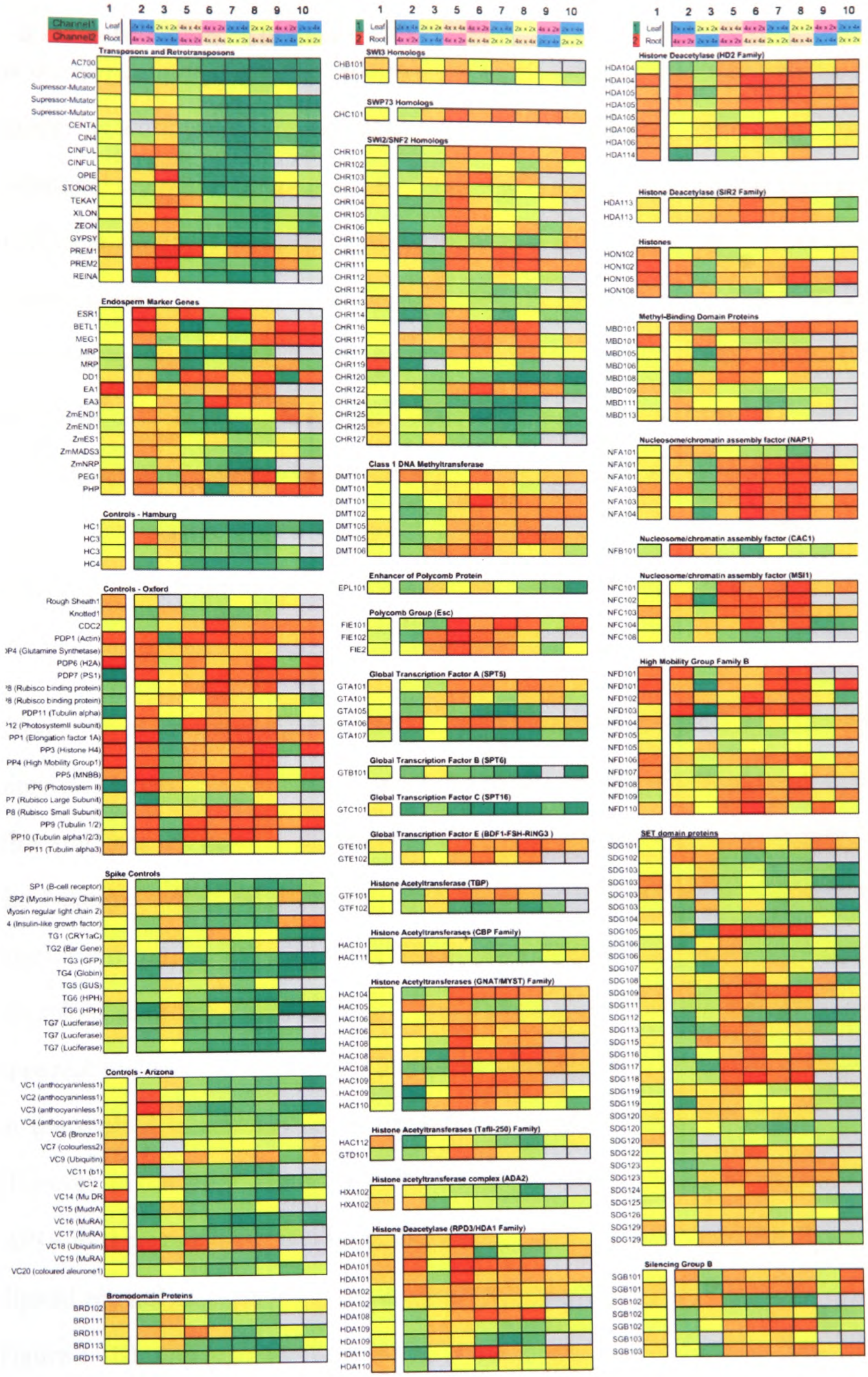


Table 21 Tabulation of gene expression changes in custom microarray

Following data normalization the median ratio value between the three replicates of each spot was tabulated, and a colour assigned based its value in comparison to all other spot ratios for that hybridisation. A yellow colour spot suggests an average ratio (which would be close to 1), whereas green suggests up-regulation in channel 1 (Cy3), and a red colour indicates up-regulation in channel 2 (Cy5). Spots are grouped into gene classes, and the hybridisation number, are listed in columns at the top of each block.

5.2.4 Verification of results by RT-PCR

An independent method was adopted to verify whether the observations from the microarray work could be confirmed by an alternative technology, and additionally, to investigate temporal differences. A subset of genes were selected and analyzed by RT-PCR (Le et al., 2005), since the most highly regulated genes included some of the BETL marker genes, BETL1,2, and 3 were tested, in addition to *ZmMRP*, *ZmESR*, and *ZmEND1*, as displayed in Figure 5-19. Additionally, a control was selected from the results of the chromatin microarrays, GTA101 an homologue of yeast SPT5, a transcriptional elongation factor, which was shown to be moderately expressed in all samples and showed the least variance across experiments (Figure 5-19A). Expression of the control was similar through samples, so verified the technique.

Expression of the transfer layer markers *BETL2* and *BETL3* were down-regulated in the maternal excess cross in comparison to the other crosses observed at all time-points. This supports the general observations from both microarrays. Surprisingly, although the balanced tetraploid cross displayed signal for *ZmMRP* at all stages, the diploid control did not, indeed signal for *BETL1*, and *BETL2* is considerably higher for the tetraploid than the diploid contol. The higher expression of *ZmMRP* supports the apparent higher expression in the tetraploid, observed in Table 12, but the lack of signal in the diploid is contrary to the reported expression pattern (Gomez et al., 2002). However, in the latter paper, the authors used a considerably higher starting amount of cDNA in their reaction. Similarly, BETL1 appears to not be expressed as highly in the diploid control as expected from the reported expression profile (Hueros et al., 1995) Figure 1-1B), although the time-points investigated here are at the early stage of its expression, when it is still quite low in abundance. The appearance of BETL1 in the tetraploid control at all time-points suggests higher expression in the tetraploid than diploid, possibly suggesting a dosage effect on the domain. Surprisingly, BETL1

appears to be expressed more highly in the paternal excess than in the diploid control, which is contrary to observations in the reporter line. The *BETL2* gene is expressed similarly between the paternal excess and diploid control, and displays an earlier expression pattern than *BETL1* (Figure 1-1C). The gene is expressed in the diploid control at all time points, but consistently at a lower level than the tetraploid control. However this supports the observation from oligo array Hyb 3 (Table 12). There is not much difference in *BETL2* expression between the diploid control and the paternal excess, which is also consistent with oligo array Hyb 1 (Table 12). A clear down-regulation of *BETL2* in the maternal excess also supports the oligomer array hyb 2 (Table 12). Interestingly, the *BETL3* gene is expressed similarly between the diploid and paternal excess, but is only temporarily active in the maternal excess, and not at all in the tetraploid (Figure 5-19D).

ZmEND1 is a marker specific to mid-stage endosperm (Dr J Gutierrez-Marcos, University of Oxford), and it is expressed at similar levels in all three endosperm samples, except the paternal excess, where it is down regulated (Figure 1-1E).

The expression pattern for the embryo surrounding region is similar between the two controls, consistent with the oligo hyb 4 (Table 12), being expressed higher at earlier stages. The paternal excess shows a lower expression pattern at 4 DAP than the diploid, also consistent with the oligo hyb 1 (Table 12), although the increased expression at other stages suggests the region does form, but is delayed. The maternal expression pattern of this gene is consistent with the oligo hyb 2, although its decreased expression by 8 DAP suggests the expression profile of this transcript is accelerated in the maternal excess.

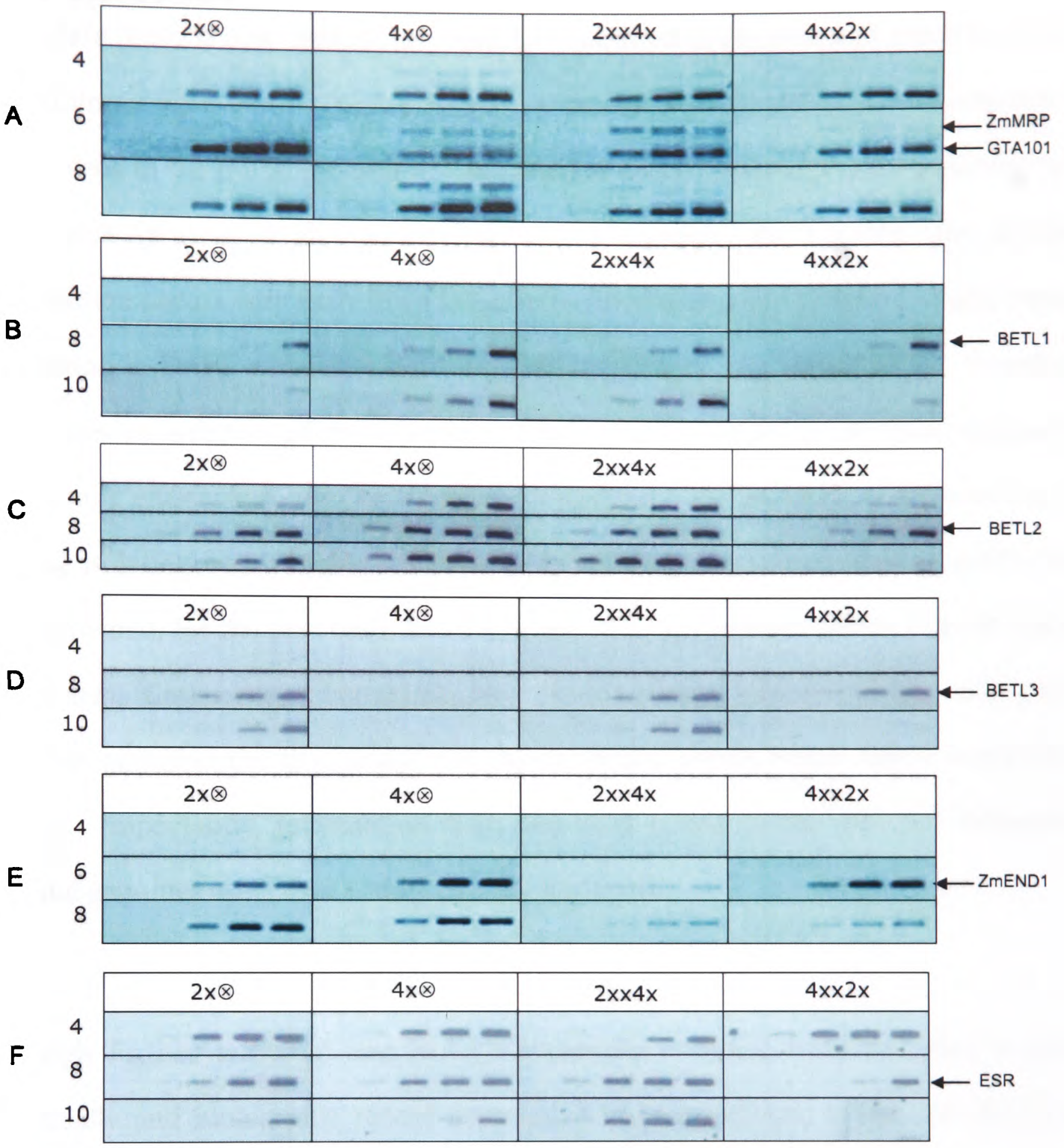


Figure 5-19 RT-PCR analysis of gene activity through endosperm development

RT-PCRs were performed on 5 ng cDNA, individual wells represent PCR product taken after 20, 25, 30, 35, and 40 cycles. Endosperm from the crosses listed at the top of the tables were taken at the number of DAP listed on the left. Primers were used to amplify **A**) GTA101 (control), and *ZmMRP*, **B**) *BETL1*, **C**) *BETL2*, **D**) *BETL3*, **E**) *ZmEND1*, and **F**) ZmESR.

5.3 Discussion

These data provide new insights into seed development in general, and specifically into maize interploidy crosses. There is limited information about global gene expression in maize seeds in the literature, and the information which does exist is for relatively late stage seeds (Hueros et al., 1999b; Hueros et al., 1995). As it seems that domain specification occurs very early in endosperm development, even prior to cellularization (at around 4 DAP), obtaining data for this stage is of key importance. However, difficulties in isolation of endosperms, and the limited amounts of RNA obtainable from young endosperm have restricted such analyses. In this work, a combination of manual isolation and a highly sensitive RNA labelling method or RNA amplification have provided, for the first time, an insight into the expressional profile of early maize endosperm. Comparisons between control (within ploidy) and interploidy endosperm RNA has provided the first analysis of the molecular events which follow interploidy crosses. Importantly, information from this work points to the different influences parental genomes exert over endosperm development.

Although further technical and biological replication would have provided a more statistically and biologically robust assessment of these effects, it was prevented by limiting resources. However, the encouraging general concordance between the oligo arrays and RT-PCR analysis, suggest that the data can be used to draw some preliminary conclusions with confidence. Further, the apparent differences between the methods are likely to be explained by the additional stages of development assessed in the RT-PCR. However, the data presented here are not intended to provide a complete assessment of the phenomena under investigation, but to provide a pointer to the general processes occurring, and also to suggest directions for future study.

5.3.1 Technological Insight

The use of the dendrimer technology was adopted for this work because it offered an opportunity to increase signal for samples using small amounts of RNA (Baggerly et al., 2004), but without the potential problems associated with RNA amplification methods (Hueros et al., 1999b; Puskas et al., 2002; Schneider et al., 2004). The results did, however, reveal a limitation of the system: namely signal-specific fluorescence of some sequences without apparent similarity to linear nucleotide sequence (Figure 5-8). In spite of this, the system has enormous advantages over more commonly used methods of fluorescent labelling, and, as part of this work, a method of identifying and eliminating these sequence specific effects was determined. This involves a test hybridization being performed on each new array type using cDNA reverse-transcribed with primers lacking the complementary sequence to the dendrimer capture sequence. The usual secondary incubation in the presence of both fluorophores would thus identify spots which fluoresce non-specifically. This test is now used as a control for all new slide types used in the laboratory.

5.3.2 Gene expression in early endosperm

This work identified the most abundant RNA populations in early-stage endosperm, which are enriched in sequences for protein biosynthesis and cytoskeleton genes (

Figure 5-12). This is logical, as ribosomal genes are most highly expressed genes in tissues, and because the endosperm is presumably involved in high levels of protein biosynthesis to support its rapid growth. Additionally, at this stage the endosperm has just undergone a phase of cellularization, which involves significant cytoskeletal dynamics (Olsen, 2001).

5.3.3 Changes in gene transcription following interploidy crosses

At the early stage of endosperm development assessed in this work the majority of genes are not affected by paternal excess (5.9% are regulated more than 2-fold), although this is a higher proportion than those genes affected by maternal excess crosses (2.2%) (Figure 5-11). A direct comparison of the two interploidy crosses suggested an almost additive effect was observed: 7.7% of genes were affected, suggesting that a different subset of genes was affected by each interploidy cross, which was also suggested by direct comparison of genes which were up-and down-regulated between the different crosses. Fewer genes are changed in maternal excess crosses, indicating that this direction of cross is affected less by the interploidy cross. This is again supported by the histology, which more closely resembles wild-type cross until late stages of development. The expression pattern of a general marker for mid stage endosperm (*END1*), more closely resembled the control crosses than the paternal excess, which also supports the notion that perturbation of development is initially less severe in maternal excess crosses.

Mis-regulation of endosperm domain markers was seen in the cell biology study (Chapter 4) and there was some correlation with expression data, both from oligo and chromatin arrays (5.2.2.2, and 5.2.3.2). Further, for the maternal excess cross, this molecular work directly supported observations in the reporter lines (Sub-chapter 4.2.6), and the suggested up-regulation of storage molecules (zeins, glutelins, and starch biosynthesis - Table 14, Table 15 and Table 17), and cytoskeletal genes also supports the histological observations of accelerated development in the CSE. It is remarkable that even at this early stage of endosperm development, when the cells have only just cellularized, genes involved in processes which would normally commence five days later are active in maternal excess embryos.

Paternal excess is characterized by the activity of a number of calcium signalling proteins, and signal transduction molecules, in addition to the up-regulation of an amino acid transporter (Table 13) and amino metabolism proteins (Table 18). A down-regulation of cytoskeletal genes, is also observed, as is up-regulation of genes involved in pre-prophase band formation (Table 17, which supports the trend towards cell proliferation rather than cell maturation apparent in this direction of cross.

One of the most frustrating aspects of the microarray work is the incomplete concordance between different comparisons. This could be simply due to technical limitations of the system, however, it is quite possible that it represents a general variability between samples, which is also observed in some of the histological work. Whereas the phenotypic effect of interploidy crosses was consistent between maternal excess kernels, significant inter-kernel variability was observed in the paternal excess kernels. For example, in the maternal excess cross marker genes of the BETL layer were one of the most conspicuously down-regulated classes. This is supported by the lack of a transfer layer in these crosses (Figure 4-12), and is corroborated by RT-PCR which suggest that all BETL transcripts are expressed less in these endosperms, and within a more confined time-frame. In contrast, investigation of paternal excess endosperms suggest no significant difference in BETL marker expression, which seems to contradict histological work (Chapter 4.2.9), in which the reporter for the BETL domain was reduced, however there was variability between samples. This variability could account for the variability between chromatin microarray experiments. Moreover, the reporter line studies point to later stage paternal excess kernels having random activation of at least three genes. The discrepancy between the gene expression results and the reporter line work might simply relate to different activation of the endogenous gene and the transgene (which will have different genomic locations and therefore might not have the same expression pattern). Custom microarray data suggests that paternal excess endosperms show an up-regulation of transposons and

retrotransposons, if these genes become active and integrate randomly though the genome, they could alter the expression pattern of the genes into which they integrate. This may explain the apparent variation between gene expression patterns between kernels and the later stage histological observations.

Another reason that no significant regulation was found for the chromatin genes between these classes of genes could be due to the fact the system is not sensitive enough to detect the kinds of changes that might be anticipated. It is commonly accepted that only measured levels of differential regulation greater than two-fold are significant in microarray studies. Many chromatin-associated factors are known to work in complexes with other subunits, and therefore require a tight stoichiometry, they will be dosage sensitive. Hence, subtle changes in copy number would be reflected in subtle changes in expression levels, possibly no more than two-fold: this would not be detected with microarrays. A more sensitive approach, such as real-time RT-PCR, would be more appropriate for assessing these kinds of effects.

The RT-PCR time-series experiments revealed that the three BETL genes investigated showed different profiles in response to interploidy crosses. This suggests that even though the genes appear to be expressed in the same seed domain (Hueros et al., 1999b), and have similar mid-term expression profiles, they have different control of expression, which could imply different biological functions (which have so far been poorly defined). *BETL1* and *BETL2* are most similarly expressed between paternal excess and tetraploid selfed kernels, and the diploid control and maternal excess show quite similar profiles for *BETL1,* but not *BETL2*. In contrast, *BETL3*, displays a similar expression profile between the diploid and paternal excess, but differ from the maternal excess and tetraploid control.

Another significant finding is that there are a sizeable number of genes which are regulated differently between the diploid and tetraploid controls, although fewer genes are affected than with interploidy comparisons. Such changes in gene expression following polyploidization have been reported elsewhere (Madlung et al., 2002; Osborn et al., 2003; Pikaard, 2001). Two classes of changes have been observed, those which occur consistently between polyploidization events, and those which show variability between plants (Wang et al., 2004). It has been shown that recent polyploidization events can confer altered expression patterns on genes and these might be the consequence of dosage-effects on sensitive genes (Birchler, 1993), and chromosomal and epigenetic changes which occur following polyploidization (Pikaard, 2001). Further, an apparent difference in seed size between the diploid and tetraploid kernels suggests physiological differences that must also be reflected transcriptionally.

6 General Discussion

6.1 Summary and main conclusions

The work presented in this thesis has provided new information into the phenotypic consequences of interploidy crosses, and also reinforced data from earlier studies (Charlton et al., 1995; Cooper and Brink, 1945; Leblanc et al., 2002). New insights have also been provided into endosperm development in terms of the specification of two of the cellular domains. Additionally, this work can also be viewed as lending support to the parental conflict model of imprinting as it relates to the effect of interploidy crosses.

6.2 Insights into the specification of the aleurone and BETL

A detailed description of the activity of the promoter for the *VP1* gene through normal development has been provided, as assessed by a GUS reporter gene (Section 4.2.1). The transgene acts as an accurate predictor of aleurone fate, although its activity precedes by around 5 days the appearance of cells with a typical aleurone morphology (Section 4.2.2).

A temporal profile of activity of the BETL promoter has also been provided by utilisation of a GUS reporter (Section 4.2.3). This promoter acts as a marker of mid-stage BETL cell fate in normal crosses, as BETL morphology is first observed four days prior to GUS activity (Section 4.2.4).

The different effects of interploidy crosses on these two cell domains demonstrate that different mechanisms are involved in their specification. The aleurone fate is the more robust of the two cell types, being defined in both maternal and paternal excess crosses, albeit at altered intervals. In contrast, the BETL cell fate is disrupted in both interploidy crosses, but evidently for different reasons.

Maternal excess endosperms suggest that their entire periphery, including what would normally form the BETL, is capable of forming an aleurone (Figure 4-11). This necessitates that any 'aleurone inducing factor' is present around the entire periphery in these crosses and from at least as early as 6 DAP. The accelerated differentiation into cells with an aleurone morphology in maternal excess crosses is indicative of a more general accelerated differentiation process observed in this direction of cross, as suggested by premature starch accumulation and aleurone differentiation. The failure to form a BETL in maternal excess crosses throws some light on the interplay between aleurone and BETL cell identity. Possibly the inducing signal for BETL fate is delayed in maternal excess crosses, such that naïve cells located at the base of the endosperm instead respond to the aleurone inducing signal. Studies on a defective endosperm mutant, which showed failure to correctly form a transfer layer proposed that this domain might be specified within a short window of time (Costa et al., 2003), in which case the maternal excess crosses might affect the timing of this window, or the time at which cells become responsive to it. It has been suggested that a BETL-inducing signal is likely to originate from the maternal tissue via the pedicel (Thompson et al., 2001), in which case the timing of the inducing signal would be the same, implying the capacity to respond to it is altered in the endosperm is altered. Since the maternal excess endosperms display a general accelerated development it is most likely that this 'window of receptivity' occurs too early, before the maternal tissue has started to unload resources at the placenta. Cells would hence be undefined, and become responsive to other inducing signals, such as that for aleurone fate. An alternative hypothesis to account for observed replacement of the BETL with aleurone cells is that the presence of an aleurone inducing signal may be accelerated and basally located cells therefore adopt an aleurone cell fate, this is consistent with the premature differentiation of the aleurone. Also, GUS activity for the BETL reporter is seen in maternal excess crosses at a low frequency, but when it is, it starts no sooner than in

control kernels. It is interesting that the cells located close to the embryo are those least likely to form an 'ectopic aleurone'. This might suggest that the embryo is the source of BETL inducing signal; that it confers a lowered threshold of perception onto nearby cells; or somehow interferes with the aleurone inducing signal.

Paternal excess endosperms are characterized by specification of aleurone cells at approximately the correct time (as shown by GUS activity first occurring about the same time as in the balanced control), but a delayed differentiation of these cells to true aleurone morphology. This suggests an opposite effect on the rate of differentiation as seen in the maternal excess endosperms, and is also reflected in the underlying cells which appear poorly differentiated and permit substantial GUS diffusion until later in development. However, the failure of paternal excess endosperms to form a normal BETL layer is more difficult to explain. Intriguingly, GUS appears prior to the appearance of typical BETL morphology, contrary to control endosperms. For this reason it seems likely that these cells are not actually responding correctly to their normal developmental cues, but for some reason, the cells located in this area of the endosperm are more susceptible to the processes which are evident throughout the endosperm from later stages of development namely sporadic ectopic GUS activity. This is also supported by the fact when crossed with one of the tetraploid genotypes (H99-4x) 'ectopic' pBETL::GUS activity were observed from the same stage at which it was found in the basally located cells. Throughout development domains of GUS activity tend to increase in size, which probably reflects clonal development of cells which had GUS expression induced at an earlier stage (Costa et al., 2004).

An alternative proposal for the distorted formation of the transfer layer in interploidy crosses could involve dosage sensitive factors, the balance of which could be disrupted following interploidy crosses. As the transfer layer is the main route for uptake of

nutrients into the endosperm (Thompson et al., 2001), it may be the target of action of genes from the two parents with opposing effects. In arabidopsis, the *medea* mutant, which displays an over-proliferative phenotype similar to paternal excess crosses also shows an enlarged chalazal region - the functionally equivalent domain to the BETL. Significantly, the RT-PCR work carried out in this thesis (Section 5.2.4), suggests that the tetraploid has a much increased transcripts of basal layer genes than the diploid, which might reflect a dosage effect controlling the size of the BETL. If a similar process was to occur in maize, the reduction or lack of a transfer layer in the maternal excess makes sense, although its disruption in the paternal excess is more difficult to explain. However, an interploidy cross results in an enormous disruption to the system, and if processes are disturbed too much, then they may fail to give any further insight into their normal roles in development. In other words, the trend for the paternal genome to increase the size of the transfer layer may be confounded by other processes that are disrupted following this cross. The histological analysis of the basal region of the endosperm showed that whilst the cells located here were usually not typically transfer-like (except for a few later in development), they were also clearly not aleurone-like or typical of the CSE. It seems they must have undergone some developmental program, and by the apparently normal levels of expression of BETL marker genes (Figure 5-19 and (Hueros et al., 1999b)), it was presumably a BETL cell program, which must have become modified. Conceivably these cells possessed an excessive amount of a signal which induces the promotion of the BETL, provided by the additional paternal genome. Such a gene might be imprinted on the female side, and expressed paternally to maximize nutrient uptake, this could explain the up-regulation of several amino acid transporters in this cross. However, the program conferred might be simply too extreme, and instead a modified BETL program ensues, which leads to the formation of tissue which stains deeply for carbohydrate, but is not obviously BETL-like. As maize homologues for E(z) proteins have been identified (Springer et al., 2002), it will be interesting to see whether knockouts in this gene lead to a phenotype on the basal layer

which mimics paternal excess interploidy crosses, as *mea* mutants do in arabidopsis, caution must be payed though, as these genes might be more general repressors of transcription, and the closest E(z) homologue of maize is more similar to the *CURLYLEAF* gene than *MEA*.

6.3 Support for the conflict hypothesis?

The effects of interploidy crosses have been reported in several species, and where alternate developmental effects have occurred (Lin, 1982; Scott et al., 1998) a role for imprinted genes whose balance is distorted has been proposed. The phenotypic consequences of these crosses have lent support to the parental conflict hypothesis for genomic imprinting (Haig and Westoby, 1989). The predictions of this hypothesis were that imprinted genes were the ultimate consequence of differential interests of the mother, father and offspring on resource allocation. As such, the paternal parent, whose offspring are likely to be competing for maternal resources alongside unrelated individuals, would gain advantage by increasing the loading of nutrients. In contrast, the maternal parent shares an equal relatedness to all her offspring, and should partition her resources equitably. This, it is proposed, would lead to the evolution of paternally expressed, maternally silenced nutrient transporters and growth promoters, and paternally silenced but maternally expressed repressors of those genes (Haig, 2000).

Although the molecular work presented here does suggest a small number of amino acid transporters are up-regulated in the paternal excess cross, supporting this argument (Table 13 and Table 15), it is not the grand suite of such genes that would have been expected. Additionally, the up-regulation of storage proteins in the maternal excess is almost counter to the hypothesis (Table 14). However, perhaps it is a naïve assumption that this is the only way by which the paternal genome might manifest its

acquisition of maternal resources. In fact, it is reported that in maize the ultimate seed mass is not determined by the amount of nutrients loaded per cell, as an upper limit of cell loading capacity appears to exist. Instead, heavier grains are generally associated with a greater number of cells (Reddy and Daynard, 1983). Hence, if it is considered that the mechanism by which the mother controls endosperm development is by the rate of differentiation, and the time of initiation of loading of resources, whereas the paternal influence is to maximize cell proliferation prior to loading, the data presented here lend real support to the theory. This work provides a demonstration of a much higher rate of mitosis and the production of many more cells in the paternal excess cross, but the opposite in maternal excess crosses. This situation is analogous to arabidopsis, where it was seen that paternal excess endosperm show a higher rate of nuclear proliferation, with more cells ultimately to account for the larger size, in contrast, the smaller maternal excess seeds had reduced cellular proliferation (Scott et al., 1998). Furthermore, the transcriptome of paternal excess endosperms appears to be more focussed towards proliferation, and less towards cellular development. For example, the adgeminal base of the paternal excess endosperms often develops into a mitotically active tumour-like region, whereas normally the endosperm would have long since ceased mitotic activity and be embarking on a period of programmed cell death. This suggests an inherently proliferative nature conferred by the paternal genome. Such a scenario is not restricted to plants however, as it has been demonstrated in mouse cells that cell lines with an exclusively paternally derived genome composition have a shorter cell cycle, and show spontaneous tumour formation, whereas those with just female genomes show the opposite effects (Hernandez et al., 2003).

The success of an interploidy cross is assessed by the number of recoverable seedlings, and in arabidopsis it is very high in crosses between diploids and tetraploids, but low in crosses between diploids and hexaploids (Scott et al., 1998). In maize, very few

seedlings can be recovered, although it is reported in this study that a few percentage can be recovered (Section 3.2.5). It is usually reported that the maternal excess yields a higher number of seedlings (Birchler, 1993; Nishiyama and Inomata, 1966). The sample size used in the current study was not large enough to see whether the larger proportion in the paternal excess was contradictory to this assumption. However, based on histological work, the paternal excess actually very seldom produces an embryo (Section 4.2.14), so the frequency of seedlings recovered from kernels may be higher for the mother, but a considerably higher proportion of those embryos which develop from paternal excess crosses are geminable. This might have implications on the abilities of the different endosperms on supporting the development of the embryo, which might be connected with nutrient allocation between the mother plant, endosperm and embryo.

6.4 Paternal excess confers genomic stress on developing endosperm

Perhaps the most novel insight into interploidy crosses obtained in this work is the apparent mis-expression of genes in paternal excess crosses, which is first observed at 10 DAP with the unusual expression of the pBETL::GUS marker. It becomes increasingly severe throughout development and is also observed for the pVP1::GUS reporter and an endogenous gene apparently involved in anthocyanin biosynthesis. Such irregular ectopic expression is similar to events which follow the formation of some interspecific hybrids, and has been proposed to result from a form of 'genomic shock' (Madlung and Comai, 2004) which might invoke the activation and random integration of transposable elements, as also occurs in some inter-specific crosses, with a resulting mis-expression of genes into which they integrate. It is therefore exciting to see that in some cases an up-regulation of several retrotransposons was observed following paternal but not maternal excess crosses.

A potential explanation of such phenomena might be provided by the differential dosage hypothesis proposed by Dilkes and Comai (Dilkes and Comai, 2004). In this they broadened the concept of parental conflict effects to not just imprinted genes (which show exclusively mono-allelic expression), but also those genes displaying other parental effects. Their proposal accepts that many genes are parentally differentially expressed in the seed, whereas only few genes have been shown to be truly imprinted. Many of these differentially expressed genes are involved in regulatory complexes, which as according to Birchler (Birchler et al., 2001), may be involved in 'stoichiometric complexes' (i.e. those requiring strictly regulated ratios between components). Hence, if a differentially expressed (or truly imprinted) regulator is provided by one parent, an interploidy cross would alter the stoichiometry with the partners of its complex, disrupting its function. Since DNA could be a component of such a complex, then a maternally expressed regulator which shows dosage effects would be 'precipitated out' by the additional DNA molecules supplied by a paternal excess cross. If such regulators were involved in silencing of genes or transposons, mis-expression of that gene, or retrotransposon activation could occur. In this scenario, the same would not be true of maternal excess crosses, since there would be an excess of the dosage dependent gene for each DNA molecule. Given that the majority of dosage-dependent genes are silenced on the paternal side, this is quite plausible. An extension of this hypothesis is that perhaps any gene might be prone to ectopic expression randomly through the endosperm, and the ones observed in this thesis are merely those which were conspicuous. Clearly it would be interesting to see if *in situ* localization of randomly selected genes are found throughout the endosperm. Preliminary RT-PCR analysis to see whether genes apparently not expressed in the endosperm can be detected in late stage paternal excess endosperm have also been promising.

6.5 General description of interploidy crosses

The biological observations reported in this work support those reported by Cooper (Cooper, 1951), and many of the observations which he made through presumably laborious and time-consuming histology have been more rapidly confirmed using effective visual markers. For example, his suggestion that the maternal excess endosperms show the development of a highly reduced 'absorbing tissue' near the suspensor is revealed without the need for any microscopy, as blue staining clearly marks this suppressed domain in a low percentage of kernels. The current study, however, extends this work to correlate biological events occurring in the seed with the establishment and maturation of two seed domains, as reported by two specific transcripts. A general summary of the data presented in this chapter is presented in Table 22.

	Paternal Excess	**Maternal Excess**
General observations	Delayed differentiation of aleurone and CSE. Aleurone is irregular and found on ectopic surfaces. BETL appears to form aberrantly. Delayed starch accumulation. Highly proliferative, condensed nuclei. Rarely produces an embryo. Central region disintegrates. Proliferative zone forms in adgerminal margin. Mis-expression of genes from 10 DAP onwards. High variability of seed size	Premature differentiation of aleurone and CSE. Aleurone cell morphology is slightly distorted, and layer invades BETL region. Highly reduced BETL domain. Accelerated starch accumulation. Early cessation of mitosis. Embryo usually present but often delayed. Seed size relatively regular. Endosperm gradually shrinks.
pVp1::GUS	Throughout periphery, even where misshapen.	Throughout periphery. Invades region of BETL
pBETL::GUS	Starts as irregular clusters of a few cells, predominately in BETL region. Later becomes more widespread throughout whole endosperm.	Consistently found in reduced domain at basal germinal corner, but low frequency.

Table 22 Summary of effects of interploidy crosses

6.6 Future Prospects

This work has produced a series of candidate genes which may be important in the regulation of endosperm development, and as such could be worthy of further investigation. Variability between microarray hybridisations suggests that in order to identify the consistent, 'real' effects of interploidy crosses (rather than stochastic events), a considerably higher number of replicate experiments will be needed than

were possible in this work. Additionally, one of the most powerful uses of microarray technology is to apply clustering techniques to profiles of gene expression. The small number of arrays used in this work, and the single time-point investigated, did not lend itself to such analysis, future analyses should sensibly look at developmental profiles through a number of timepoints to enable the identification of similarly regulated genes. Another possibility would be to perform crosses between hexaploid, tetraploid, and diploid lines (similar to arabidopsis experiments (Scott et al., 1998)). This would produce a series of endosperm genome ratios which would be large enough for clustering analysis (Figure 6-1), and would provide a very powerful discriminator between parentally-controlled effects and more sporadic and random events. Obviously, crosses between diploids and hexaploids would be predicted to be very severe, however, endosperm tissue was recoverable from interploidy crosses until as late as 30 DAP, so it is likely that some developmental program commences following these crosses. Future microarray experiments would also benefit from performing a hybridization of DNA extracted from the same samples for which RNA comparisons are being made. This technique should facilitate the identification of any genes which appear to be increased or decreased in copy number, as might occur in tetraploids , and to relate this to the expression levels of each transcript. This approach might also identify whether there are any regions of the chromosome are preferentially replicated in endoreduplication and whether this affects their levels of gene expression (Leiva-Neto et al., 2004).

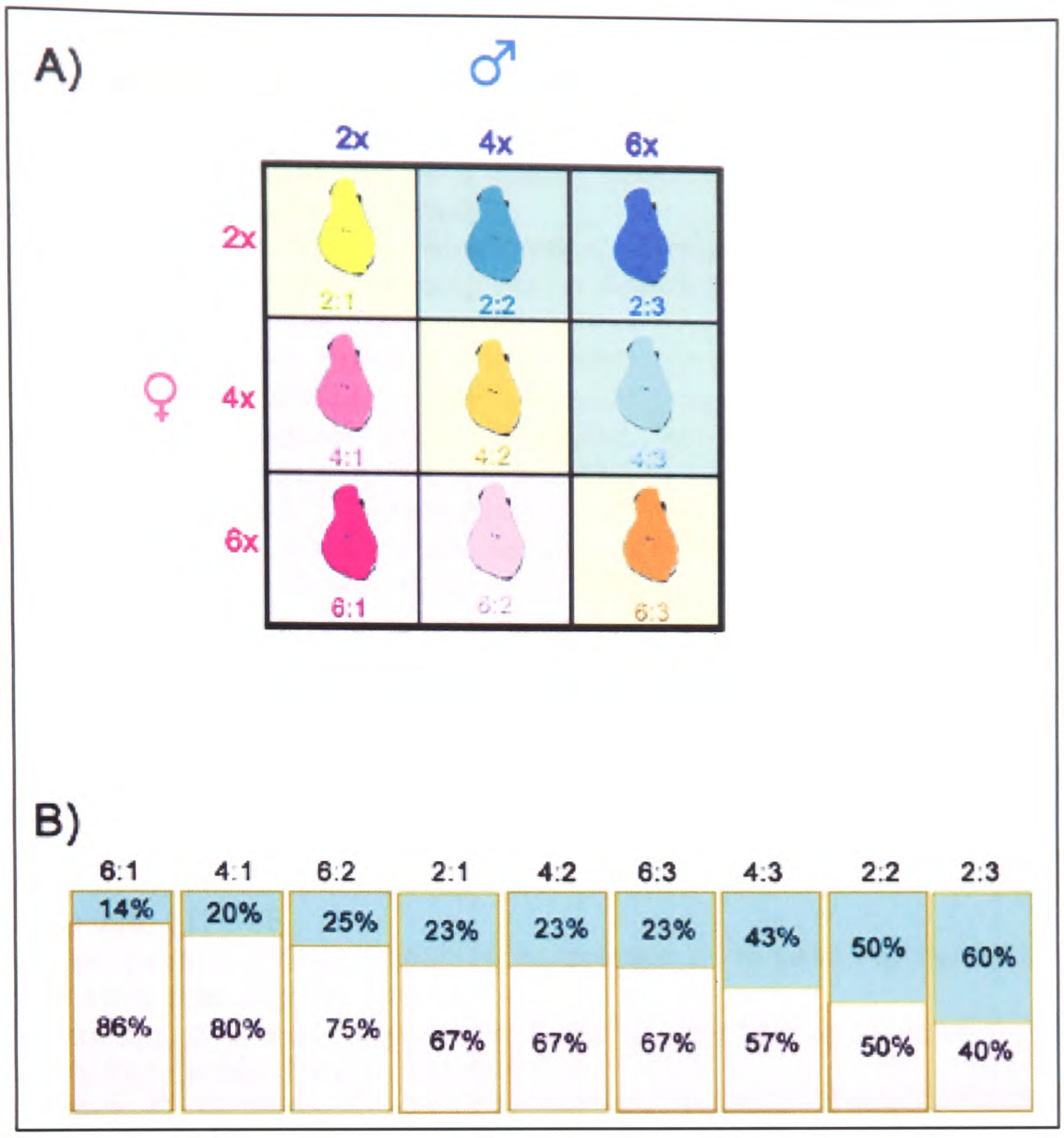


Figure 6-19 Plan of interploidy series for future microarray investigation

A) Details a table of cross directions, paternal ploidy along the top, and maternal down the left. The resultant endosperms at the intersect demonstrates the absolute parental genome ratio in the resultant endosperm. **B)** Table details the proportion of maternal : paternal genomes in the endosperms generated from table A, expressed as a percentage (blue = paternal, pink = maternal).

7 References

Aalfs, J. D. and Kingston, R. E. (2000). What does `chromatin remodelling` mean? *Trends in Biochemical Sciences* **25**, 548-555.
Adams, S., Vinkenoog, R., Spielman, M., Dickinson, H. G. and Scott, R. J. (2000). Parent-of-origin effects on seed development in *Arabidopsis thaliana* require DNA methylation. *Development* **127**, 2493-502.
Ahringer, J. (2000). NuRD and SIN3. *Trends in Genetics* **16**, 351-356.
Alleman, M. and Doctor, J. (2000). Genomic imprinting in plants: observations and evolutionary implications. *Plant Mol Biol* **43**, 147-61.
Altschul, S. F., Madden, T. L., Schaeffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389-3402.
Armstrong, C. L., Green, C. M. and Phillips, R. L. (1991). Development and availability of germplasm with high Type II culture formation response. *Maize Genetics Cooperative Newsletter* **65**, 92-3.
Arumuganathan, K. and Earle, E. D. (1991). Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter* **9**, 208-218.
Auger, D. L. and Birchler, J. A. (2002). Maize tertiary trisomic stocks derived from B-A translocations. *J Hered* **93**, 42-7.
Baggerly, K., Mitra, R., Grier, R., Medhane, D., Lozano, G. and Kapoor, M. (2004). Comparison of sample-labelling techniques in DNA microarray experiments. *Analytica Chimica Acta* **506**, 117-125.
Balkanjieva, J. (1983). Hybridization between autotriploid and diploid forms of barley. *Genetics and Plant Breeding* **16**, 43-51.
Bannister, A. J. and Kouzarides, T. (1996). The CBP co-activator is a histone acetyltransferase. *Nature*, 641-643.
Baroux, C., Spillane, C. and Grossniklaus, U. (2002). Evolutionary origins of the endosperm in flowering plants. *Genome Biology* **3**, reviews1026.1 - reviews1026.5.
Bate, N. J., Niu, X., Wang, Y., Reimann, K. S. and Helentjaris, T. G. (2004). An Invertase Inhibitor from Maize Localizes to the Embryo Surrounding Region during Early Kernel Development. *Plant Physiology* **134**, 246-254.
Becker, H. A., Hueros, G., Maitz, M., Serna, A. and Thompson, R. (1998). Domains of gene expression in developing endosperm. In *The molecular and cell biology of Sexual Plant Reproduction*, (ed. M. Cresti), pp. 361-375. Berlin: Springer Verlag.
Becraft, P. W. (2001). Cell fate specification in the cereal endosperm. *Semin Cell Dev Biol* **12**, 387-94.
Becraft, P. W. and Asuncion-Crabb, Y. (2000). Positional cues specify and maintain aleurone cell fate in maize endosperm development. *Development* **127**, 4039-48.
Becraft, P. W., Kang, S. H. and Suh, S. G. (2001). The maize CRINKLY4 receptor kinase controls a cell-autonomous differentiation response. *Plant Physiol* **127**, 486-96.
Becraft, P. W., Stinard, P. S. and McCarty, D. R. (1996). CRINKLY4: A TNFR-Like Receptor Kinase Involved in Maize Epidermal Differentiation. *Science*, 1406-1408.
Bedinger, P. and Russell, S. D. (1993). Gametogenesis in maize. In *The Maize Handbook*, (ed. M. Freeling and V. Walbot), pp. 759. New York: Springer Verlag.
Bennetzen, J. L. (1996). The contributions of retroelements to plant genome organization, function and evolution. *Trends in Microbiology* **4**, 347-353.
Bennetzen, J. L., Chandler, V. L. and Schnable, P. (2001). National Science Foundation-Sponsored Workshop Report. Maize Genome Sequencing Project. *Plant Physiology* **127**, 1572-1578.
Berger, F. (2003). Endosperm: the crossroad of seed development. *Current Opinion in Plant Biology* **6**, 42-50.
Birchler, J. (1980). On the nonautonom of the small-kernel phenotype produced by B-A translocations in maize. *Genetical Research* **36**, 111-116.
Birchler, J. A. (1979). A study of enzyme activities in a dosage series of the long arm of chromosome one in maize. *Genetics* **92**, 1211-1229.
Birchler, J. A. (1993). Dosage analysis of maize endosperm development. *Annu Rev Genet* **27**, 181-204.

Birchler, J. A., Bhadra, U., Bhadra, M. P. and Auger, D. L. (2001). Dosage-dependent gene regulation in multicellular eukaryotes: implications for dosage compensation, aneuploid syndromes, and quantitative traits. *Dev Biol* **234**, 275-88.
Birchler, J. A., Chalfoun, D. J. and Levin, D. M. (1990). Recombination in the B chromosome of maize to produce A-B-A chromosomes. *Genetics* **126**, 723-33.
Birchler, J. A. and Newton, K. J. (1981). Modulation of protein levels in chromosomal dosage series of maise: the biochemical basis of aneploid syndromes. *Genetics* **99**, 247-266.
Birchler, J. A., Riddle, N. C., Auger, D. L. and Veitia, R. A. (2005). Dosage balance in gene regulation: biological implications. *Trends in Genetics* **21**, 219-226.
Bommert, P. and Werr, W. (2001). Gene expression patterns in the maize caryopsis: clues to decisions in embryo and endosperm development. *Gene* **271**, 131-42.
Bonello, J. F., Opsahl-Ferstad, H. G., Perez, P., Dumas, C. and Rogowsky, P. M. (2000). Esr genes show different levels of expression in the same region of maize endosperm. *Gene* **246**, 219-27.
Bonello, J. F., Sevilla-Lecoq, S., Berne, A., Risueno, M. C., Dumas, C. and Rogowsky, P. M. (2002). Esr proteins are secreted by the cells of the embryo surrounding region. *Journal of Experimental Botany* **53**, 1559-1568.
Bowe, L. M., Coat, G. and dePamphilis, C. W. (2000). Phylogeny of seed plants based on all three genomic compartments: Extant gymnosperms are monophyletic and Gnetales' closest relatives are conifers. *PNAS* **97**, 4092-4097.
Brannan, C. I. and Bartolomei, M. S. (1999). Mechanisms of genomic imprinting. *Current OPinion in Genetics and Development* **9**, 164-170.
Breitling, R., Amtmann, A. and Herzyk, P. (2004). Iterative Group Analysis (iGA): A simple tool to enhance sensitivity and facilitate interpretation of microarray experiments. *BMC Bioinformatics* **5**, 34.
Brink, R. A. and Cooper, D. C. (1947). Effect of the DE_{17} allele on development of the maize caryopsis. *Genetics* **32**, 350-368.
Brown, R. C., Lemmon, B. E. and Olsen, O. A. (1994a). Endosperm Development in Barley: Microtubule Involvement in the Morphogenetic Pathway. *Plant Cell* **6**, 1241.
Brown, R. C., Lemmon, B. E. and Olsen, O. A. (1994b). Endosperm Development in Barley: Microtubule Involvement in the Morphogenetic Pathway. *Plant Cell* **6**, 1241-1252.
Brunner, S., Fengler, K., Morgante, M., Tingey, S. and Rafalski, A. (2005). Evolution of DNA Sequence Nonhomologies among Maize Inbreds. *Plant Cell* **17**, 343-360.
Cai, G., Faleri, C., Del Casino, C., Hueros, G., Thompson, R. D. and Cresti, M. (2002). Subcellular localisation of BETL-1, -2 and -4 in Zea mays L. endosperm. *Sexual Plant Reproduction* **15**, 85-98.
Call, D. R., Chandler, D. P. and Brockman, F. (2001). Fabrication of DNA Microarrays Using Unmodified Oligonucleotide Probes. *Biotechniques* **30**, 368-379.
Cao, X. and Jacobsen, S. E. (2002). Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. *Proc Natl Acad Sci U S A.*
Casati, P. and Walbot, V. (2003). Gene Expression Profiling in Response to Ultraviolet Radiation in Maize Genotypes with Varying Flavonoid Content. *Plant Physiology* **132**, 1739-1754.
Causton, H. C., Brazma, A. and Quackenbush, J. (2003). Image processing, normalisation and data transformation. In *Begineer's Guide: Microarray Gene Expression Data Analysis*, pp. 40-71. Blackwell Science.
Chandler, V. L., Eggleston, W. B. and Dorweiler, J. E. (2000). Paramutation in maize. *Plant Mol Biol* **43**, 121-45.
Charlton, W. L., Keen, C. L., Merriman, C., Lynch, P., Greenland, A. J. and Dickinson, H. G. (1995). Endosperm development in Zea mays; implications of gametic imprinting and paternal excess in regulation of transfer layer development. *Development* **121**, 3089-3097.
Chaudhuri, S. and Messing, J. (1994). Allele-specific parental imprinting of dzr1, a posttranscriptional regulator of zein accumulation. *Proc Natl Acad Sci U S A* **91**, 4867-71.
Chaudhury, A. M., Koltunow, A., Payne, T., Luo, M., Tucker, M. R., Dennis, E. S. and Peacock, W. J. (2001). Control of Early Seed Development. *Annual Review of Cell and Developmental Biology* **17**, 677-700.
Chaudhury, A. M., Ming, L., Miller, C., Craig, S., Dennis, E. S. and Peacock, W. J. (1997). Fertilization-independent seed development in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* **94**, 4223-8.

Chaw, S.-M., Parkinson, C. L., Cheng, Y., Vincent, T. M. and Palmer, J. D. (2000). Seed plant phylogeny inferred from all three plant genomes: Monophyly of extant gymnosperms and origin of Gnetales from conifers. *PNAS* **97**, 4086-4091.
Cheng, P.-C. and Pareddy, D. R. (1994). Morphology and development of the tassel and ear. In *The Maize Handbook*, (ed. M. Freeling and V. Walbot), pp. 759. New York: Springer Verlag.
Cheng, W.-H., Taliercio, E. W. and Chourey, P. S. (1996). The *Miniature1* Seed locus of maize encodes a cell wal invertase required for normal development of endosperm and maternal cells in the pedicel. *Plant Cell* **8**, 971-983.
Cheung, A. Y. and Wu, H.-m. (2001). PLANT BIOLOGY: Pollen Tube Guidance--Right on Target. *Science* **293**, 1441-1442.
Cho, Y., Fernandes, J., Kim, S.-H. and Walbot, V. (2002). Gene-expression profile comparisons distinguish seven organs of maize. *Genome Biology* **3**, research0045.1 - research0045.16.
Choi, Y., Gehring, M., Johnson, L., Hannon, M., Harada, J. J., Goldberg, R. B., Jacobsen, S. E. and Fischer, R. L. (2002). DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in arabidopsis. *Cell* **110**, 33-42.
Chou, C. C., Chen, C. H., Lee, T. T. and Peck, K. (2004). Optimization of probe length and the number of probes per gene for optimal microarray analysis of gene expression. *Nucleic Acids Research* **32**, e99-e99.
Churchill, G. A. (2002). Fundamentals of experimental design for cDNA microarrays. *Nature Genetics* **32**, 490-495.
Coe, E. H. (2001). The origins of maize genetics. *Nature Reviews Genetics* **2**, 898-905.
Colombo, L., Franken, J., Van der Krol, A. R., Wittich, P. E., Dons, H. and Angenent, G. C. (1997). Downregulation of Ovule-Specific MADS Box Genes from Petunia Results in Maternally Controlled Defects in Seed Development. *Plant Cell* **9**, 703-715.
Cooper, D. C. (1951). Caryopsis development following matings between diploid and tetraploid strains of *Zea mays*. *American Journal of Botany* **38**, 702-708.
Cooper, D. C. and Brink, R. A. (1945). Seed collapse following matings between diploid and tetraploid races of *Lycopersicon pimpinellifolium*. *Genetics* **30**, 376-401.
Costa, L. M., Gutierrez-Marcos, J. F., Brutnell, T. P., Greenland, A. J. and Dickinson, H. G. (2003). The globby1-1 (glo1-1) mutation disrupts nuclear and cell division in the developing maize seed causing alterations in endosperm cell fate and tissue differentiation. *Development* **130**, 5009-5017.
Costa, L. M., Gutierrez-Marcos, J. F. and Dickinson, H. G. (2004). More than a yolk: the short life and complex times of the plant endosperm. *Trends in Plant Science* **9**, 507-514.
Dangl, M., Brosch, G., Haas, H., Loidl, P. and Lusser, A. (2001). Comparative analysis of HD2 type histone deacetylases in higher plants. *Planta* **213**, 280-285.
Danilevskaya, O. N., Hermon, P., Hantke, S., Muszynski, M. G., Kollipara, K. and Ananiev, E. V. (2003). Duplicated fie Genes in Maize: Expression Pattern and Imprinting Suggest Distinct Functions. *Plant Cell* **15**, 425-438.
Davis, R. W. and Smith, J. D. (1990). A light and electron microscope investigation of the transfer cell regions of maize caryopses. *Can. J. Bot.* **68**, 471-479.
Dechiara, T. M., Efstratiadis, A. and Robertson, E. J. (1990). A Growth-Deficiency Phenotype in Heterozygous Mice Carrying an Insulin-Like Growth Factor-Ii Gene Disrupted by Targeting. *Nature* **345**, 78-80.
Diboll, A. G. and Larson, D. A. (1966). An electron microscopic study of the mature megagametophyte in *Zea mays*. *Am J Bot* **53**.
Diehl, F., Grahlmann, S., Beier, M. and Hoheisel, J. D. (2001). Manufacturing DNA microarrays of high spot homogeneity and reduced background signal. *Nucl. Acids Res.* **29**, e38-.
Dilkes, B. P. and Comai, L. (2004). A Differential Dosage Hypothesis for Parental Effects in Seed Development. *Plant Cell* **16**, 3174-3180.
Doan, D. N., Linnestad, C. and Olsen, O. A. (1996). Isolation of molecular markers from the barley endosperm coenocyte and the surrounding nucellus cell layers. *Plant Mol Biol* **31**, 877-86.
Dolezcaronel, J., Binarova, P. and Lucretti, S. (1989). Analysis of nuclear DNA content in plant cells by flow cytometry. *Biologia Plantarum* **31**, 113-120.
Dolezcaronel, J., Greilhuber, J., Lucretti, S., Meister, A., Lysak, M. A., Nardi, L. and Obermayer, R. (1998). Plant Genome Size Estimation by Flow Cytometry: Inter-laboratory Comparison. *Ann Bot* **82**, 17-26.

Duvick, D. N. (2001). Biotechnology in the 1930s: the development of hybrid maize. *Nature Reviews Genetics* **2**, 69-73.
Engel, M. L., Chaboud, A., Dumas, C. and McCormick, S. (2003). Sperm cells of Zea mays have a complex complement of mRNAs. *Plant Journal* **34**, 697-708.
Esen, A. and Soost, R. K. (1973). Seed development in Citrus with special reference to 2X x 4X crosses. *American Journal of Botany* **60**, 448-462.
Federoff, N. V. (1983). Controlling elements in maize. In *Mobile Genetic Elements*, (ed. J. Shapiro), pp. 1-63. New York: Academic Press.
Feldman, A. L., Costouros, N. G., Wang, E., Qian, M., Marincola, F. M., Alexander, H. R. and Libutti, S. K. (2002). Advantages of mRNA Amplification for Microarray Analysis. *Biotechniques* **33**, 906-915.
Fernandes, J., Brendel, V., Gai, X., Lal, S., Chandler, V. L., Elumalai, R. P., Galbraith, D. W., Pierson, E. A. and Walbot, V. (2002). Comparison of RNA Expression Profiles Based on Maize Expressed Sequence Tag Frequency Analysis and Micro-Array Hybridization. *Plant Physiology* **128**, 896-910.
Fernandes, J., Dong, Q., Schneider, B., Morrow, D. J., Nan, G. L., Brendel, V. and Walbot, V. (2004). Genome-wide mutagenesis of Zea mays L. using RescueMu transposons. *Genome Biology* **5**, 2004/5/10/R82.
Finnegan, E. J. and Kovac, K. A. (2000). Plant DNA methyltransferases. *Plant Molecular Biology* **43**, 189-201.
Francis, N. J. and Kingston, R. E. (2001). Mechanisms of transcriptional memory. *Nature Reviews Molecular Cell Biology* **2**, 409-421.
Friedman, W. E. (1992). Evidence of a pre-angiosperm origin of endopserm: implications for the evolution of flowering plants. *Science* **255**, 336-339.
Friedman, W. E. and Williams, J. H. (2004). Developmental Evolution of the Sexual Process in Ancient Flowering Plant Lineages. *Plant Cell* **16**, S119-S132.
Fu, H. and Dooner, H. K. (2002). Intraspecific violation of genetic colinearity and its implications in maize. *Proceedings- National Academy of Sciences USA* **99**, 9573-9578.
Galbraith, D. W. H., Kristi R.; Maddox, Joyce M.; Ayres, Nicola M.; Sharma, Dharam P.; Firoozabady, Ebrahim SO. (1983). Rapid Flow Cytometric Analysis of the Cell Cycle in Intact Plant Tissues. *Science* **220**, 1049-1051.
Gasser, C. S., Broadhvest, J. and Hauser, B. A. (1998). Genetic Analysis of Ovule Development. *Annual Review of Plant Physiology and Plant Molecular Biology* **49**, 1-24.
Gehring, M., Choi, Y. and Fischer, R. L. (2004). Imprinting and Seed Development. *Plant Cell* **16**, S203-S213.
Gibbon, B. C. and Larkins, B. A. (2005). Molecular genetic approaches to developing quality protein maize. *Trends in Genetics* **21**, 227-233.
Gomez, E., Royo, J., Guo, Y., Thompson, R. and Hueros, G. (2002). Establishment of cereal endosperm expression domains: identification and properties of a maize transfer cell-specific transcription factor, ZmMRP-1. *Plant Cell* **14**, 599-610.
Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E. M. and Coupland, G. (1997). A polycomb-group gene regulates homeotic gene expression in Arabidopsis. *Nature* **386**, 44-51.
Graessle, S., Loidl, P. and Brosch, G. (2001). Histone acetylation: plants and fungi as model systems for the investigation of histone deacetylases. *Cellular and Molecular Life Sciences* **58**, 704-720.
Grant-Downton, R. T. and Dickinson, H. G. (2005). Epigenetics and its implications for plant biology. Part 1 - The epigenetic network in plants. *Annals of Botany* **Submitted**.
Grimanelli, D., Perotti, E., Ramirez, J. and Leblanc, O. (2005). Timing of the Maternal-to-Zygotic Transition during Early Seed Development in Maize. *Plant Cell* **17**, 1061-1072.
Grossniklaus, U., Nogler, G. A. and van Dijk, P. J. (2001). How to avoid sex: the genetic control of gametophytic apomixis. *Plant Cell* **13**, 1491-8.
Grossniklaus, U., Vielle-Calzada, J.-P., Hoeppner, M. A. and Gagliano, W. B. (1998). Maternal Control of Embryogenesis by MEDEA, a Polycomb Group Gene in Arabidopsis. *Science* **280**, 446-450.
Guignard, L. (1899). Sur les antherozoides et la double copulation sexualle chez les vegetaux angiosperms. *Crit. Rev. Acad. Sci. Paris* **128**, 864-871.
Guivarch, A., Caissard, J. C., Azmi, A., Elmayan, T., Chriqui, D. and Tepfer, M. (1996). In situ detection of expression of the gus reporter gene transgenic plants: Ten years of blue genes. *Transgenic Research* **5**, 281-288.

Guo, M. and Birchler, J. B. (1994). Trans-acting dosage effects on the expression of model gene systems in maize aneuploids. *Science* **266**, 1999-2002.
Guo, M., Davis, D. and Birchler, J. A. (1996). Dosage effects on gene expression in a maize ploidy series. *Genetics* **142**, 1349-55.
Gupta, V., Cherkassky, A., Chatis, P., Joseph, R., Johnson, A. L., Broadbent, J., Erickson, T. and DiMeo, J. (2003). Directly labeled mRNA produces highly precise and unbiased differential gene expression data. *Nucleic Acids Res* **31**, e13.
Gurdon, J. B., Mitchell, A. and Mahoney, D. (1995). Direct and continuous assessment by cells of their position in a morphogen gradient. *Nature*, 520.
Gutierrez-Marcos, J. F., Costa, L. M., Biderre-Petit, C., Khbaya, B., O Sullivan, D. M., Wormald, M., Perez, P. and Dickinson, H. G. (2004). maternally expressed gene1 Is a Novel Maize Endosperm Transfer Cell-Specific Gene with a Maternal Parent-of-Origin Pattern of Expression. *Plant Cell* **16**, 1288-1301.
Haan, N. (2004). A Bayesian appraoch to array image variability. *Microarray Technology*, 51-59.
Haig, D. (2000). The Kinship Theory of Genomic Imprinting. *Annual Review of Ecology and Systematics* **31**, 9-32.
Haig, D. and Westoby, M. (1989). Parent-specific gene expression and the triploid endosperm. *American Naturalist* **134**, 147-155.
Haig, D. and Westoby, M. (1991). Genomic Imprinting in endosperm: its effect on seed development in corsses between species, and between different ploidies of the same species, and its implications for the evolution of apomixis. *Philosophical Transactions of the Royal Society of London* **333**.
Hames, B. D. and Higgins, S. J. (1985). Nucleic acid hybridisation. Oxford, Washington DC: IRL Press.
Haseloff, J. (1999). GFP variants for multispectral imaging of living cells. In *Methods in Cell Biology, Vol 58*, vol. 58, pp. 139-+.
Hasseman, J. (2001). Aminoallyl labeling of RNA for microarrays, vol. 2005 (ed.: The Institue of Genomic Research.
Havas, K. and Whitehouse, I.
Hegde, P., Qi, R., Abernathy, K., Gay, C., Dharap, S., Gaspard, R., Hughes, J. E., Snesrud, E., Lee, N. and Quackenbush, J. (2000). A Concise Guide to cDNA Microarray Analysis. *Biotechniques* **29**, 548-563.
Henikoff, S. and Comai, L. (1998). A DNA methyltransferase homolog with a chromodomain exists in multiple polymorphic forms in Arabidopsis. *Genetics* **149**, 307-318.
Hernandez, L., Kozlov, S., Piras, G. and Stewart, C. L. (2003). Paternal and maternal genomes confer opposite effects on proliferation, cell-cycle length, senescence, and tumor formation. *Proceedings- National Academy of Sciences USA* **100**, 13344-13349.
Hessner, M. J., Wang, X., Katie, H., Meyer, L., Yan, W., Nye, S., Guo, S.-W. and Ghosh, S. (2003). Three color cDNA microarrays: quantitative assessment through the use of fluorescein-labeled probes. *Nucleic Acids Res* **31**, e14.
Higashiyama, T., Yabe, S., Sasaki, N., Nishimura, Y., Miyagishima, S.-y., Kuroiwa, H. and Kuroiwa, T. (2001). Pollen Tube Attraction by the Synergid Cell. *Science* **293**, 1480-1483.
Holloway, A. J., van Laar, R. K., Tothill, R. W. and Bowtell, D. D. L. (2002). Options available-from start to finish-for obtaining data from DNA microarrays II. *Nature Genetics* **32**, 481-489.
Horn, P. J. and Peterson, C. L. (2002). MOLECULAR BIOLOGY: Chromatin Higher Order Folding--Wrapping up Transcription. *Science* **297**, 1824-1827.
Hu, L., Wang, J., Baggerly, K., Wang, H., Fuller, G., Hamilton, S., Coombes, K. and Zhang, W. (2002). Obtaining reliable information from minute amounts of RNA using cDNA microarrays. *BMC Genomics* **3**, 16.
Huang, B.-Q. and Russell, D. A. (1992). Female germ unit: organization, isolation and function. *International Review of Cytology* **140**, 233-293.
Huber, M., Wei, T. F., Muller, U. R., Lefebvre, P. A., Marla, S. S. and Bao, Y. P. (2004). Gold nanoparticle probe-based gene expression analysis with unamplified total human RNA. *Nucleic Acids Research* **32**, e137-e137.
Hueros, G., Gomez, E., Cheikh, N., Edwards, J., Weldon, M., Salamini, F. and Thompson, R. D. (1999a). Identification of a Promoter Sequence from the BETL1 Gene Cluster Able to Confer Transfer-Cell-Specific Expression in Transgenic Maize. *Plant Physiology* **121**, 1143-1152.

Hueros, G., Royo, J., Maitz, M., Salamini, F. and Thompson, R. D. (1999b). Evidence for factors regulating transfer cell-specific expression in maize endosperm. *Plant Molecular Biology* **41**, 403-414.
Hueros, G., Varotto, S., Salamini, F. and Thompson, R. D. (1995). Molecular characterization of BET1, a gene expressed in the endosperm transfer cells of maize. *Plant Cell* **7**, 747-57.
Hughes, T. R., Mao, M., Jones, A. R., Burchard, J., Marton, M. J., Shannon, K. W., Lefkowitz, S. M., Ziman, M., Schelter, J. M. and Meyer, M. R. (2001). Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nature Biotechnology* **19**, 342-347.
Hulskamp, M., Schneitz, K. and Pruitt, R. E. (1995). Genetic Evidence for a Long-Range Activity That Directs Pollen Tube Guidance in Arabidopsis. *Plant Cell* **7**, 57-64.
Hurst, L. D. and McVean, G. T. (1997). Growth effects of uniparental disomies and the conflict theory of genomic imprinting. *Trends in Genetics* **13**, 436-443.
Imai, S.-i., Armstrong, C. M., Kaeberlein, M. and Guarente, L. (2000). Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. *Nature* **403**, 795-800.
Jacobsen, S. E., Sakai, H., Finnegan, E. J., Cao, X. and Meyerowitz, E. M. (2000). Ectopic hypermethylation of flower-specific genes in Arabidopsis. *Current Biology* **10**, 179-186.
Jeddeloh, J. A., Stokes, T. L. and Richards, E. J. (1999). Maintenance of genomic methylation requires a SWI2/SNF2-like protein. *Nature Genetics* **22**, 94-97.
Jones, D. O., Cowell, I. G. and Singh, P. B. (2000). Mammalian chromodomain proteins: their role in genome organisation and expression. *Bioessays* **22**, 124-37.
Kaneko-Ishino, T., Kohda, T. and Ishino, F. (2003). The Regulation and Biological Significance of Genomic Imprinting in Mammals. *J Biochem (Tokyo)* **133**, 699-711.
Kankel, M. W., Ramsey, D. E., Stokes, T. L., Flowers, S. K., Haag, J. R., Jeddeloh, J. A., Riddle, N. C., Verbsky, M. L. and Richards, E. J. (2003). Arabidopsis MET1 cytosine methyltransferase mutants. *Genetics* **163**, 1109-1122.
Kato, A. (1997). An improved method for chromosome counting in maize. *Biotech Histochem* **72**, 249-252.
Kato, A. (1998). Nitrous oxide (N~2O) is effective for chromosome counting in maize. *Maize Genetics Cooperation Newsletter* **72**, 32.
Kato, A. (2002). Chromosome doubling of haploid maize seedlings using nitrous oxide gas at the flower primordial stage. *Plant Breeding* **121**, 370-377.
Kellogg, E. A. (2001). Evolutionary History of the Grasses. *Plant Physiology* **125**, 1198-1205.
Kermicle, J. L. (1970). Dependence of the R-mottled aleurone phenotype in maize on mode os sexual transmisson. *Genetics* **66**, 69-85.
Kermicle, J. L. (1978). Imprinting of gene action in the maize endosperm. In *Maize breeding and genetics*, (ed. D. B. Walden), pp. 357-371. New York: John Wiley.
Kiesselbach, T. A. (1949a). The structure and reproduction of corn. *Nebraska Agricultural Experimental Station Research Bulletin* **161**, 1-96.
Kiesselbach, T. A. (1949b). The structure and reproduction of corn. *Nebraska Agricultural Experimental Station Research Bulletin*.
Kinoshita, T., Yadegari, R., Harada, J. J., Goldberg, R. B. and Fischer, R. L. (1999). Imprinting of the MEDEA polycomb gene in the Arabidopsis endosperm. *Plant Cell* **11**, 1945-52.
Kiyosue, T., Ohad, N., Yadegari, R., Hannon, M., Dinneny, J., Wells, D., Katz, A., Margossian, L., Harada, J. J., Goldberg, R. B. et al. (1999). Control of fertilization-independent endosperm development by the MEDEA polycomb gene in Arabidopsis. *Proc Natl Acad Sci U S A* **96**, 4186-91.
Knox, R. B. and Heslop-Harrison, J. (1970). Direct demonstration of the low permeability of the angiosperm meiotic tetrad using a fluorgenic ester. *Z Pflanzenphsiol Bd* **62**, 451-459.
Kouzarides, T. (2002). Histone methylation in transcriptional control. *Current OPinion in Genetics and Development* **12**, 198-209.
Kowles, R. V. and Phillips, R. L. (1988). Endosperm development in maize. *International Review of Cytology* **112**, 97-136.
Kranz, E., Loerz, H., Digonnet, C. and Faure, J. E. (1992). In Vitro Fusion of Gametes and Production of Zygotes. *International Review of Cytology*, 407.
Kranz, E., von Wiegen, P., Quader, H. and Lorz, H. (1998). Endosperm development after fusion of isolated, single maize sperm and central cells in vitro. *Plant Cell* **10**, 511-24.

Kuo, W. P., Jenssen, T. K., Butte, A. J., Ohno-Machado, I. and Kohane, I. S. (2002). Analysis of matched mRNA measurements from two different microarray technologies. *Bioinformatics* **18**, 405-412.
Lachner, M. and Jenuwein, T. (2002). The many faces of histone lysine methylation. *Current Opinion in Cell Biology* **14**, 286-298.
Larkins, B., Dilkes, B. P., Dante, R. A., Coelho, C. M., Woo, Y. and Liu, Y. (2001). Investigating the hows and whys of DNA endoreduplication (Review article). *Journal of Experimental Botany* **52**, 183-192.
Lauria, M., Rupe, M., Guo, M., Kranz, E., Pirona, R., Viotti, A. and Lund, G. (2004). Extensive Maternal DNA Hypomethylation in the Endosperm of Zea mays. *Plant Cell* **16**, 510-522.
Laurie, D. A. and Bennett, M. D. (1985). Nuclear DNA content in the genera Zea and Sorghum. INtergeneric, interspecific, and intraspecific variation. *Heredity* **55**, 307-313.
Le, Q., Gutierrez-Marcos, J. F., Costa, L. M., Meyer, S., Dickinson, H. G., Lorz, H., Kranz, E. and Scholten, S. (2005). Construction and screening of subtracted cDNA libraries from limited populations of plant cells: a comparative analysis of gene expression between maize egg cells and central cells. *The Plant Journal* **44**, 167-178.
Leblanc, O., Pointe, C. and Hernandez, M. (2002). Cell cycle progression during endosperm development in Zea mays depends on parental dosage effects. *Plant Journal* **32**, 1057-1066.
Lee, H.-S., Wang, J., Tian, L., Jiang, H., Black, M. A., Madlung, A., Watson, B., Lukens, L., Pires, J. C., Wang, J. J. et al. (2004). Sensitivity of 70-mer oligonucleotides and cDNAs for microarray analysis of gene expression in Arabidopsis and its related species. *Plant Biotechnology Journal* **2**, 45-57.
Leiva-Neto, J. T., Grafi, G., Sabelli, P. A., Dante, R. A., Woo, Y. m., Maddock, S., Gordon-Kamm, W. J. and Larkins, B. A. (2004). A Dominant Negative Mutant of Cyclin-Dependent Kinase A Reduces Endoreduplication but Not Cell Size or Gene Expression in Maize Endosperm. *Plant Cell.* **16**, 1854-1869.
Leung, Y. F. and Cavalieri, D. (2003). Fundamentals of cDNA microarray data analysis. *Trends in Genetics* **19**, 649-659.
Lid, S. E., Al, R. H., Krekling, T., Meeley, R. B., Ranch, J., Opsahl-Ferstad, H. G. and Olsen, O. A. (2004). The maize disorganized aleurone layer 1 and 2 (dil1, dil2) mutants lack control of the mitotic division plane in the aleurone layer of developing endosperm. *Planta* **218**, 370-378.
Lid, S. E., Gruis, D., Jung, R., Lorentzen, J. A., Ananiev, E., Chamberlin, M., Niu, X., Meeley, R., Nichols, S. and Olsen, O. A. (2002). The defective kernel 1 (dek1) gene required for aleurone cell development in the endosperm of maize grains encodes a membrane protein of the calpain gene superfamily. *Proc Natl Acad Sci U S A* **99**, 5460-5.
Lin, B.-Y. (1982). Association of endosperm reduction with parental imprinting in maize. *Genetics* **100**, 475-486.
Lin, B.-Y. (1984). Ploidy barrier to endosperm development in maize. *Genetics* **107**, 103-115.
Lipshutz, R. J., Fodor, S. P. A., Gingeras, T. R. and Lockhart, D. J. (1999). High density synthetic oligonucleotide arrays. *Nature Genetics* **21**, 20-24.
Lolle, S. J., Victor, J. L., Young, J. M. and Pruitt, R. E. (2005). Genome-wide non-mendelian inheritance of extra-genomic information in Arabidopsis. **434**, 505-509.
Lopes, M. A. and Larkins, B. A. (1993). Endosperm origin, development, and function. *Plant Cell* **5**, 1383-99.
Lund, G., Ciceri, P. and Viotti, A. (1995a). Maternal-specific demethylation and expression of specific alleles of zein genes in the endosperm of Zea mays L. *Plant J* **8**, 571-81.
Lund, G., Messing, J. and Viotti, A. (1995b). Endosperm-specific demethylation and activation of specific alleles of alpha-tubulin genes of Zea mays L. *Mol Gen Genet* **246**, 716-22.
Luo, M., Bilodeau, P., Dennis, E. S., Peacock, W. J. and Chaudhury, A. (2000). Expression and parent-of-origin effects for FIS2, MEA, and FIE in the endosperm and embryo of developing Arabidopsis seeds. *Proc Natl Acad Sci U S A* **97**, 10637-42.
Luo, M., Bilodeau, P., Koltunow, A., Dennis, E. S., Peacock, W. J. and Chaudhury, A. M. (1999). Genes controlling fertilization-independent seed development in Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **96**, 296-301.
Lusser, A., Brosch, G., Loidl, A., Haas, H. and Loidl, P. (1997). Identification of Maize Histone Deacetylase HD2 as an Acidic Nucleolar Phosphoprotein. *Science*, 88-90.
Lusser, A., Kölle, D. and Loidl, P. (2001). Histone acetylation: lessons from the plant kingdom. *Trends in Plant Science* **6**, 59-65.

Madlung, A. and Comai, L. (2004). The Effect of Stress on Genome Regulation and Structure. *Annals of Botany* **94**, 481-495.
Madlung, A., Masuelli, R. W., Watson, B., Reynolds, S. H., Davison, J. and Comai, L. (2002). Remodeling of DNA Methylation and Phenotypic and Transcriptional Changes in Synthetic Arabidopsis Allotetraploids. *Plant Physiology* **129**, 733-746.
Maere, S., Heymans, K. and Kuiper, M. (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks. *Bioinformatics* **21**, 3448-3449.
Magnard, J. L., Le Deunff, E., Domenech, J., Rogowsky, P. M., Testillano, P. S., Rougier, M., Risueno, M. C., Vergne, P. and Dumas, C. (2000). Genes normally expressed in the endosperm are expressed at early stages of microspore embryogenesis in maize. *Plant Mol Biol* **44**, 559-74.
Magnard, J. L., Lehouque, G., Massonneau, A., Frangne, N., Heckel, T., Gutierrez-Marcos, J. F., Perez, P., Dumas, C. and Rogowsky, P. M. (2003). ZmEBE genes show a novel, continuous expression pattern in the central cell before fertilization and in specific domains of the resulting endosperm after fertilization. *Plant Molecular Biology* **53**, 821-836.
Maitz, M., Santandrea, G., Zhang, Z., Lal, S., Hannah, L. C., Salamini, F. and Thompson, R. D. (2000). rgf1, a mutation reducing grain filling in maize through effects on basal endosperm and pedicel development. *Plant J* **23**, 29-42.
Mangelsdorf, P. C. (1974). Corn: Its Origin, Evolution, and Improvement. Cambridge, Massachusetts: Harvard University Press.
Martienssen, R. A., Rabinowicz, P. D., O'Shaughnessy, A. and McCombie, W. R. (2004). Sequencing the maize genome. *Current Opinion in Plant Biology* **7**, 102-107.
McCarty, D. R. and Carson, C. B. (1991). The molecular genetics of seed matuartion in maize. *Physiologia Plantarum* **81**, 267-272.
McCarty, D. R., Hattori, T., Carson, C. B., Vasil, V., Lazar, M. and Vasil, I. K. (1991). The Viviparous-1 developmental gene of maize encodes a novel transcriptional activator. *Cell* **66**, 895-905.
McClintock, B. (1978). Development of the maize endosperm as revealed by clones. In *The Clonal Basis of of Development.*, (ed. S. Subtelny and I. M. Sussex), pp. 217-237. New York: Academic Press Inc.
Messing, J., Bharti, A. K., Karlowski, W. M., Gundlach, H., Kim, H. R., Yu, Y., Wei, F. S., Fuks, G., Soderlund, C. A., Mayer, K. F. X. et al. (2004). Sequence composition and genome organization of maize. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 14349-14354.
Meyers, B. C., Tingey, S. V. and Morgante, M. (2001). Abundance, Distribution, and Transcriptional Activity of Repetitive Elements in the Maize Genome. *Genome Research* **11**, 1660-1676.
Miller, M. E. and Chourey, P. S. (1992). The Maize Invertase-Deficient miniature-1 Seed Mutation Is Associated with Aberrant Pedicel and Endosperm Development. *Plant Cell* **4**, 297-305.
Mizzen, C. A., Yang, X. J., Kokubo, T., Brownell, J. E., Bannister, A. J., Owen-Hughes, T., Workman, J., Wang, L., Berger, S. L. and Kouzarides, T. (1996). The TAF(II)250 Subunit of TFIID Has Histone Acetyltransferase Activity. *Cell* **87**, 1261-1270.
Moll, R. H., Lonnquist, J. H., Veley Fortuna, J. and Johnson, E. C. (1965). The relationship of heterosis and genetic divergence in maize. *Genetics* **52**, 139-144.
Muhitch, M. J. (1998). Characterization of pedicel beta-glucuronidase activity in developing maize (Zea mays) kernels. *Physiologia Plantarum* **104**, 423-430.
Muller, J., Hart, C. M., Francis, N. J., Vargas, M. L., Sengupta, A., Wild, B., Miller, E. L., O Connor, M. B., Kingston, R. E. and Simon, J. A. (2002). Histone Methyltransferase Activity of a Drosophila Polycomb Group Repressor Complex. *Cell* **111**, 197-208.
Munshi, N., Merika, M., Yie, J., Senger, K., Chen, G. and Thanos, D. (1998). Acetylation of HMG I(Y) by CBP Turns off IFN[beta] Expression by Disrupting the Enhanceosome. *Molecular Cell* **2**, 457-467.
Nawaschin, S. G. (1898). Resultate einer Revision der Befruchtungsvorgaenge bei *Lilium martagon* und *Fritillaria tenella. Bul. Acad. Imp. Sci. St Petersburg* **9**, 377-382.
Neuwald, A. E. and Landsman, D. (1997). GCNS-related histone N-acetyltransferases belong to a diverse superfamily that includes the yeast SPT10 protein. *Trends in Biochemical Sciences* **22**, 154.

Nishiyama, I. and Inomata, N. (1966). Embryological studies on cross-incompatibility between 2x and 4x in *Brassica*. *Japanese Journal of Genetics* **41**, 27-42.
Ogryzko, V. V. (2001). Mammalian histone acetlytransferases and their complexes. *Cellular and Molecular Life Sciences* **58**, 683-692.
Ohad, N., Margossian, L., Hsu, Y. C., Williams, C., Repetti, P. and Fischer, R. L. (1996). A mutation that allows endosperm development without fertilization. *Proc Natl Acad Sci U S A* **93**, 5319-24.
Ohad, N., Yadegari, R., Margossian, L., Hannon, M., Michaeli, D., Harada, J. J., Goldberg, R. B. and Fischer, R. L. (1999). Mutations in FIE, a WD polycomb group gene, allow endosperm development without fertilization. *Plant Cell* **11**, 407-16.
Olsen, O. A. (2001). Endosperm development: Cellularization and cell fate specification. *Annu Rev Plant Physiol Plant Mol Biol* **52**, 233-267.
Olsen, O. A. (2004a). Dynamics of maize aleurone cell formation: the "surface" rule. *Maydica* **49**, 37-40.
Olsen, O. A. (2004b). Nuclear Endosperm Development in Cereals and Arabidopsis thaliana. *Plant Cell* **16**, S214-S227.
Olsen, O. A., Lemmon, B. and Brown, R. (1998). A model for aleurone development. *Trends in Plant Science* **3**, 168-169.
Olsen, O. A., Potter, R. H. and Kalla, R. (1992). Histo-differentiation and molecular biology of developing cereal endosperm. *Seed Science Research* **2**, 117-131.
Opsahl-Ferstad, H. G., Le Deunff, E., Dumas, C. and Rogowsky, P. M. (1997). ZmEsr, a novel endosperm-specific gene expressed in a restricted region around the maize embryo. *Plant Journal* **12**, 235-246.
Osborn, T. C., Chris Pires, J., Birchler, J. A., Auger, D. L., Jeffery Chen, Z., Lee, H. S., Comai, L., Madlung, A., Doerge, R. W. and Colot, V. (2003). Understanding mechanisms of novel gene expression in polyploids. *Trends in Genetics* **19**, 141-147.
Otte, A. P. and Kwaks, T. H. (2003). Gene repression by Polycomb group protein complexes: a distinct complex for every occasion? *Current Opinion in Genetics and Development* **13**, 448-454.
Pandey, R., Mueller, A., Napoli, C. A., Selinger, D., Pikaard, C., Richards, E., Bender, J., Mount, D. W. and Jorgensen, R. A. (2002). Analysis of histone acetyltransferase and histone deacetylase families of Arabidopsis thaliana suggests functional diversification of chromatin modification among multicellular eukaryotes. *Nucleic Acids Research* **30**, 5036-5055.
Papa, C. M., Springer, N. M., Muszynski, M. G., Meeley, R. and Kaeppler, S. M. (2001). Maize Chromomethylase Zea methyltransferase2 Is Required for CpNpG Methylation. *Plant Cell* **13**, 1919-1928.
Pejic, I., Ajmone-Marsan, P., Morgante, M., Kozumplick, V., Castiglioni, P., Taramino, G. and Motto, M. (1998). Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. *TAG Theoretical and Applied Genetics* **97**, 1248-1255.
Pikaard, C. S. (2001). Genomic change and gene silencing in polyploids. *Trends in Genetics* **117**, 675-677.
Punyasingh, K. (1947). Chromosome numbers in crosses of diploid, triploid and tetraploid maize. *Genetics* **32**, 541-554.
Puskas, L. G., Zvara, A., Hackler, L. and Van Hummelen, P. (2002). RNA Amplification Results in Reproducible Microarray Data with Slight Ratio Bias. *Biotechniques* **32**, 1330-1341.
Quackenbush, J. (2002). Microarray data normalization and transformation. *Nature Genetics Supplement* **32**, 496-501.
Randolph, L. F. (1936). Developmental morphology of the caryopsis in maize. *Journal of Agricultural Research* **53**, 881-916.
Reddy, V. M. and Daynard, T. B. (1983). Endosperm characteristics associated with rate of grain filling and kernel size in corn. *Maydica* **28**, 333-355.
Reiser, L. and Fischer, R. L. (1993). The Ovule Sac and the Embryo Sac. *The Plant Cell* **5**, 1291-1301.
Revollo, J. R., Grimm, A. A. and Imai, S. i. (2004). The NAD Biosynthesis Pathway Mediated by Nicotinamide Phosphoribosyltransferase Regulates Sir2 Activity in Mammalian Cells. *Journal of Biological Chemistry* **279**, 50754-50763.
Rinn, J. L., Euskirchen, G., Bertone, P., Martone, R., Luscombe, N. M., Hartman, S., Harrison, P. M., Nelson, F. K., Miller, P. and Gerstein, M. (2003). The transcriptional activity of human Chromosome 22. *Genes and Development* **17**, 529-541.

Robichaud, C. S., Sedee, N. J. A. and Sussex, I. M. (1980). Control and growth of vivaparous embryo mutants by abscisic acid. *Developmental Genetics* **1**, 325-330.
Rolletschek, H., Koch, K., Wobus, U., Borisjuk, L. (2005). Positional cues for the starch/lipid balance in maize kernels and resource partitioning to the embryo. *Plant Journal* **42**, 69-83.
Russell, S. D. (1979). Fine structure of megagametophyte development in *Zea mays*. *Canadian Journal of Botany* **57**, 1093-1110.
Saeed, A. I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T. and Thiagarajan, M. (2003). TM4: A Free, Open-Source System for Microarray Data Management and Analysis. *Biotechniques* **34**, 374-379.
Sarkar, K. R. and Coe, E. H., Jr. (1971). Origin of parthenogenesis in maize and its implications for the production of homozygous lines. *Crop Science* **11**, 543-544.
Saurin, A. J., Shao, Z., Erdjument-Bromage, H., Tempst, P. and Kingston, R. E. (2001). A Drosophila Polycomb group complex includes Zeste and dTAFII proteins. *Nature*, 655-659.
Schel, J. H. N., Kieft, H.and van Lammeren, A. A. M. (1984a). Interactions between embryo and endosperm during early developmental stages of maize caryopses (Zea mays). *Can. J. Bot.* **62**, 2842-2853.
Schena, M., Shalon, D., Davis, R. W. and Brown, P. O. (1995). Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray. *Science* **270**, 467-470.
Schneider, J., BuneSZ, A., Huber, W., Volz, J., Kioschis, P., Hafner, M., Poustka, A. and Sultmann, H. (2004). Systematic analysis of T7 RNA polymerase based in vitro linear RNA amplification for use in microarray experiments. *BMC Genomics* **5**, 29.
Scholten, S., Lorz, H. and Kranz, E. (2002). Patemal mRNA and protein synthesis coincides with male chromatin decondensation in maize zygotes. *Plant Journal* **32**, 221-232.
Scott, R. J., Spielman, M., Bailey, J. and Dickinson, H. G. (1998). Parent-of-origin effects on seed development in Arabidopsis thaliana. *Development* **125**, 3329-41.
Serna, A., Maitz, M., O Connell, T., Santandrea, G., Thevissen, K., Tienens, K., Hueros, G., Faleri, C., Cai, G. and Lottspeich, F. (2001). Maize endosperm secretes a novel antifungal protein into adjacent maternal tissue. *Plant Journal* **25**, 687-698.
Sharman, G. B. (1971). Late DNA replication in the paternally derived X chromosome of female kangaroos. *Nature* **230**, 231-232.
Shchepinov, M. S., Udalova, I. A., Bridgman, A. J. and Southern, E. M. (1997). Oligonucleotide dendrimers: synthesis and use as polylabelled DNA probes. *Nucleic Acids Research* **25**, 4447-4454.
Sheen, J., Hwang, S., Niwa, Y. and Kobayashi, H. (1995). Green-fluorescent protein as a new vital marker in plant cells. *Plant Journal* **8**, 777.
Shen, B., Li, C., Min, Z., Meeley, R. B., Tarczynski, M. C. and Olsen, O. A. (2003). sal1 determines the number of aleurone cell layers in maize endosperm and encodes a class E vacuolar sorting protein. *Proceedings- National Academy of Sciences USA* **100**, 6552-6557.
Sheridan, W. F. and Clark, J. K. (1994). Fertilization and embryogeny in maize. In *The Maize Handbook*, (ed. M. Freeling and V. Walbot). New York: Springer Verlag.
Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A. and Casero, R. A. (2004). Histone Demethylation Mediated by the Nuclear Amine Oxidase Homolog LSD1. *Cell* **119**, 941-953.
Soave, C. and Salamini, F. (1984). The role of structural and regulatory genes in the development of maize endosperm. *Developmental Genetics* **5**, 1-25.
Song, R., Llaca, V., Linton, E. and Messing, J. (2001). Sequence, Regulation, and Evolution of the Maize 22-kD alpha Zein Gene Family. *Genome Research* **11**, 1817-1825.
Spillane, C., MacDougallDagger, C., StockDagger, C., Kohler, C., Vielle-Calzada, J. P., Nunes, S. M., Grossniklaus, U. and Goodrich, J. (2000). Interaction of the Arabidopsis Polycomb group proteins FIE and MEA mediates their common phenotypes. *Current Biology* **10**, 1535-1538.
Springer, N. M., Danilevskaya, O. N., Hermon, P., Helentjaris, T. G., Phillips, R. L., Kaeppler, H. F. and Kaeppler, S. M. (2002). Sequence relationships, conserved domains, and expression patterns for maize homologs of the polycomb group genes E(z), esc, and E(Pc). *Plant Physiol* **128**, 1332-45.
Springer, N. M. and Kaeppler, S. M. (2005). Evolutionary divergence of monocot and dicot Methyl-CpG-binding domain proteins. *Plant Physiology* **138**, 92-104.
St Johnston, D. and Nüsslein-Vollhard, C. (1992). The origins of pattern and polarity in the *Drosophila* embryo. *Cell* **68**, 201-19.

Stears, R. L., Getts, R. C. and Gullans, S. R. (2000). A novel, sensitive detection system for high-density microarrays using dendrimer technology. *Physiological Genomics* **3**, 93-100.
Steiglitz, H. (1977). Role of ß-1,3-glucanase in post-meiotic microspore release. *Developmental Biology* **57**, 87-97.
Surani, A., Barton, S. C. and Norris, M. L. (1984). Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature* **308**, 548-550.
Murphy, S.K., Wylie, A. A., Jirtle, R. L. (2003). Imprinting evolution and the price of silence. *Bioessays* **25**, 577-588.
Swigonova, Z., Lai, J., Ma, J., Ramakrishna, W., Llaca, V., Bennetzen, J. L. and Messing, J. (2004). Close Split of Sorghum and Maize Genome Progenitors. *Genome Research* **14**, 1916-1923.
Tagaki, N. and Sasaki, M. (1975). Preferential inactivation of the paternally derived X chromosome in the extra embryonic membranes of the mouse. *Nature* **256**.
Tariq, M. and Paszkowski, J. (2004). DNA and histone methylation in plants. *Trends in Genetics* **20**, 244-251.
Thompson, R. D., Hueros, G., Becker, H. A. and Maitz, M. (2001). Development and functions of seed transfer cells. *Plant Science* **160**, 775 - 783.
Tian, G. W., You, R. L., Guo, F. L. and Wang, X. C. (1998). Microtubular Cytoskeleton of Free Endosperm Nuclei During Division in Wheat. *Cytologia- International Journal of Cytology* **63**, 427-433.
Tie, F., Prasad-Sinha, J., Birve, A., Rasmuson-Lestander, A. and Harte, P. J. (2003). A 1-Megadalton ESC/E(Z) Complex from Drosophila That Contains Polycomblike and RPD3. *Molecular and Cellular Biology* **23**, 3352-3362.
van Gelder, R., von Zastrow, M., Yool, A., Dement, W., Barchas, J. and Eberwine, J. (1990). Amplified RNA Synthesized from Limited Quantities of Heterogeneous cDNA. *PNAS* **87**, 1663-1667.
Vaughn, M. W. and Martienssen, R. (2005). It's a Small RNA World, After All. *Science* **309**, 1525-1526.
Vielle-Calzada, J. P., Baskar, R. and Grossniklaus, U. (2001). Response to "Early paternal gene activity in Arabidopsis". *Nature* **414**, 710.
Vielle-Calzada, J.-P., Baskar, R. and Grossniklaus, U. (2000). Delayed activation of the paternal genome during seed development. **404**, 91-94.
Vielle-Calzada, J. P., Thomas, J., Spillane, C., Coluccio, A., Hoeppner, M. A. and Grossniklaus, U. (1999). Maintenance of genomic imprinting at the Arabidopsis medea locus requires zygotic DDM1 activity. *Genes Dev* **13**, 2971-82.
Vijayaraghavan, M. R. and Prabhakar, K. (1984). The endosperm. In *Embryology of Angiosperms*, (ed. J. BM), pp. 319-376. Berlin Heidelberg, New York: Springer-Verlag.
Vinkenoog, R. and Scott, R. J. (2001). Autonomous endosperm development in flowering plants: how to overcome the imprinting problem? *Sexual Plant Reproduction* **14**, 189-194.
von Wangenheim, K. H. (1962). Zur Ursache der Abortion von Samenlagen in Diploid-Polyuploid-Kreuzungen. *Z. Vererbungsl.* **93**, 319-34.
von Wangenheim, K. H. and Peterson, H. P. (2004). Aberrant endosperm development in interploidy crosses reveals a timer of differentiation. *Developmental Biology* **270**, 277-289.
Walbot, V. (1994). Overview of key steps in aleurone development. In *The Maize Handbook*, (ed. M. Freeling and V. Walbot), pp. 78-80. New York: Springer.
Walbot, V. (2000). Saturation mutagenesis using maize transposons. *Current Opinion in Plant Biology* **3**, 103-107.
Walbot, V. and Evans, M. M. S. (2003). Unique features of the plant life cycle and their consequences. *Nature Reviews Genetics* **4**, 369-379.
Walling, J. G., Pires, J. C. and Jackson, S. A. (2005). Preparation of samples for comparative studies of plant chromosomes using in situ hybridization methods. In *Molecular Evolution: Producing the Biochemical Data, Part B*, vol. 395, pp. 443-460.
Wang, H., Miyazaki, S., Kawai, K., Deyholos, M., Galbraith, D. W. and Bohnert, H. J. (2003a). Temporal progression of gene expression responses to salt shock in maize roots. *Plant Molecular Biology* **52**, 873-891.
Wang, H.-Y., Malek, R., Kwitek, A., Greene, A., Luu, T., Behbahani, B., Frank, B., Quackenbush, J. and Lee, N. (2003b). Assessing unmodified 70-mer oligonucleotide probe performance on glass-slide microarrays. *Genome Biology* **4**, R5.

Wang, J., Tian, L., Madlung, A., Lee, H. S., Chen, M., Lee, J. J., Watson, B., Kagochi, T., Comai, L. and Chen, Z. J. (2004). Stochastic and epigenetic changes of gene expression in Arabidopsis polyploids. *Genetics* **167**, 1961-1973.
Washida, H., Sugino, A., Messing, J., Esen, A. and Okita, T. W. (2004). Asymmetric Localization of Seed Storage Protein RNAs to Distinct Subdomains of the Endoplasmic Reticulum in Developing Maize Endosperm Cells. *Plant and Cell Physiology* **45**, 1830-1837.
Waterborg, J. H., Harrington, R. E. and Winicov, I. (1990). Dynamic histone acetylation in alfalfa cells : Butyrate interference with acetate labeling. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1049**, 324-330.
Weatherwax, P. (1923). The Story of the Maize Plant. Chicago: The University of Chicago Press.
Weijers, D., Geldner, N., Offringa, R. and Jurgens, G. (2001). Seed development: Early paternal gene activity in Arabidopsis. *Nature*, 709.
Wilkins, J. F. and Haig, D. (2003). What good is genomic imprinting: the function of parent-specific gene expression. *Nature Reviews Genetics* **4**, 359-368.
Williams, J. H. and Friedman, W. E. (2002). Identification of diploid endosperm in an early angiosperm lineage. *Nature*, 522-525.
Wilson, L. M., Whitt, S. R., Ibanez, A. M., Rocheford, T. R., Goodman, M. and Iv, E. S. B. (2004). Dissection of Maize Kernel Composition and Starch Production by Candidate Gene Association. *Plant Cell* **16**, 2719-2733.
Wisniewski, J. P. and Rogowsky, P. M. (2004). Vacuolar H+ -translocating inorganic pyrophosphatase (Vpp1) marks partial aleurone cell fate in cereal endosperm development. *Plant Molecular Biology* **56**, 325-337.
Wolf, H. J. and al, e. (1972). Maize with multilayer aleurone of high protein content. *Crop Science* **12**, 440-442.
Wolpert, L. (1971). Positional information and pattern formation. *Curr. Top. Dev. Biol.* **6**, 183-224.
Woodell, S. R. J. and Valentine, D. H. (1961). Studies in British Primulas. *New Phytologist* **60**, 282-294.
Wu, K., Tian, L., Malik, K., Brown, D. and Miki, B. (2000). Functional analysis of HD2 histone deacetylase homologues in Arabidopsis thaliana. *The Plant Journal* **22**, 19-27.
Yadegari, R., Kinoshita, T., Lotan, O., Cohen, G., Katz, A., Choi, Y., Nakashima, K., Harada, J. J. and Goldberg, R. B. (2000). Mutations in the FIE and MEA Genes That Encode Interacting Polycomb Proteins Cause Parent-of-Origin Effects on Seed Development by Distinct Mechanisms. *Plant Cell* **12**, 2367-2381.
Yang, I., Chen, E., Hasseman, J., Liang, W., Frank, B., Wang, S., Sharov, V., Saeed, A., White, J., Li, J. et al. (2002). Within the fold: assessing differential expression measures and reproducibility in microarray assays. *Genome Biology* **3**, research0062.1 - research0062.12.
Yang, Y. H. and Speed, T. (2002). Design issues for cDNA microarray experiments. *Nature Reviews Genetics* **3**, 579-588.
Young, T. E. and Gallie, D. R. (2000). Regulation of programmed cell death in maize endosperm by abscisic acid. *Plant Mol Biol* **42**, 397-414.
Young, T. E., Gallie, D. R. and DeMason, D. A. (1997). Ethylene-Mediated Programmed Cell Death during Maize Endosperm Development of Wild-Type and shrunken2 Genotypes. *Plant Physiology* **115**, 737-751.
Yu, L. X. and Setter, T. L. (2003). Comparative Transcriptional Profiling of Placenta and Endosperm in Developing Maize Kernels in Response to Water Deficit. *Plant Physiology* **131**, 568-582.
Zemach, A. and Grafi, G. (2003). Characterization of Arabidopsis thaliana methyl-CpG-binding domain (MBD) proteins. *Plant Journal* **34**, 565-572.
Zhao, H., Hastie, T., Whitfield, M. L., Børresen-Dale, A.-L. and Jeffrey, S. S. (2002). Optimization and evaluation of T7 based RNA linear amplification protocols for cDNA microarray analysis. *BMC Genomics* **3**.

8 Appendices

8.1 Contents of enclosed CD

Extensive data was generated in this thesis, which has been included in an accompanying CD. This CD also contains several programs written during the course of this work, a database of biological samples harvested, and details of primers used for different applications in this thesis.

8.2 Publication

Gutierrez-Marcos, J. F., Pennington, P. D., Costa, L. M., Dickinson, H. D. (2003) Imprinting in the endosperm: a possible role in preventing wide hybridization. Phil. Trans. R. Soc. Lond. B **358**, 1105-1111.

8.3 Table of sequence data and primers used to produce custom microarray

THE ROYAL SOCIETY

Published online 2 May 2003

Imprinting in the endosperm: a possible role in preventing wide hybridization

Jose F. Gutierrez-Marcos*, Paul D. Pennington, Liliana M. Costa and Hugh G. Dickinson

Department of Plant Sciences, University of Oxford, Oxford, OX1 3RB, UK

Reproductive isolation is considered to play a key part in evolution, and plants and animals have developed a range of strategies that minimize gene flow between species. In plants, these strategies involve either pre-zygotic barriers, such as differences in floral structure and pollen–stigma recognition, or post-zygotic barriers, which are less well understood and affect aspects of seed development ranging from fertilization to maturation.

In most angiosperms, a double fertilization event gives rise to a zygote and the endosperm: a triploid tissue with an unequal parental genomic contribution, which, like the placenta of mammals, provides reserves to the developing embryo. Interestingly, many aspects of endosperm development, again like the placenta, are regulated by a range of epigenetic mechanisms that are globally termed imprinting. Imprinted genes are characterized by their uniparental expression, the other parental allele being silenced. Normal development of the endosperm thus requires a highly specific balance of gene expression, from either the maternal or paternal genomes. Any alteration of this balance resulting from changes in allelic copy number, sequence or epigenetic imprints can cause endosperm failure and eventual seed abortion. In its widest sense, the endosperm thus serves as an accurate 'sensor' of compatibility between parents. A first step in understanding this important, yet complex system must clearly be the isolation and characterization of as wide a range as possible of imprinted genes.

Keywords: silencing; epigenetic; imprinting; endosperm; hybridization

1. INTRODUCTION

Since the discovery of double fertilization (Nawaschin 1898; Guignard 1899) and the existence of the endosperm just over a century ago, the role of this structure in angiosperm reproduction has remained puzzling. Certainly, development of the endosperm is necessary for the correct development of the embryo and, as early as 1942, Cooper & Brink (1942) proposed a role for endosperm failure in reproductive isolation and angiosperm speciation. The genetic basis of endosperm failure remains unclear, although gene dosage effects (Birchler 1993) and imprinting of regulatory genes (Haig & Westoby 1991) have been proposed as contributory factors.

Imprinting is a mitotically stable epigenetic modification of parental alleles that leads to the differential expression of genes in a parent-of-origin manner. Imprinting therefore renders parental genomes non-equivalent, and for this reason the 'correct' contributions of maternal and paternal genomes are essential for post-fertilization development (Spielman *et al.* 2001; Surani 2001). Interestingly, whereas mammalian embryo development requires the participation of both parental genomes, the same does not hold for all angiosperms, where maternally or paternally derived haploid embryos can produce viable seedlings—as can maternal diploid embryos of apomictic angiosperms (Grimanelli *et al.* 2001). Imprinting thus appears to play a far more significant part in endosperm development (Haig & Westoby 1991) than in embryogenesis. Haig & Westoby (1991) also proposed that the unusual triploid composition of the endosperm possibly led to the selection of parentally imprinted genes that play an important part in the function of the endosperm in reproductive strategies.

2. GENETIC EVIDENCE FOR IMPRINTING IN THE ENDOSPERM

Evidence that imprinting is, in part, responsible for endosperm failure is largely derived from studies of intraspecific crosses between different ploidy levels, and crosses between related species. Reciprocal crosses performed between diploid and tetraploid individuals in a range of species have been reported to result in abnormal endosperm development (Brink & Cooper 1947; Redei 1964; Johnston *et al.* 1980). Likewise, interploidy crosses carried out in maize using *indeterminate gametophyte* (*ig*), a mutant that generates embryo sacs with a variable number of cells and nuclei, indicate that deviations from the normal two maternal to one paternal (2m : 1p) genomic balance in the endosperm resulted in abnormal development, perhaps owing to an imbalance of imprinted sequences within the endosperm (Lin 1984). Although most studies suggest

*Author for correspondence (jose.gutierrez@plants.ox.ac.uk).

One contribution of 21 to a Discussion Meeting Issue 'Mechanisms regulating gene flow in flowering plants'.

© 2003 The Royal Society

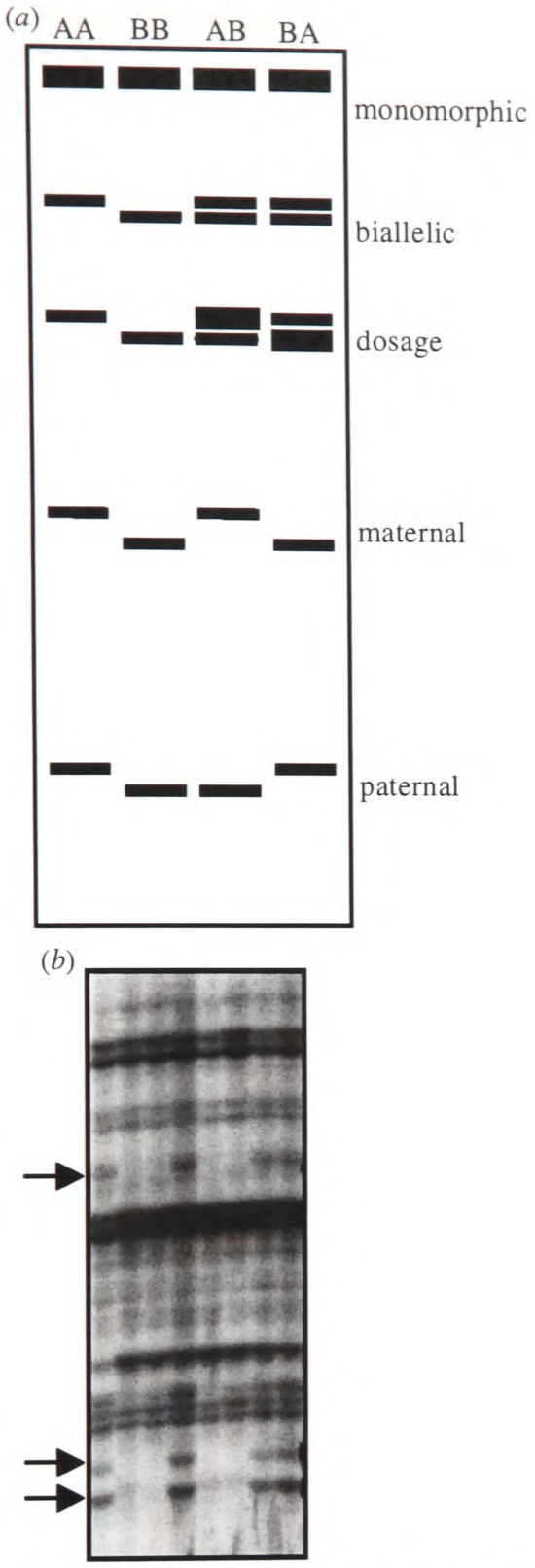


Figure 1. AMD strategy for the identification of imprinted genes in maize endosperm. (*a*) Schematic diagram of band patterns expected after RT-PCR amplification and separation by electrophoresis, showing cDNA polymorphisms between endosperm mRNA samples isolated from two selfed and reciprocally crossed maize inbred lines, A and B. Monomorphic cDNAs are represented by the same sized band located in all four endosperm samples. Among polymorphic cDNAs, biallelic expression is represented by the presence of both parental bands in reciprocal endosperm samples (AB and BA); dosage expression gives a similar pattern to biallelic expression only with a stronger band signal for the maternal allele in reciprocal endosperm samples; monoallelic expression is represented by the absence of either a maternal or paternal band in reciprocal endosperm samples and are good candidates for genes subjected to a paternal and maternal parent-of-origin expression pattern, respectively. (*b*) Autoradiograph showing a typical AMD gel. Each lane displays different cDNAs from selfed and reciprocally crossed maize inbred lines (as above), amplified by RT-PCR using a range of random primer combinations. Note the presence of a paternal parent-of-origin pattern of expression (arrows).

that the 2m : 1p ratio is essential for endosperm development, there are some exceptions. Lin (1984) showed that 3m : 1p endosperms developed, whereas 6m : 2p endosperms did not. In another approach, Rhoades & Dempsey (1966) generated tetraploid plants from *elongate* mutant lines, which produce unreduced gametes. When crossed with diploid, tetraploid and hexaploid males, endosperms were formed with a diverse number of ploidies ranging from $3x$ to $11x$. The only endosperms that developed normally were those of diploid and tetraploid males, although a few kernels were obtained from crosses with a hexaploid male. Interestingly, self-pollination of hexaploid and octoploid maize results in high female sterility (Randolph 1932). One possible explanation is that high ploidy in the endosperm might have caused other developmental abnormalities, separate from those resulting from deviation from the 2m : 1p ratio that affect lower endosperm ploidies (Birchler 1993). More convincing evidence of imprinting operating in the endosperm came from chromosomal translocation studies in maize. Lin (1982) found that the absence of a paternally inherited chromosome arm in the endosperm resulted in reduced seed size, and that extra female copies of this chromosome arm failed to compensate for the absence of the paternally transmitted copy. Conversely, evidence of a dosage mechanism operating in the endosperm was provided by Birchler & Hart (1987), who demonstrated that extra maternal copies of chromosome arms *other* than the arm absent from the paternal genome also accentuated a reduction in kernel size.

Taken together, the evidence found to date strongly suggests that dosage effect and imprinting of different loci are the main contenders for regulating endosperm development in maize, and most probably in other angiosperms.

3. AN OVERVIEW OF IMPRINTED GENES IN PLANTS

The first plant gene that was demonstrated to be imprinted was the endosperm-expressed *r1* gene of maize (Kermicle 1970, 1978). In a series of outstanding studies, Kermicle not only identified imprinting in this gene, but also genetically demonstrated the existence of a regulatory element, the maternal derepressor of *r* (MDR) (Kermicle 1978).

Almost 20 years later, additional examples of imprinting in plants were identified including genes encoding zeins, tubulins and a posttranscriptional regulator of zeins (*dzr1*) (Chaudhuri & Messing 1994; Lund *et al.* 1995*a*,*b*). A common feature of these imprinted genes is that their expression is endosperm-specific, they are members of multi-copy gene families and the imprinting seems to be allele-specific only in some inbred lines.

Recently, locus-specific imprinting has been identified in genes of the FIS class in *Arabidopsis*. MEDEA was identified as a lethal gametophytic maternal effect mutant (*mea*) and encodes an enhancer of zeste (E[Z]) polycomb protein (Grossniklaus *et al.* 1998; Kinoshita *et al.* 1999). Later studies suggest that MEDEA expression is imprinted in both embryo and endosperm (Vielle-Calzada *et al.* 1999). The other two genes of the FIS class (*FIE*, *FIS2*) characterized in *Arabidopsis* are exclusively maternally

Table 1. (*a*) Sequences revealed by AMD screening of 6 and 15 DAP endosperms, and (*b*) putative functions of some characterized sequences.

(*a*) samples	paternal	maternal	dosage	percentage
6 DAP	1	15	65	2–5
15 DAP	0	31	125	3–6

(*b*) putative function	paternal	maternal	dosage	total
structural	0	6	4	10
metabolic	1	2	10	13
transcription	0	2	1	3
chromatin	0	1	0	1

expressed in the endosperm, and belong to the polycomb group of proteins (Luo *et al.* 1999; Yadegari *et al.* 2000). *FIE* seemingly interacts with MEDEA, and among other roles, may be involved in the establishment of the anterior–posterior axis in early endosperm development (Spillane *et al.* 2000; Sorensen *et al.* 2001). Surprisingly, different mechanisms may regulate the parent-of-origin expression seen in *mea* and *fie* mutants, since plants with reduced global methylation used as pollen donors can rescue *fis2* and *mea* but not *fie* (Luo *et al.* 2000; Vinkenoog *et al.* 2000). Similarly, pollen from a mutant that affects chromatin conformation and DNA methylation (*ddm1*) can rescue *mea* but not *fie* (Yadegari *et al.* 2000). Most recently, *DEMETER* (*DME*), a DNA glycosylase, has been demonstrated to be required for the specific activation of *MEA* in the megagametophyte (Choi *et al.* 2002).

A delay of paternal genome activation in the *Arabidopsis* embryo and endosperm has been reported (Vielle-Calzada *et al.* 2000), and it has been proposed that this genome-wide silencing of paternally inherited genes may affect the vast majority of loci—excluding those with a major role in early seed development (Baroux *et al.* 2002). However, *in vitro* data from maize suggest that transcription of paternally transmitted transgenes occurs in zygotes shortly after fertilization, pointing to activity of the male genome, at least in maize, almost immediately after karyogamy (Scholten *et al.* 2002). Whether the entire paternal genome, or just a fraction of it, is available for early zygotic and endosperm development thus remains unresolved.

These findings suggest that imprinting in plants is a complex process common, but perhaps not exclusive, to the endosperm. Further, it appears that imprinting operates via multiple mechanisms: allele-specific (most sequences described in maize to date), gene-specific (*MEDEA*, FIS and *FIE* in *Arabidopsis*), and genome-wide (the reported paternal genome silencing in *Arabidopsis*). The isolation and characterization of further imprinted genes in an extended range of species should assist our understanding of the operation, and the evolutionary origin of this intriguing and important process.

4. GENOME-WIDE SCREENING FOR IMPRINTED GENES IN PLANTS

Mutant screens have been of great value in isolating imprinted genes in *Arabidopsis* (Baroux *et al.* 2002) but have proven less useful in other species with a more complex genomic organization. To identify novel imprinted sequences in the maize endosperm we have modified a molecular screen termed AMD, originally developed to isolate imprinted sequences in mouse (Hagiwara *et al.* 1997). In principle, AMD is based on a differential display PCR technique that allows visualization of multiple 'transcripts'—in reality 3′-ends of arbitrary cDNAs derived from tissue- or stage-specific mRNAs—regions rich in polymorphisms among maize inbred lines (Bhattramakki *et al.* 2002). AMD can therefore reveal polymorphic alleles and their parental expression levels in endosperm samples isolated from self-pollinated and reciprocally crossed parental inbred lines (figure 1). Using a combination of arbitrary and anchor primers, it is possible to generate a unique 'fingerprint' of polymorphic bands representing four expression categories: biallelic, dosage-dependent, exclusively paternal and exclusively maternal (figure 1*a*). In a comparative analysis of two developmental stages (6 and 15 DAP), between 2% and 6% of the amplified polymorphic sequences exhibited parent-of-origin expression. A large proportion of these sequences belonged to structural and biochemical gene families, whereas a smaller number were transcriptional regulators and genes involved in chromatin architecture (table 1). Significantly, most of these sequences were expressed in a dosage-dependent manner with only a small minority of sequences showing exclusively maternal or paternal expression patterns (table 1). Further molecular characterization of the one clearly paternally imprinted sequence detected, *PEG1*, showed that its parent-of-origin expression pattern is restricted to the endosperm—the gene being biallelically expressed in the embryo (figure 2*a*). This is the first report, to our knowledge, of a gene showing a paternal parent-of-origin expression pattern in plants. We also identified a small gene family, homologous to the *Arabidopsis FIE* gene (Springer *et al.* 2002), which showed an exclusively maternal pattern of expression. Of this family, *FIE101* is only expressed in the endosperm and imprinted throughout endosperm development (figure 2*a*), whereas *FIE102* is biallelically expressed in the embryo, but maternally expressed only at early stages of endosperm development (figure 2*a*). The differences found in expression patterns and parent-of-origin effects strongly indicate that the maize *FIE*-like sequences may have evolved to assume divergent roles during maize evolution (Danilevskaya *et al.* 2003).

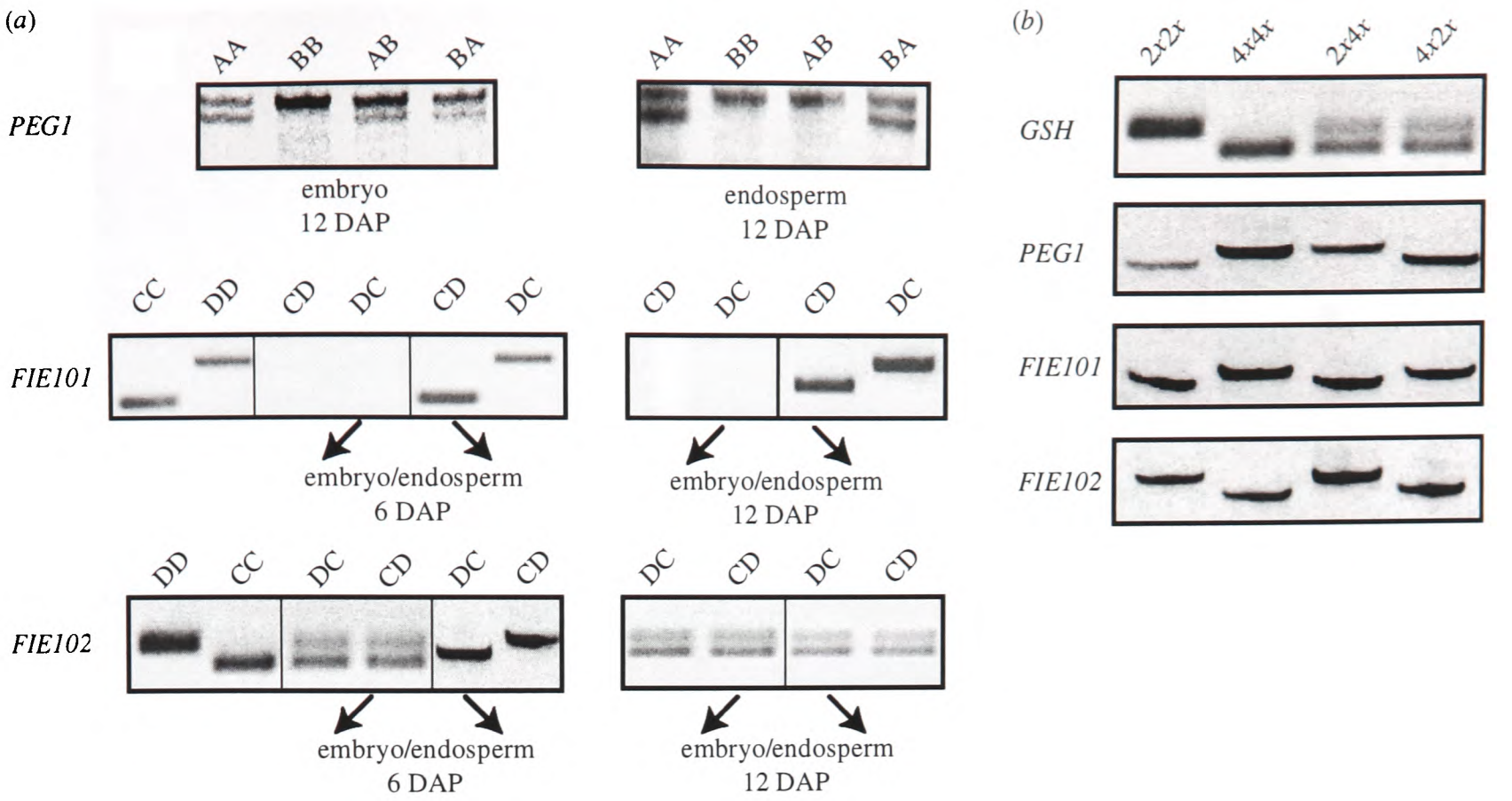


Figure 2. Several types of parent-of-origin expression patterns operate in maize endosperm. (*a*) Embryo and endosperm (polymorphic) mRNA accumulation of parental inherited alleles after selfing and reciprocally crossing four maize inbred lines (where A is W22; B, Tx303; C, Mo17; and D, B73). *PEG1* shows a paternal parent-of-origin pattern of expression in endosperm and biallelic expression in embryo. The endosperm-specific gene, *FIE101*, shows a maternal pattern of expression throughout endosperm development, whereas *FIE102* shows a maternal parent-of-origin pattern of expression only at early stages of endosperm development and is biallelically expressed in embryo. (*b*) Early endosperm (6 DAP) mRNA expression of parental alleles after selfed and reciprocally crossed diploid and tetraploid maize inbred lines. Glutathione synthase expression was used as a control for biallelic expression.

Our study clearly demonstrates how a genomic screening approach can be effective in identifying maize endosperm sequences expressed in a parent-of-origin pattern. While the sequences we have characterized to date are unambiguously imprinted in the maize endosperm, much further study is required before a clear understanding can be gained of the mechanisms by which imprinting operates during early embryo and endosperm development. Significantly, the parent-of-origin expression patterns of *PEG1*, *FIE101* and *FIE102* in the endosperm are not affected by interploidy crosses (figure 2*b*), which suggests that the imprinting mechanism continues to function in the endosperm despite alterations to the maternal to paternal genomic ratios.

5. FAILURE OF EARLY ENDOSPERM DEVELOPMENT IN INTERPLOIDY CROSSES AND HYBRIDS

Cooper & Brink (1942) were among the first to ascribe a major role to the endosperm in the failure of seed production after interspecific hybridization. Histological studies on hybrid maize showed abnormalities at very early stages of endosperm development ranging from defects at the chalazal pole (Cooper & Brink 1940), to abnormal free nuclear division rates (Brink & Cooper 1940). Interploidy crosses in monocot species also exhibit altered rates of endosperm cellularization (Kihara & Nishiyama 1932; Wakakuwa 1934; Cooper 1951; Håkanson 1953), while similar crosses in dicots reveal cellularization to occur early with the production of fewer cells in maternal excess. Dicot endosperms with excess paternal genomes exhibit delayed cellularization and elevated rates of cytokinesis (Valentine 1955; Nishiyama & Inomata 1966; Scott *et al.* 1998).

Evidence is thus accumulating that failure of the polygonum-type endosperm in interspecific and interploidy hybrids results from abnormal development at, or about, the free nuclear stage, since hybridizations involving species that do not undergo a free nuclear stage of development do not respond in this way (Sansome *et al.* 1942; Cooper & Brink 1945).

While many of the regulators of embryo and endosperm development have now been identified (comprehensively reviewed in Jurgens (2001) and Olsen (2001)), the molecular basis of these abnormalities in young endosperm development during the free nuclear stage is not well understood. Previous studies have however, revealed that the BETL cells develop abnormally in endosperms resulting from diploid by tetraploid crosses (Charlton *et al.* 1995). To explore these early developmental changes in a structure that is largely inaccessible to the normal techniques of molecular analysis, we have employed a maize transgenic line carrying an endosperm-specific reporter construct that permits the identification of BETL cells in the developing endosperm (figure 3*a–c*). Preliminary analysis has shown dramatic changes in the distribution pattern of this marker following interploidy crosses, ranging from limited expression in the BETL, to a few aberrant cells localized in the apical region of the endosperm (figure 3*b*,*c*). Further, *in situ* hybridization analysis of an early transcriptional activator of BETL-specific

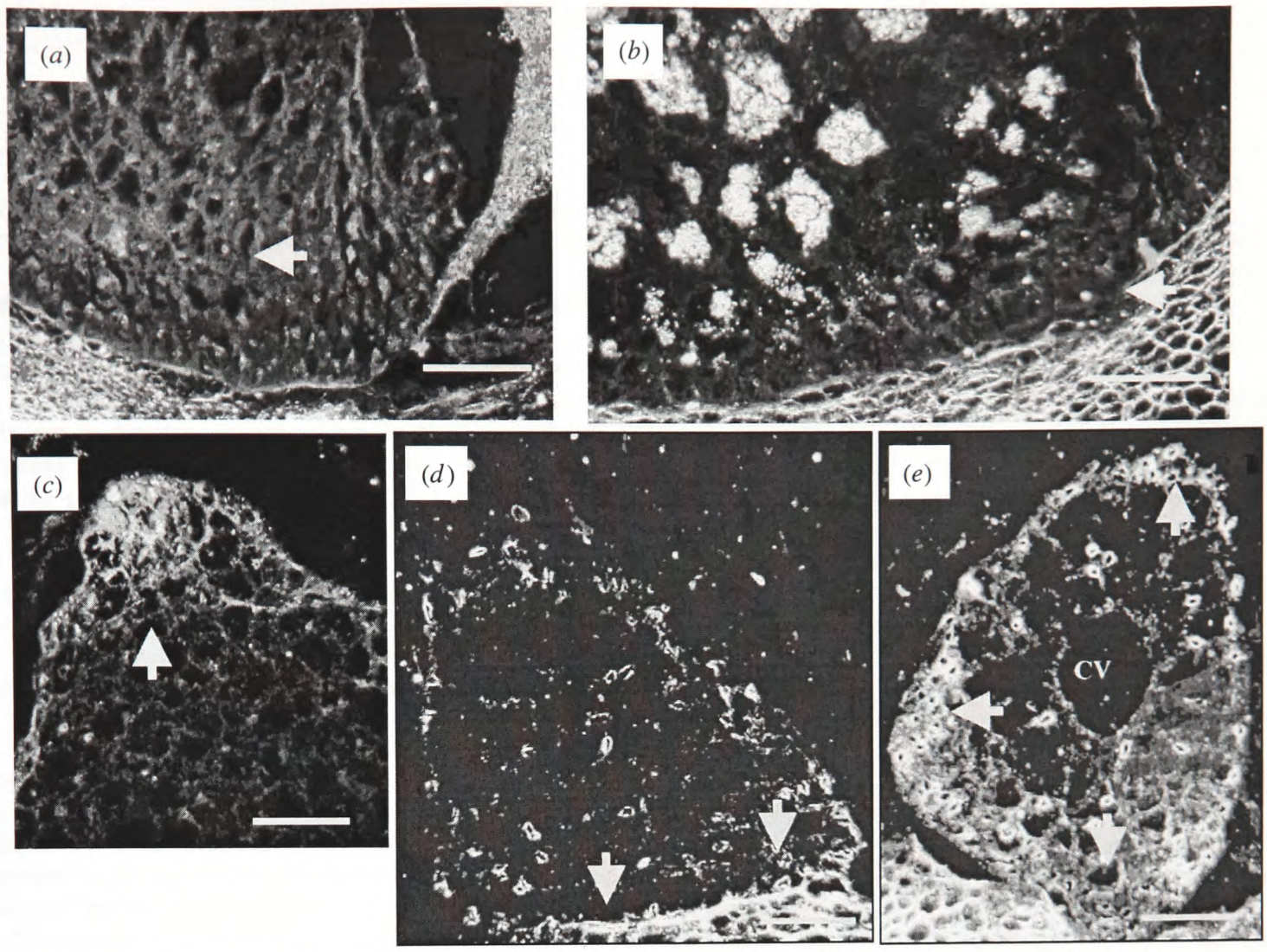


Figure 3. Abnormal localization of BETL transcripts in interploidy maize kernels. β-glucoronidase localization at 10 DAP in balanced (2m : 1p) triploid endosperms is confined to basal transfer cells (*a*), restricted to a small group of transfer cells in the maternal excess (4m : 1p) interploidy endosperms (*b*), and limited to a few cells in the apex in a paternal excess (2m : 2p) interploidy endosperm (*c*). Distribution of ZmMRP transcript during early endosperm development was found in the basal portion of maternal excess interploidy endosperms (*d*), and in few basal and apical cells of the early, cellularizing, paternal excess interploidy endosperm (*e*). Scale bars, 100 μm.

genes, *MRP* (Gomez *et al.* 2002), in interploidy endosperms undergoing cellularization showed abnormal patterns of expression (figure 3*d*,*e*).

Interestingly, specification of BETL in cereals is held to occur in the free nuclear or coenocytic endosperm (Olsen 2001), and mRNAs for MRP have been localized in the basal portion of the coenocyte a few hours after fertilization (Gomez *et al.* 2002). Taken together with our data, these observations suggest that following interploidy crosses, developmental regulators become aberrantly localized within the coenocytic endosperm, leading to an inappropriate distribution of cells with BETL identity in the cellular endosperm.

The parental balance of genomes in these interploidy crosses is also clearly of importance since expression of BETL marker genes is very disorganized in endosperms with a 2m : 2p genomic balance when compared with endosperms with a 4m : 1p genomic balance (figure 3*b*,*c*). While there are features of these endosperms that almost certainly result from parental conflict operating through imprinting (Haig & Westoby 1991; Scott *et al.* 1998), they also highlight the pivotal role played by maternally transmitted factors in establishing functional domains during early endosperm development. How this maternal control is exerted at a molecular level remains unclear, but data currently available suggest that it is achieved through a complex series of interactions between maternally inherited factors and imprinted genes activated after fertilization. It is hoped that molecular analysis of the striking maternal effect mutants found in maize (Gavazzi *et al.* 1997; Evans & Kermicle 2001) and *Arabidopsis* (reviewed in Chaudhury & Berger 2001; Berger 2003) will help in understanding this process.

6. CONCLUSIONS

Little is currently known of the molecular and genetic mechanisms responsible for the endosperm acting as a hybridization barrier in plants. Gene dosage and imprinting effects in the endosperm are currently considered the 'gatekeepers' of endosperm development, yet conclusive evidence linking these processes with hybrid failure remains patchy. Analysis of the molecular mechanisms regulating endosperm development in hybrids will not only reveal the parts played by maternal determinants and parental imprinting in this complex process, but may also shed some light onto the evolutionary events that have led to the development, solely in the angiosperms, of this unusual triploid structure.

This work was funded by the University of Oxford, the UK BBSRC, and EU Frameworks IV and V. The authors thank the Department of Plant Sciences at the University of Oxford for greenhouse facilities and technical support, and gratefully acknowledge the invaluable technical and scientific contribution made by Biogemma and Syngenta (JHI) to this project.

They thank Dr Richard Thompson for providing them with the pBET1-GUS transgenic line.

REFERENCES

Baroux, C., Spillane, C. & Grossniklaus, U. 2002 Genomic imprinting during seed development. *Adv. Genet.* **46**, 165–214.

Berger, F. 2003 Endosperm: the crossroad of seed development. *Curr. Opin. Plant. Biol.* **6**, 42–50.

Bhattramakki, D., Dolan, M., Hanafey, M., Wineland, R., Vaske, D., Register 3rd, J. C., Tingey, S. V. & Rafalski, A. 2002 Insertion-deletion polymorphisms in 3′ regions of maize genes occur frequently and can be used as highly informative genetic markers. *Plant Mol. Biol.* **48**, 539–547.

Birchler, J. & Hart, J. R. 1987 Interaction of endosperm size factors in maize. *Genetics* **117**, 309–317.

Birchler, J. A. 1993 Dosage analysis of maize endosperm development. *A. Rev. Genet.* **27**, 181–204.

Brink, R. A. & Cooper, D. C. 1940 Double fertilization and development of the seed in angiosperms. *Bot. Gaz.* **102**, 1–25.

Brink, R. A. & Cooper, D. C. 1947 The endosperm in seed development. *Bot. Rev.* **13**, 423–541.

Charlton, W. L., Keen, C. L., Merriman, C., Lynch, P., Greenland, A. J. & Dickinson, H. G. 1995 Endosperm development in *Zea mays*; implications of gametic imprinting and paternal excess in regulation of transfer layer development. *Development* **121**, 3089–3097.

Chaudhuri, S. & Messing, J. 1994 Allele-specific parental imprinting of *dzr1*, a posttranscriptional regulator of zein accumulation. *Proc. Natl Acad. Sci. USA* **91**, 4867–4871.

Choi, Y., Gehring, M., Johnson, L., Hannon, M., Harada, J. J., Goldberg, R. B., Jacobsen, S. E. & Fischer, R. L. 2002 DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in arabidopsis. *Cell* **110**, 33–42.

Cooper, D. C. 1951 Caryopsis development following matings between diploid and tetraploid strains of *Zea mays*. *Am. J. Bot.* **38**, 702–708.

Cooper, D. C. & Brink, R. A. 1940 Somatoplastic sterility as a cause of seed failure after interspecific hybridization. *Genetics* **25**, 593–617.

Cooper, D. C. & Brink, R. A. 1942 The endosperm as a barrier to interspecific hybridization in flowering plants. *Science* **95**, 75–76.

Cooper, D. C. & Brink, R. A. 1945 Seed collapse following matings between diploid and tetraploid races of *Lycopersicum pimpinellifolium*. *Genetics* **30**, 276–401.

Danilevskaya, O. N., Hermon, P. S. H., Muszynski, M. G., Kollipara, K. & Ananiev, E. V. 2003 Duplicated *fie* genes in maize: expression pattern and imprinting suggest distinct functions. *Plant Cell* **15**, 425–438.

Evans, M. M. & Kermicle, J. L. 2001 Interactions between maternal effect and zygotic effect mutations during maize seed development. *Genetics* **159**, 303–315.

Gavazzi, G., Dolfini, S., Allegra, D., Castiglioni, P., Todesco, G. & Hoxha, M. 1997 Dap (Defective aleurone pigmentation) mutations affect maize aleurone development. *Mol. Gen. Genet.* **256**, 223–230.

Gomez, E., Royo, J., Guo, Y., Thompson, R. & Hueros, G. 2002 Establishment of cereal endosperm expression domains: identification and properties of a maize transfer cell-specific transcription factor, ZmMRP-1. *Plant Cell* **14**, 599–610.

Grimanelli, D., Leblanc, O., Perotti, E. & Grossniklaus, U. 2001 Developmental genetics of gametophytic apomixis. *Trends Genet.* **17**, 597–604.

Grossniklaus, U., Vielle-Calzada, J. P., Hoeppner, M. A. & Gagliano, W. B. 1998 Maternal control of embryogenesis by MEDEA a *Polycomb* group gene in *Arabidopsis*. *Nature* **280**, 446–450.

Guignard, L. 1899 Sur les antherozoides et la double copulation sexualle chez les vegetaux angiospermes. *Crit. Rev. Acad. Sci. Paris* **128**, 864–871.

Hagiwara, Y., Hirai, M., Nishiyama, K., Kanazawa, I., Ueda, T., Sakaki, Y. & Ito, T. 1997 Screening for imprinted genes by allelic message display: identification of a paternally expressed gene impact on mouse chromosome 18. *Proc. Natl Acad. Sci. USA* **94**, 9249–9254.

Haig, D. & Westoby, M. 1991 Genomic imprinting in the endosperm: its effect on seed development in crosses between species, and between different ploidies of the same species, and its implications for the evolution of apomixis. *Phil. Trans. R. Soc. Lond.* B **333**, 1–13.

Håkanson, A. 1953 Endosperm formation after 2*x*, 4*x* crosses in certain cereals, especially in *Hordeum vulgare*. *Hereditas* **39**, 57–64.

Johnston, S. A., de Nijs, T. P. M., Peloquin, S. J. & Hanneman, R. E. J. 1980 The significance of genic balance to endosperm development in interspecific crosses. *Theor. Appl. Genet.* **57**, 5–9.

Jurgens, G. 2001 Apical-basal pattern formation in *Arabidopsis* embryogenesis. *EMBO J.* **20**, 3609–3616.

Kermicle, J. L. 1970 Dependence of the R-mottled aleurone phenotype in maize on the mode of sexual transmission. *Genetics* **66**, 69–85.

Kermicle, J. L. 1978 Imprinting of gene of action in maize endosperm. In *Maize breeding and genetics* (ed. D. B. Walden), pp. 357–371. New York: Wiley.

Kihara, H. & Nishiyama, I. 1932 The genetics and cytology of cereals. III. Different compatibility in reciprocal crosses of *Avena*, with special reference to tetraploid hybrids between hexaploid and diploid species. *Jpn. J. Bot.* **6**, 245–305.

Kinoshita, T., Yadegari, R., Harada, J. J., Goldberg, R. B. & Fischer, R. L. 1999 Imprinting of the MEDEA polycomb gene in the *Arabidopsis* endosperm. *Plant Cell* **11**, 1945–1952.

Lin, B.-Y. 1982 Association of endosperm reduction with parental imprinting in maize. *Genetics* **100**, 475–486.

Lin, B.-Y. 1984 Ploidy barrier to endosperm development in maize. *Genetics* **107**, 103–115.

Lund, G., Ciceri, P. & Viotti, A. 1995*a* Maternal-specific demethylation and expression of specific alleles of zein genes in the endosperm of *Zea mays* L. *Plant J.* **8**, 571–581.

Lund, G., Messing, J. & Viotti, A. 1995*b* Endosperm-specific demethylation and activation of specific alleles of alpha-tubulin genes of *Zea mays* L. *Mol. Gen. Genet.* **246**, 716–722.

Luo, M., Bilodeau, P., Koltunow, A., Dennis, E. S., Peacock, W. J. & Chaudhury, A. M. 1999 Genes controlling fertilization-independent seed development in *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA* **96**, 296–301.

Luo, M., Bilodeau, P., Dennis, E. S., Peacock, W. J. & Chaudhury, A. 2000 Expression and parent-of-origin effects for FIS2, MEA, and FIE in the endosperm and embryo of developing *Arabidopsis* seeds. *Proc. Natl Acad. Sci. USA* **97**, 10 637–10 642.

Nawaschin, S. G. 1898 Resultate einer Revision der Befruchtungsvorgaenge bei *Lilium martagon* und *Fritillaria tenella*. *Bul. Acad. Imp. Sci. St. Petersburg* **9**, 377–382.

Nishiyama, I. & Inomata, N. 1966 Embryological studies on cross-incompatibility between 2*x* and 4*x* in *Brassica*. *Jpn. J. Bot.* **41**, 27–42.

Olsen, O. A. 2001 Endosperm development: cellularization and cell fate specification. *A. Rev. Plant Physiol. Plant Mol. Biol.* **52**, 233–267.

Randolph, L. F. 1932 Some effects of high temperature on polyploidy and other variations in maize. *Proc. Natl Acad. Sci. USA* **18**, 222–229.

Redei, G. P. 1964 Crossing experiences with polyploids. *Arabidopsis Inf. Serv.* **1**, 13.

Rhoades, M. M. & Dempsey, E. 1966 Induction of chromosome doubling at meiosis by the *elongate* gene in maize. *Genetics* **54**, 505–522.

Sansome, E. R., Stina, S. & Blakeslee, A. F. 1942 Disintegration of ovules in diploid–tetraploid crosses in *Datura*. *Bull. Torrey Bot. Club* **69**, 405–420.

Scholten, S., Lorz, H. & Kranz, E. 2002 Paternal mRNA and protein synthesis coincides with male chromatin decondensation in maize zygotes. *Plant J.* **32**, 221–231.

Scott, R. J., Spielman, M., Bailey, J. & Dickinson, H. G. 1998 Parent-of-origin effects on seed development in *Arabidopsis thaliana*. *Development* **125**, 3329–3341.

Sorensen, M. B., Chaudhury, A. M., Robert, H., Bancharel, E. & Berger, F. 2001 Polycomb group genes control pattern formation in plant seed. *Curr. Biol.* **11**, 277–281.

Spielman, M., Vinkenoog, R., Dickinson, H. G. & Scott, R. J. 2001 The epigenetic basis of gender in flowering plants and mammals. *Trends Genet.* **17**, 705–711.

Spillane, C., MacDougall, C., Stock, C., Koehler, C., Vielle-Calzada, J. P., Nunes, S. M., Grossniklaus, U. & Goodrich, J. 2000 Interaction of the *Arabidopsis* polycomb group proteins FIE and MEA mediates their common phenotypes. *Curr. Biol.* **10**, 1535–1538.

Springer, N. M., Danilevskaya, O. N., Hermon, P., Helentjaris, T. G., Phillips, R. L., Kaeppler, H. F. & Kaeppler, S. M. 2002 Sequence relationships, conserved domains, and expression patterns for maize homologs of the polycomb group genes E(z), esc and E(Pc). *Plant Physiol.* **128**, 1332–1345.

Surani, M. A. 2001 Reprogramming of genome function through epigenetic inheritance. *Nature* **414**, 122–128.

Valentine, D. H. 1955 Studies in British primulas. IV. Hybridization between *Primula vulgaris* Huds. and *P. veris* L. *New Phytol.* **54**, 70–80.

Vielle-Calzada, J. P., Thomas, J., Spillane, C., Coluccio, A., Hoeppner, M. A. & Grossniklaus, U. 1999 Maintenance of genomic imprinting at the *Arabidopsis* MEDEA locus requires zygotic DDM1 activity. *Genes Dev.* **13**, 2971–2982.

Vielle-Calzada, J. P., Baskar, R. & Grossniklaus, U. 2000 Delayed activation of the paternal genome during seed development. *Nature* **404**, 91–94.

Vinkenoog, R., Spielman, M., Adams, S., Fischer, R. L., Dickinson, H. G. & Scott, R. J. 2000 Hypomethylation promotes autonomous endosperm development and rescues postfertilization lethality in fie mutants. *Plant Cell* **12**, 2271–2282.

Wakakuwa, S. 1934 Embryological studies on the different seed-development in reciprocal interspecifc crosses of wheat. *Jpn. J. Bot.* **7**, 151–184.

Yadegari, R. (and 10 others) 2000 Mutations in the FIE and MEA genes that encode interacting polycomb proteins cause parent-of-origin effects on seed development by distinct mechanisms. *Plant Cell* **12**, 2367–2382.

GLOSSARY

AMD: allelic message display
BETL: basal endosperm transfer layer
DAP: days after pollination
FIS: fertilization-independent seed

ID	Forward Primer (5'-3')	Reverse Primer (5'-3')	GeneClass	Sequence /GenBank ID
BETL	ATTGGCTGTTGCCATTCTG	CTTTGTGCTCCACGCTTTC	Endosperm Marker	ATTGGCTGTTGCCATTCTGTCCTCACTGTCTCCATGCTACGAAGCCGGCGGCTGCATCGGG AAACCCAAGAAGTCACCGCCACCGCCGAGAAGGCCATACTTCTCGTCCTATTCCGAGGACC ACCAGAATTGCCGCTTGATCTGTAGCTCCAAGGGTTTCAAGGACGGTGGCTGGTGTGACGA AAGCGTGGAGCACAAAG
BRD111	CTTGAACGGAGCCTTGCTCC	CTGGCTTGCTAATGCTTCGG	Chromatin	BG410594
BRD113	TGTTGAACGAGTTGCCCTCC	CAGACTTCTGTTTCTTTGACTTG	Chromatin	TGTTGAACGAGTTGCCCTCCAACCTTCCAGTTTTATTGCTTGGAACATCATCAGTGGTTTTTA CTGACCTTGAAGAAGAGTGTGCTTCCATATTCTCTTCTTGCAATGTGTATCAAGTGGATCAAC CAAGTTTTGACGATAGATTGAGGTACTTCAGTATATTGTTTGAGTCATTGCTCTCTTTCCAAA TGGAAGAATCAAGAAACAAGTCAAAGAAACAGAAGTCTG
CDC2_Cell division cycle2	ATTTAGCAGGGCACTGGAC	ATTTAGCAGGTCACTGTACCA	Control (positive)	ATTTAGCAGGGCACTGGACCACTTCAAGGGCCTTGAAGTACTCATGCTCAAGGGCTTGCCT CGCTGGGATTCTTTTGCTTGGCTCGGATCGAAACATTTTAAAAAGAAGGGCCCACCCAGCA GGGGCAAAATTTGGGACTACTGGTGCCAGGTCCTGAGCTTGCCACCTGGGGAAAGCTGTCT TGAAGTCAGGCAAACAACTGACTCCTGGCCAACTTTGGTCATTCGGGGGACCTAGTATCCT GAATATCTTAAAAAGTTCGTCGATCTCAGAATCGCCAGGGAATAGTGGCTTTTGGGTCACCA TTTCCGCAAAGATACAGCCCGCAGACCACACATCAACTGGTGTGGAATACTGCCGCGCTCC AAGCAGAATTTCTGGAGCTCTGTACCATAATGTCACTACCTCATGAG
CENTA	CACCTAAAAACAGTTTGAGCC	GCACGAACAACAACAACAG	Retrotransposon	CACCTAAAAACAGTTTGAGCCTTTGCTTTACTACTTAGTTTTCGACTTGTTGAATTCCGGTAG CTGCATTTGGGTCGAGTTGCTGGTCTAAAGTTTTCTTACCGCTAGAGTTTCGAGTTCGCGCC ACCTTGTTTCAATCACCAGTTTAGACCTCTTGCTGCAATTCAACCAAAAAGAAGAGAAAGCAA AAGGCGAGTGCACAAAAAAAGCCGCACTAATCAGCAAAACAAAAAAAGACACGTGCAAAAC AAAAGAGAGAGAAAAAAACCAGTTCTGAATTTTGGTAGATAAAATTTGTAAGTGCAACAAAAC AAAAGGCAGTTTGTGTGCCTTCTTTTTATAGTTTCAGAAATCAGATTGTTGTTCTGAGCTTTT GGTGATACTATTTGTGTAACGGCTCGCGTCTCTATTACGGTTTGGACTAGGACCAGCACAAC ACCTTGTGGAACGTTTATTCAACTTGTTGTGGCTAACGTGGTACTAGCTATTCCTTGGAACTA TTGTTTAAACAGCCACCTATAAATCCACAAAATTTTCTACAACACCACCAGGTTGTGCTAGCA GCCACTGTTGTTGTTGTTCGTG
CHB101	M13_For	M13_Rev	Chromatin	BG873812
CHC101	GAGGCGTGGATTTCTCCTAGC	CTACTATCATAATCCTGCATTCC	Chromatin	GAGGCGTGGATTTCTCCTAGCTATTTTAGTGTATAAACAAATTATCTGTTCTGTATTGGCAGT AAATTTCAGGCGTATTAGTGAAGATAATTGTTCAGTGTTGTGAAATGATGTTCAGGCTAGTCA TTGTATGTTTATGTTGTGTTCTATATTGTCAAGTGCTGAACTGAACCGTAATGTTGCTTGGTG TGGCCCTGAAAATAGTAAGCCTAAATCACATTCTAAAAAACTGGAATGCAGGATTATGATAGT AG
CHR101	CAGGAGGACTCTAGGCCATG	GCCGAACAAGAGAGCGTCAG	Chromatin	CAGGAGGACTCTAGGCCATGTGTGGTCGAGTCTCCGTACGGTGGTCGTTGCGTTAGATTAC GTGGGTGGGTGATGGGACGGTTGGTTGAACCAGACCCGACGGCAGCATGGACACACGCCG GTTGCTTTGTAAAGTTAGTATCGCGTCGTCCGTGATACAATGCGGTCCTCCGTCTACATGGG TGCTGACGCTCTCTTGTTCGGC
CHR102	GGCTGGTTGGTGGTGGTAG	AATTGAGGAAGTGAGGACCC	Chromatin	GGCTGGTTGGTGGTGGTAGGAGTGATGCTTGTGATATAGCTCTGGCGGTAGAAGTTGTAGC TAGGGCGATGGAGTTTCAGTGTAGGCATAGGGTAGGGCTGGAAGAAAACGAAATCCAGGCA CGCGTGCATGTATCTCTGATGCTAAACATTAGCACTAGATCAGTGCTAAATGATTGGGTCCT CACTTCCTCAATT
CHR103	AGGCTGGAGGTGTTGGAATC	TCTTGGTCAGTCAAAGCTCC	Chromatin	AGGCTGGAGGTGTTGGAATCAATCTTACTGCCGCTTCTAATGCATTTGTCATGGACCCTTGG TGGAATCCTGCTGTGGAAGAACAAGCTGTCATGCGCATCCATCGAATTGGTCAAACGAAGA CTGTTTCTATCAGAAGATTCATCGTGAAGGGCACTGTCGAAGAACGAATGGAGGCTGTGCA GGCTCGAAAGCAGCGAATGATTTCTGGAGCTTTGACTGACCAAGA

ID	Forward Primer (5'-3')	Reverse Primer (5'-3')	GeneClass	Sequence /GenBank ID
CHR104	ATGCATGACGTAGTTAAGCATGAG	TACTCCAACGGAAAGCCTC	Chromatin	ATGCATGACGTAGTTAAGCATGAGTAGGGAGGGCCAGGCTATGCTAGGACTAGGAGCACTG ACTTGTTGGAAGTTGGAACCCTGACGCTGCCATAACACCAGGACGGACGGGGTCGGCGGT GCGTCGACTGCTGCTGTGCCTGTTGGAACTGATGTACTTTTGGCAGTGTGTGTAACATCGAT GGCAGCGACTAGCGAGGCTTTCCGTTGGAGTA
CHR105	TCTAAAGCAACAATTCACATTGG	GCACAAATCTACGGTAGAATCC	Chromatin	TCTAAAGCAACAATTCACATTGGTTGTTGGTAGCACTAACTGGATTCGTGATTTCGGGGCCG TGACTCGTGTTTTTCTCATGGAACCAATATGGGATAGGAGTATGGAGGAACAGGTTATTAGT CGGGCACATCGAATGGGGGCTACGCGTCCAATACATGTAGAAACTTTGGCTATGCGTGGTA CCATTGAAGAGCAAATGCTTAAGCTTTTGCAGGATTCTAACGCCTGCAGAAAAATAGTGGAT AAAGGAACAGGCAGCACTGACAATGAAGGAGGGCGGTCCCATCGGAGCCTCCATGATTTTG CTGAAAGCAGCTATCTGGCGCAACTCAGCTCTGTGTAGACAGGATCTAAAGCAACAATTCAC ATTGGTTGTTGGTAGCACTAACTGGATTCGTGATTTCGGGGCCTTCCTTCTGATCATTCCTTT GCAATAGTGATAGTATTTGCTTCGAAGCCAGTTGTTGTTCTCGATTCCAAACCAGTGTTGATC GGGGTTCTTTTGTGTAAATAACTGCGTTCCTTTTGTTAAACGCGGTATTCCCAGTTGAAGAAG CCTTATACTGGGCAATATCAGGATTCTACCGTAGATTTGTGC
CHR106	AGGTCGTGGTCCCAACCAAG	ACTTGCGTGTCACTGGCACG	Chromatin	AGGTCGTGGTCCCAACCAAGAGCGGTGGCGGTATGCTCTCGTCGCTCACCAGCTGAGCGA TGCCGGCACCATGTTTGGGGTGCCCATTTTGCTAGTCACTCATCATGTGGACGATGGGGCG CTTGTTGGCCGGTCGCCGATCAGGCGGATATTTTGCATGTATGCCCCCCTACAAACTGTTGA TCTGGGTGTTGAGCCTGCAACAGTAGGGTTAGGTTATTTCCTCCCGTGTCCGTGCCAGTGA CACGCAAGT
CHR110	ATTTATGCAATGATGACATAGG	GAGTGCAACAGTTACTACTGCC	Chromatin	ATTTATGCAATGATGACATAGGTTAGTTATTAATCCTAAATGGAGCATGTATGTGCATGCTGT GTTAGTGTCGTAGATCTGGACAACTGTTGCTGCCCAGCTGACACAGATGCCTTTCGCTGCAA TGCTTCCGGTCCTGCATTTTGCCGCAACTGATTTTGCAACCAAGGCTAAACCCAACGCGGG CAGTAGTAACTGTTGCACTC
CHR111	TGGGAACAGGACCTGCTC	CAAACGCCAGCAGTTCATCC	Chromatin	TGGGAACAGGACCTGCTCCCAGTTAACGGAAGCATTTTAGAATCATTAGGAGCTATGAAGC GGTGTCATTAGAACCGTTTCACCTGTTTAGCGTACCGGTGCGCGTGCCGGCTGTTGTATCTT CTCTCGACGTGTGTTGTAAACTGGGCAGGGACATGGTAGCTAGTGGTGATGTGTTGTGTTA GCAGGCTGACTGCTGTCATCGTAGTGACGGATGAACTGCTGGCGTTTGGAAAAAAAAAAAA AAAAAA
CHR112	TCATTGCCGATGTGTATAGC	TCATTCTCAGTCTGATTCACG	Chromatin	AW257911
CHR113	GAGGCTGAGGCTGAAAATTCTG	TGAAATGTCATCCTCCAATGAG	Chromatin	GAGGCTGAGGCTGAAAATTCTGAAAGAAACTACTGGGATGAACTATTGAAGGATAGATATGA TGTACAGAAAGTTGAAGAACATACTGCTATGGGAAAAGGGAAAAGAAGCCGCAAACAGATG GCTGCCGCTGATGAAGATGACATTCATGATTTAAGTTCCGAAGATGNNAGGATTACTCATTG GAGGATGACATTTC
CHR114	GGCTTCAATACCTGGAGTGC	GATGAGCTGCTGATCGCACC	Chromatin	GGCTTCAATACCTGGAGTGCAAGATTATGCTGATATTGGTGGGTGTGGTGAGATATCTGGAT GTTTGCAGAAAATGAATAGCACACATCAGCAGATTGGAAGACGCTACTGAAGAAGACTTAGG AAGTGGGGACATCACTGTGACCCATCAACTGTACCTGATACTATACTGCAGTGTTCTGCTGG CGAAGAGGTTTCTTTTGTTTTCACCAATCAGGTTGATGGAAAACTTGTTCCAGTTGAATCTAT GGTGCGATCAGCAGCTCATC
CHR117	AACCCATGAACGCAGGAAGC	ATATTCGACAAAACATGAGTTCC	Chromatin	AI622799
CHR119	AACTGGCTCCTCTTGCGAGC	CCCAACTACTGCTAAGTACC	Chromatin	AACTGGCTCCTCTTGCGAGCATGGATGCATCGCATCGAGTTTCTGCTAATACAGCCGACGA AGATGGGTGGATGAAATGATCATTAGACTCGAGCTAGTTAGCTCAACATTCCTTTCTTTTGC GTTCAGCATAACACTGGGATTTTCCCTCTCCCAATGTTTCCAGTGTTGATCTCGTGTTGTGTA ATGCTTTAGTGCAGGTCTCGTTGCAGATAGGATGAAACGCAAAGGGCTAGTGGTACTTAGC AGTAGTTGGG

ID	Forward Primer (5'-3')	Reverse Primer (5'-3')	GeneClass	Sequence /GenBank ID
CHR120	TCTTATTCTTGCAAAGCATGACC	TATGGCTGTATTGCCTGATTC	Chromatin	TCTTATTCTTGCAAAGCATGACCATATTTTGGATAGTAATATTGTGAATATAACCCCTAGCCT GAGCCATTGTCTGCTCAGTTGGGGCGCATCATTTCTCTTCAATAGACTTGAGGAGTTGCAGC AACATAGCTACTCAAATGTCTCTGGTGATGAGCTTTTCATGGATAATGTCGATTTGGAGTTTT TGACAAAATTACTCAGCAAGGTTGAACTCAGAACTGAATCAGGCAATACAGCCATA
CHR122	GAAGAAGGTCAAGCCACTTC	AGTTACAAATTCATGCCAACAAG	Chromatin	AW231410
CHR124	CTGAAGAAGCTACCCAATCC	TGCTACCAAATGTATTCTCAGC	Chromatin	CTGAAGAAGCTACCCAATCCATTAAAAAGGCCGATGAATACCTCCAGTCTTGCAGACCTGTA CGGAGAAAACATGGGAGGTTGCTGGAGGAAGGTGCTAGTGGTTTTCTTGCTGGGAAAGTTC CTGTTGGGGATGATGATTCTGTTAAGTGTCATGAGAAAAGTTGGAGCTCTTTCAATGAACTTA TTAAATCTAAGGAGAGTGCTGAGAATACATTTGGTAGCA
CHR125	GAAGGAACCTTCGGTAATTATG	CTCATACTGCTCTTTGTCTACA	Chromatin	GAAGGAACCTTCGGTAATTATGAAAGAAGAGTCTGAAAAGACTTACAAACTATGGCAGTCTC TTGCTACAAAAAGCATACCAAAGCATCACAAGAATTTTACTGCTTTGATAAAGAAACGACAAG TAGATGCGAAGCGCTTCTCTGATAGTTGTCAACGAGAGGTAAAGCTCAAAGTAAGCAGGTCA CTGAAACTGATGAGGTGTGCTGCAGTACGCACAAGAAAGTTTGTTAAGGATATGCAGATGTT TTGGAAACGTGTAGACAAAGAGCA
CHR127	M13_For	M13_Rev	Chromatin	BM379675
CIN4	CTCCACTACCCAAAATGCCTTC	CTCTAACAAACACCCCCGCATC	Retrotransposon	CTCCACTACCCAAAATGCCTTCATTCAAAAGCGATCAATTCATGATAACTTCCTCTATGTCCA GAAAGTGATTAAAAAACTTCACAAGAGCAAACAAGCAGCCCTATTTGTCAAGCTAGACATAT CAAAAGCTTTTGACTCTCTTAATTGGGCATATCTTCTAGACGTCTTAAAAGCCCTAGGCTTCA CCCAAAAATGGAGAGATTGGATTGCTACCATCCTTGGATCATCTTCATCCAAAATCATCATTA ATGGTCAGCAAACCAAAGAAATCAAGCACATGAGAGGGGTGCGTCAGGGTGACCCACTATC GCCTTTTCTCTTCATCCTCGCAATGGATCCCCTACAACGAATGATTGAAAGGGCAGCTCATG AGGGGCTTCTGGGCCAGGTATTACCCAATGGGGCAAAGTTTCGTTGTTCTCTTTATGCAGAC GATGCGGGGGTGTTTGTTAGA
CINFUL	GAGTACCACGACCAATGATTGC	AGGATTACCCACACAACGATGC	Retrotransposon	GAGTACCACGACCAATGATTGCATTGTAAGGGTATTCCATGTCAACAATATCAAATACAACTT GTTCAGTCCTTGTGTTGTTAATGAATCCGAAGGTTATTGGCATTGTGATTTTTCCAAGTGCTA CGATCTGCCTTCCTCCGAAGCCACAAAGGGGATGTGTGGCATCATGGATTTTGTCTTCGGG CTCTTGCATCTGTCTGAAGGCCTTAGCAAATATGATATCAGCTGCGCTGCCAGTATCAACCA AAACATTATGAACCAGGAATCCTTTGATCACACAAGAGATAACCATAGCATCGTTGTGTGGG TAATCC
CRD109	TAAGGATGCATTGACGCTGC	AGTGACAATATCGAGGAGTTAG	Chromatin	TAAGGATGCATTGACGCTGCTGGGTTTCAGTTGCATGTGCTGCACCTATGTTTGGAACAAAT GGAAATGTATTTAAATGTAGCATTTTTTGTGCTAGATGCATTTGTAGTATTATTTGAGTAAGTA GATAATTAGTAAGTAGCCAATGCTGAGTGCAAGTGAAACATGATAGCATGGTTTGTCAAACT AACTCCTCGATATTGTCACT
DAGN1	M13_For	M13_Rev	Endosperm Marker	Gift from G Marcos
DD1	#N/A	TCTGATAATGTTACATATGGCGAGC C	Endosperm Marker	ATAACCTGGGGTGGTTATACTAGGAGTCCCAATGGAAATGCAAGTCCTCCCATGGGCAATG GGCAATGGCCTGGAAAAAACTCAGCAAGTGTTCAAAATGTTCAATTTGTTGATTCTACTGGC CAGGGATATGCACTACCGGCTTGGGCTCTTCATGTCTCTATAAGCAACAAGAAATGCTACCA AGTTAGCACCTTTTTTGATAGCATGTTCTACTACGGGGGTCCAGGCGGTTGCACTGATTAGC AACCTTCAACATTTGAAGCACCCATATAGCTTCTCTAATGAAATAAAGAGCTTCTTTGCCAAG GCTATATAGTTTATCACCAGTTTTTTTCCCGATAACCTTGTCACCAGTCTTGTATCATCCTCTA TTCAAATC

ID	Forward Primer (5'-3')	Reverse Primer (5'-3')	GeneClass	Sequence /GenBank ID
DMT101	CTTCCATCAAACCATGACCC	TCACCAAAACTAGTTGAAAG	Chromatin	CTTCCATCAAACCATGACCCATGAAGCTAAGCGCTGAGGTCGTCCTTGAGGACCAGTTAATT TTGGTTTTATCAGTCTTAATGGACTCCTGAATGTATATGTTAGAGAAGTGTCGATTGTTGATT GTTACCCTGATTCAGGGTAGCGGTTATATCTAAAAACTTGAGAAAATCTAGTGTACTCTAGTT GCTATGTGTTCCATTTTGTTGACTCTAAACTTTCAACTAGTTTTGGTGATTAATGACAACATGA GATTAACTTAAATTTTGTAG
DMT102	AAGGATGACAGATCTGAGCTG	CAATACACATACATGACGAAGC	Chromatin	AAGGATGACAGATCTGAGCTGAGCTGGGCAACATCCAGCGGCAGGAGCATTTCTGGTTCGG TTCGATTCGGGCTCACGAGCGCATTGTGCTGATTCGCAGGGTGCACTCGAGCTCGAGTGCG GAGTTGTTGTAATGTAGGTTGAGGTTTCAAAANNNNNNNNNNNNNNNGGCCCCGCCTATT TGTGGATGGTGTTTCTCGTGCCATTTTAGTTCAGCTGTATTGATGCTTCGTCATGTATGTGTA TTG
DMT104	TCATAGAGAAGTCCAGTCAG	CATTCACCATGACAGCATAGG	Chromatin	TCATAGAGAAGTCCAGTCAGCCCTGGTGACATTAGCATGGCACAAACACTGTGGCTGCGAT CTGCCAAAATTTCTTCATACGTGATTTAAGCAAGGTTGGTATTTAAGTATGGTTTTGGGGACT ACAGTTTGGCCATTTGATGAGCACACCGTGTAGGCACAACCTCACAATGTATCAACAAAAGT GTACCATACAGTAGAGAAGCTGTTTCTGCGCATACTTCCTCCCCCTCATGCCTATGCTGTCA TGGTGAATG
DMT105	TTTGATTCGAGCCCATCCAG	GTCTTTGGAAACAATAGTGG	Chromatin	TTTGATTCGAGCCCATCCAGTAGCATGTTTACCAATAAATAATCATTGGTCGTGCTGATTCTT ATGGTTGGAGATGAATGTATGTAGGGTGTACTCGAGCTCGAGTGCTTGTTGTACTGTAGGTT GAGGTTTCTCATCCATTGGCCTGCCTATTTGTGGATGACGTTTCATTTCAGATTAGCAATGTG CTTATTTAAGGTTTCGTCATGTACCTGTATTCTACAATCCACTATTGTTTCCAAAGAC
DMT105	TTTGATTCGAGCCCATCCAG	GTCTTTGGAAACAATAGTGG	Chromatin	TTTGATTCGAGCCCATCCAGTAGCATGTTTACCAATAAATAATCATTGGTCGTGCTGATTCTT ATGGTTGGAGATGAATGTATGTAGGGTGTACTCGAGCTCGAGTGCTTGTTGTACTGTAGGTT GAGGTTTCTCATCCATTGGCCTGCCTATTTGTGGATGACGTTTCATTTCAGATTAGCAATGTG CTTATTTAAGGTTTCGTCATGTACCTGTATTCTACAATCCACTATTGTTTCCAAAGAC
DMT106	CAATGGGCATGGATTCGAAC	TTGCTTGCAATAATCAAGTCC	Chromatin	CAATGGGCATGGATTCGAACCGCTTCTATGAGTATAACCGTGTTGTCAAAAGGGTAAGGACT GCAGTTGGACTCAGTTGAAGAGTCCATCTTTAATGAAGATCGCACGTCTTCATAAAACTAGC CTGTAGGAGGATATAGAATCATCTGATCCGTAAATGTAGGATGCTACAAGTAAACTTTACAA GTCTTGGGACTTGATTATTGCAAGCA
EA1	T7	T3	Endosperm Marker	Gift from G Marcos
EA3	T7	T3	Endosperm Marker	Gift from G Marcos
END1	#N/A	CTAAGAGGAACTGTATAACTCC	Endosperm Marker	CAGCATGTTGTGCTGTGTGGAAGAGGGCAGATGTTCCATGCCTATGCAAATATGTTAACAAG GAAGTAGAGAAGGTATGGTGCATGGAGAAGGTTGTGTACGTCGCCAACTATTGTAAGAGGC CGTTTCAACCTGGCTACAAGTGCGGGAGTTATACAGTTCCTCTTAGTATGTAAAAGGAGTGG TTGGGACATGTGGTACTCTCTCCTAGAACTTTAATTTGAGATTTGTAACATGTGAGAGACTAT TGAGAATAATATGTATGGAGTAATATGCTTCCTTCTTTTAAACTGATCTAATTACTTCCTTGGT TTTGTGTTTT
EPL101	M13_For	M13_Rev	Chromatin	1Cgh1

ID	Forward Primer (5'-3')	Reverse Primer (5'-3')	GeneClass	Sequence /GenBank ID
ES1	#N/A		Endosperm Marker	GAATAGTTCCACCACGTTACTTCCATATATATTTCCCTTGGATTGGATTGGATCGTCGGCGC CCAAACGAATAATAATCCGGCAATGGAGCCTTCACGAGGGAAGCTGTCTGCCGCCGCCGTC CTCCTGCTGATGACGACGCTCCTCGTGGTGGCCGCCATGCGGGCGGTTGAGGCACGCGAC TGCCTGACACAGAGCACCCGGTTACCGGGGCACCTGTGCGTGCGGTCGGACTACTGCGCG ATCGGGTGCAGGGCGGAGGGCAAGGGCTACACGGGCGGCAGGTGCCTCATCTCTCCCATC CCGCTCGACGGGATTCTCTGCTACTGCGTCAAGCCGTGCCCATCCAACACGACGACATAAT GATGAGACAAAGGGCGGTGGGTGCAGTGCACGCTGGCCGGGGGTTATCAGTCCACACATC CTACCGTACGTGTCTGTGTTAATAACTTTTTTTTGTCTTGGAAGTCTGTGGTGGTAATGACTT TTAAATGTCTTGGAATAAACCGGGTTCTAGTCCCTTATAAGCTAGCAGTACTGTAACAATTCA GATCATCAAAGACAGCAGGAAACCGGTGGTTGAGT
ESR2	T7	T3	Endosperm Marker	Gift from G Marcos
FIE101	M13_For	M13_Rev	Chromatin	Gift from G Marcos
FIE102	M13_For	M13_Rev	Chromatin	Gift from G Marcos
GTA101	CACACAGGAAAACTTATAGG	GCAAACGAGGGCATTGTTAC	Chromatin	CACACAGGAAAACTTATAGGTATAGATGGTTCTGACGGAATAGTAAGGCTTGACGACACATA CGAAGTTAAGATATTAGATATGGTGATTTTGGCCAAACTGGCGACTTGAAATGGGATGTATA TTTTGTTTTGTAACGTGTCCCCCATAGAACATATTTTGTAACAATGCCCTCGTTTGC
GTA105	T7	T3	Chromatin	BG349445
GTA106	GCCGAAGAAGAGCGACAGG	TAAGCAAAGCTTTATTGCAAACG	Chromatin	BG841733
GTA107	M13_For	M13_Rev	Chromatin	BM895641
GTB101	CGAGACACATGGCAGCAGTTG	CTCACGTGCTTCAGACATGAG	Chromatin	CGAGACACATGGCAGCAGTTGAGATAGGTTGTGAGCCAATTGTGAGAAAGCATGTACGAGG CATCTTCATGAATAAAGCAGTTGTGTCAACAAGTCCCACATCTGAGGGAAATACAATCATAG ATCCATATCACCAACTATCGGGTGTTAAGTGGCTGAGAGAGAAACCACTTAGCAAGTTTGTT GATGCACAGTGGCTCCTTATTCAGAAAGCAGAGGAAGAAAAGCTCCTCAAGGTTACCGTAAA ATTGCCAGAAGATGCAAAGAAAAAACTCATGTCTGAAGCACGTGAG
GTC101	GTAGCTATTGCTCAAATGTTGC	TACGATCATTCTTGGCATTCAG	Chromatin	GTAGCTATTGCTCAAATGTTGCTCGGACATATCTTATCGACGCTACCCCTACACAGAATAAA GCATATGAGACCCTTAGGAAAGCTCATGAAGCTGCTATTCAACAAGTGAAACCAGGCAATCA GATGAGTGCAGTTTATCAAGCTGCAGTGGCTGTGATTGAGAGGGATGCACCTGAACTTCTTC CTAATCTAACAAAGTCAGCTGGAACTGGAATAGGCCTTGAATTCCGAGAATCTGGCTTAAAT CTGAATGCCAAGAATGATCGTAGGATAAAAAAGGGAATGGTTTTTAATGTATCCCTTGGTTTG CATAACATCCAGGCAGAAACAACCAGTGAGAAGACAAAGCAGTTTTCACTATTATTAGCTGA TACAGTTTTAGTCAATGAAAGAGGACATGAGATCTTAACAGCACCCTGCTCCA
GTD101	CCCAACTCTATGGATATAAACC	AGATCTACTGTCTGTAACATCC	Chromatin	CCCAACTCTATGGATATAAACCATTTGCAAGCAAAGACACCTAAAAGGTTGATAACAAAAGTT TCCTCTGAAGTTACTGAAACTGAAGGGCCAGAAGGTATTGAGAAGACAAAATCTGTCCCAGT AAAGTTCAAAGTAGGGGCACCTGATAAATCCTTGGAGAGGAACATGTCAGTCTCCGTTTCAC TCGTTTCTGATAAACGTGTCATGGATGTTACAGACAGTAGATCTACCGGAAAGGTA
GTE101	GATATAGGTGAGTTAGAATATC	CCCATATCTTATTTAATCAG	Chromatin	GATATAGGTGAGTTAGAATATCAACATTTTGTGCTTTTAAGGTTGAACATATATAACTCCATG CAAATGCAGTACTTTTGTTTCAAGGAATATTGAAAGCGGCANAANGTATTANNCTNCATTGTT AGGTGTGAACTCCTAGTTGCAATATTTATGGTTAACCACCATTTAATAGTTCTTTTTACCAGTA CTTGTACACTTGTCTATCTGATTAAATAAGATATGGG
GTE102	AGGTAGACGATGAGGAAGAC	CAAAATGGACGGCTCTCACC	Chromatin	AGGTAGACGATGAGGAAGACGACGAGCAGTGTTCCCTGCATTAAAGAAAGAAGAGAAAGGA GAAATCAGCTGATGGTGCTGTGTTTTTCCATGCCCTGTCACGGGCTAGGGGCTGGTGCCAA CCTATTGATCCAAAGTGATCACCTGTTGCGGGCCATCATGTGTGGCAATGGTTCTAAACTGG GGTGGTGAGAGCCGTCCATTTTG

ID	Forward Primer (5'-3')	Reverse Primer (5'-3')	GeneClass	Sequence /GenBank ID
GTF101	CTCGAGGATTTTCGTGTGACC	CTTGACTCAAGTTAACGCTGC	Chromatin	CTCGAGGATTCTCGTGTGACCATGAATTCTAAATGTACATAAAAGTCTGGCGGAACTTATAG CATCCTGTAATCTCTGTGAGCCGCGGTCTGTCAGTATTGGGTGGGCTTGCTGTGGTAACTG GTGGACATGTCGAGTTCATCCATTTGGGTAGGCCTTGGTGAACACGGACGCCCCACATCAT GTCTCAGTTTGAGTAGTTTGTGTGAACTGAAGGGTACGTGCTTCTCCCATGTGGTTGCAGCG TTAACTTGAGTCAAG
GTF102	AAGAGCTAAGATGAGAGATG	GTGAGGGTCCCAGTTAATGG	Chromatin	AAGAGCTAAGATGAGAGATGAGACGTTGCTTTTGAAATTATCTATCCAGTCCTCATAAAGTCA TAAAAGTCTAGCAATGCTAATTACATGGTAATTGCTAATGGGCAAGAATAATTAGGGTATGTT GTTCAAGTGTTTGCCCAGTTGTTGTCCCCCCTTTTATCATTTTGTTTTCTGATTTCGACGATA CCATTAACTGGGACCCTCAC
GYPSY	GCGGTCTAACAGGTCATTTTGG	GCTCACAAACTTCACATCACGG	Retrotransposon	GCGGTCTAACAGGTCATTTTGGACAAAAGAAGACTTATGAAATGCTCAGTGATCATTTCTATT GGCCCAAGATGAGGAGAGACGTCATCAGATTTGTGGAGCGTTGTGTCACCTGTCACAAAGC TAAGTCAAAATTGAACCTTCATGGTTTATATACTCCTTTACCCGCTCCTAAAATTCCTTGGGA AGACATAAGCATGGATTTTGTTTTAGGACTGCCGATGACTCGGAAAAAGAGAGATTCTATTTT TGTTGTGGTTGATAGATTTTCAAAAATGGCTCATTTTATTCCATGCAACAAGAACGACGATGC TTCACATGTTGCTAATTTGTTTTTCAGAGAAATCGTGAGACTACATGGAGTGCCCAAGACCAT CGTATCTGACCGTGATGTGAAGTTTGTGAGC
HAC101	TTGTGCATGCATCGAGGTGC	GCTATGAATGCCTTTACATCCC	Chromatin	TTGTGCATGCATCGAGGTGCTATGATCCACGTAACTGCACCCATCCACTCTGTTTAACCCTG AAGAAACTGTTCGCCCATGGTGCACGATGCCATGTCCGTGCCTGCAATAGCGGCGGTTGCA AAAAGTGTCTCTTTATGTGGAAACTCCTGTTCCGCCACTCAAAAGATTGCAACGACGGAGAC TGCGCGGTTCCTAGATGCAGGGATGTAAAGGCATTCATAGC
HAC104	CTGGCGTGGGCTCACTGG	GGAAACTGAAGAATCTCCTC	Chromatin	CTGGCGTGGGCTCACTGGACTGGGCACTACTGTCGCATGTTGTAAAATTATTATTGGCTCAC TGGACTGGACACTACTGTCGCATGTTGTAAAATTATTAAGTTGTATTGATATAGTTGTCACCT TGCTGAATGTTCGGCACGGTGATGAACTTTTAGTTTAGCATTATTTTAACCAGGAGGGACAC TGATTGATCTTTACATTTCGGTCTCAACCTGGCCGGCCTAATAATATAGATTGAGGA
HAC105	TTTCGCGCCAGTGGAGTGG	TTTCCAGAAGCACTCTAGTCC	Chromatin	BG355090
HAC106	CAAATCGAGGAGATTGATGG	AAGTATATAATAATCTGTCGTG	Chromatin	AI978143
HAC108	GACGTTAGCAAGCTCATATG	GGTAGCAGGGTTAATGCAAGG	Chromatin	GACGTTAGCAAGCTCATATGGACTCCCTACAAGGAGCAAGGCTAAAAGGCGAACTGCTCCG TGTCTCAGATTTAGACTCGCTGACCGATGCGATGCAGTCCCTGTACATAGCATCTCCCCACA GCTCACTTGACTTGACTTGACTCGTGTACCGTGAAACTCTACTTGTAACTTTGAACATCCACA TCGTTGTCTCCTCATTTCTCCTTGCATTAACCATGCTACCTTTTTTTTTAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAA
HAC109	ATCCAAATCGAGGAGATTGC	GTGCCTATGCTTGTGTTGG	Chromatin	AW202543
HAC109	ATCCAAATCGAGGAGATTGC	GTGCCTATGCTTGTGTTGG	Chromatin	AW202543
HAC110	TCCTTGGCAGTGCCTCTGG	CCAGCTCTGTCATCTTATCC	Chromatin	TCCTTGGCAGTGCCTCTGGAACTAACAGGAAGCGTCTGTTGCAGATGCCAACCAGTTTCAAT AAGGAGGCATCTTTTGCTGTGTATTGGACACAGGAGATTGAGGATGAGGATGAACAAACTGT TGAGCAGCAGCCAGAGGATCTCAAGACGCAGGAGCAACAGCTCAATGAGCTCGTGGACAT CCAAATCGAGGAGATTGCTGGGGTCGCCAAGAACGTCACTTCACGTTGCAAGGATAAGATG ACAGAGCTGG
HAC111	GTAGCTGCCTCTGTACAGGG	CCAAGGGTTGAGGAGACAAC	Chromatin	GTAGCTGCCTCTGTACAGGGGTATATATCTTTGTGGAAAGAGGATAGGAAGTACGCAGGGC AACAAGAAAGATTGGGGCGGTTGAATTTGAAGGGGGGGATTGGTTGGTGAAGACGAAGCG CTTCAACGTGGATTTCTTATCCTCTGCCCCGTCGTCCTCCAAAGAGCTCAAGAAGGCACACC ATTACTGATGTTAGGCAGTGCTAACCAGCCACACCCTGCTCACGTTGTCTCCTCAACCCTTG G
HAC112	T7	T3	Chromatin	BG549748

ID	Forward Primer (5'-3')	Reverse Primer (5'-3')	GeneClass	Sequence /GenBank ID
HDA101	CTGCGACCATCAAGAGTTGC	AATGACTGTACAACTGAGATC	Chromatin	CTGCGACCATCAAGAGTTGCTGGGTTTCTGCTAGGCAACTTCCTGCTTGATTTGGACAGTCC AGAAAACAGAAGTGAATCCGCAATACGTCCCTGCCATTTCAAATTTGACTTGGTTACACCTAT CTGTGACAATGTTAATGTAATGAGATCCCATGGCACTAGTGAAGTTCCCAGGTGTCCACCTG GGATTGCTTGCGGTTGAGAGAGACGATAGTCTGATGACCCATGTATACATGAAGCCCATGAT CTCAGTTGTACAGTCATTAGCCAATACAAATATGCAGTGGACTAAAGGGTGATTCATTGCGT GTT
HDA102	CTGGTGTAAATGGAAACCTGAG	AGCAATGTTACTGAGCGAGG	Chromatin	CTGGTGTAAATGGAAACCTGAGATTCTTCACTGTTGTGTGTAGTTCAGCAGCCTTGAGATGT AATAGTTTGTCAGTTGACAGCAGGGATCTAATTTAGCAGGTCAAAGTGGTTTAGACTTTTATA AGGGATTCTGACCCCTCCTGAATTACACATTGTAGTACAAGTCTGCATATTTTTAAGCATGCA AAAATTCAAATTTCTCAGTGGTATTAGTATCTCAAGAAAATACAAATTCCTCGCTCAGTAACAT TGC
HDA104	TCACTCGAAGGCGAAGCATG	CACACAACCCAAATATTGAAG	Chromatin	AI901627
HDA105	AGGCTCACTCGAGGGCGAAG	AACACACAACCCAAATATTG	Chromatin	AGGCTCACTCGAGGGCGAAGATGGGGGCAAGTGAGTCGCAGGTCCAATAGAGTCCGCAAC AAATGCGAAACATGGGAGAGGAGGGTAAGCGAGAGTCTCGAAAGAGTGTCGGTGGAAGTA GGCCTAACCTTATTTTGTTTAGAGACGGGCTATGCGTTCGATGTAGCAAAACAAGGCTGTGG TTTGTGTACTTCAATATTTGGGTTGTGTGTT
HDA106	CCACAGCTGACAAGAAATCC	CAAGGCAAGTTCCAAACTTG	Chromatin	CCACAGCTGACAAGAAATCCCCCAAATCTGGGAGCCATGGCTGCAAGTAACGCAACAAATC TTTCAGCAGTGCGTCAGCACTTTAGTCCCATCAGAAGCCAAAAAAGCATGAAGCCTAGAGGA GCATGTGCTGTACCTTATAGGTTGATCTACTTGAATTATCAAAATCAGTGGTCTTATTTTTGAA CTTGAA
HDA108	CCCAAAGTTACTGAGTTGTG	CATGCGAGAATAATCACTTC	Chromatin	CCCAAAGTTACTGAGTTGTGAATGGGATCGTCGACATGCATGTTGAAGAGTGGGGACGCAT CTGAACCTGTCCGTGCTGATATCTGCCGCTGTGCATCTGTAACAGACTGAACGTTTTGCTGG TTCCAGCTTTCCTGCGTGTTGGTAACTAGACTAGATATATGGCCCTTTGGTGGGTCCCAAAC TGTTGAGCTGGAGAAATGAAGTGATTATTCTCGCATG
HDA109	TATACCATGCTGCGGATCAC	AATGACATCCAAGGAGAGTCG	Chromatin	TATACCATGCTGCGGATCACGTGACACAAGCTTACCCAAAAGGAACGTGGTGGAGACTG TCCCTACACGTTGCTTTTTCGGAGATCGCTAACCCTTGTTGGGAACCTTGCGTGTTGCTT TCTCAACCCTGAGATAACTACATTGAAAGTCTTTGTCTGCTGTCCAAAATACGTTGGACA GTCAATGACTTTTGTTCAAGAGAATCTTTATGCTTGAGATGCAAAAACGGAAACCACGGG ATCCGACTCTCCTTGGATGTCAT
HDA110	GGAACTGGAGCTGCTCATGAGG	CTTCCATTTGAGGAACCAGC	Chromatin	GGAACTGGAGCTGCTCATGAGGTGGGAGTTCTTGATGGTCAAGGATTCTCAGTTAATATACC TTGTATCCGTGGTGGTGTTGGAGATGATGACTACATCTTTGCTTTTCAGACTGTGGTGCTTC CAATAGCTTCAGAGTTTGCCGCAGACATCACTATAATATCTGCAGGATTCGATGCAGCTAGA GGTGACCCTCTGGGTTGTTGTGACGTCACTCCAACTGGATACTCTTGGATGACATCCCTGCT AGCTGGTTCCTCAAATGGAAG
HDA113	ATGCATGGGAAGTTCGCATG	TCGGAGGCAGTACTTCATCG	Chromatin	ATGCATGGGAAGTTCGCATGCTAAGGTTGAACTCTTCTGCTGCATAGGAAAGTTGGCTGACT TCTTATATTATGCGAGCCTAGCTTAGGTACCTCCTGTGGGGAGAAGCGATTGCAATACATGG TAGCATGTTGTAGAGCTCCTCACCCCAGTTAGTCAGTGACTGATGATGGTATCCTCTTACCA GTCCTCTGGTGGTATCATCTTTTGTTCACGGAGGTCGATGAAGTACTGC
HDA114	TGTACCATCTACTCCTTATC	TTACGGGGCACAAGGCTCAG	Chromatin	TGTACCATCTACTCCTTATCAAAATCAGTGGTCTTATTTTTGAACTTGAAAATGTCCAAGTTTG GAACATGCCCAGTAGTCATTTGATTAGTGGCCTTATTTTTCTTTGAACTTGGGAATTGTCAAG TTTGGGACATAAGCAGTATGCCTTGTTAAGAACATTGGAAATGTCAGGTTAGAACAGGTCCT GCTGGCCCAGTTTTGGTCGTCGTCGAATGTTGCTAATGTTNGAACTGAGCCTTGTGCCCCG TAA
HON101	T7	T3	Control (positive)	AI691318
HON102	PGAD10	PGAD10	Control (positive)	BE056083

ID	Forward Primer (5'-3')	Reverse Primer (5'-3')	GeneClass	Sequence /GenBank ID
HON105	T7	T3	Chromatin	AI770331
HON108	PGAD10	PGAD10	Control (positive)	BE518674
HXA102	CAGCTGCTTATCTTGCATAGG	GGTCAAACTCGTGTCTCTTTGC	Chromatin	CAGCTGCTTATCTTGCATAGGGTGGAAGATGCACTTGGAGAAGGTCTAGCAGGTCGATCAC CTTCGCACATAGCTGTTGGTGCAAATAAAAAAGCTTCAAATGTGGGACATATTAAAGATGGG TCTAATGTATCAAAAGTTGAAGATGGTCATGTCGATAGAAGTGTTGGTGTGAAGAAGCCCAG ATATTCTGCAGATGAAGGCCCTTCGTTGACTGAACTGAGTGGATACAATGCAAAGAGACACG AGTTTGACC
KN1_Knotted1	M13_For	M13_Rev	Control (negative)	Gift from G Marcos
MADS3	CAGAAGGAACTTTCAGAGAGGCAG	CATCATGAAGGAAGAAGACGATGA	Endosperm Marker	CAGAAGGAACTTTCAGAGAGGCAGAAGGCGGTCGCTAGCCGGCAGCAGCAGCAGCAGCAG GTGCAGTGGGACCAGCAGACACAGGTCCAGGTCCAGACAAGCTCATCGTCTTCTTCCTTCA TGATG
MBD101	GATACTCCCAACGGTGCTTG	GCAGGTCTACACTCTACAGC	Chromatin	GATACTCCCAACGGTGCTTGTGTTATGTTTAGATGATAGTCCGCTCAAGATCATGCCCCATC CTATCTAGTGTGTACTTGTCTATCTCGACTGAAGTCGATGTTTAGGTCTTTTTAGCTGATTTG GCTCTGCATGCTTCGTTGCAACGAAGCCCCTGCGTGTCAATTCCTGGGGAATAAATGAGAT GTTGAATAAGTGCTGTAGAGTGTAGACCTGC
MBD105	TGAAGTGCACCTGAAGCCGC	CGGCGACAGCAGCCATAC	Chromatin	TGAAGTGCACCTGAAGCCGCAACGCCGGAAAAACCTGAACCCCTTCCGGCGCCGTGTATCT CATTTAGGAACAAACATTCGCATTCGCATTCACATTGTTATCAGTTAAAAAGTCTATGGAACG GTAAGCATGTTAATTAGTCAGTCTGACGCTGCTGGGTGACAGGTTGGTAATTAGGCTCTGCA TTTAGCTGTTTTGTATGGCTGCTGTCGCCG
MBD106	AGTGCATCTGAGGCCGGAAC	CATGTACGGCGGCGACAGC	Chromatin	AGTGCATCTGAGGCCGGAACGGAACACCCCTTCCAGCACCGTGTATCTCATGTAGGAACAA ACATTTGCATTCGCATTGTAATCTGATGGATCGGTAAGCATGTTGATTAGTCAGTCTGACGCT GCTGGGTGACATGTTGGTAATTAGGCTCTGCATTTGAGCTCTTTTTTTTTGTATGGCTGCTGT CGCCGCCGTACATG
MBD108	AAAGTAAAAGTGGGCAGCG	TATCCAAACAAGTTCCATAGG	Chromatin	AAAGTAAAAGTGGGCAGCGTCGTGATGCCAACTTACCCCCCTGCCATGGTTTTATTGGTCCT CTGCTTGTTTCTCGTTTTCCTTGGGAGTCAGCTTCTCTGTTTTATGTGGTCGCACACGGCCTT ATTTGTAGTATTACACCACAGAAGCTGGCTGACCCAAAAACCGACATTTTGGGGGCAGCTAC TTAACGAAATGTTTGCACGAGAGTTTAGAGCCCTATGGAACTTGTTTGGATA
MBD109	M13_For	M13_Rev	Chromatin	1Cgy2
MBD111	GGAGACGTCTACCTGATAAG	CATACTTATCCAAAGAAGCACC	Chromatin	GGAGACGTCTACCTGATAAGTAAAAGTTAGCAGCATCGTGATGCCAACTGCCCCCTATCATG GTTTTATTGGCCCTCTGCTTGTTTCTCATTTCCCTTTGGAGCCAGTTTCTCTGTTTTGCGAGG TCACACAAGCCTTATTTGTAGTATTACAACACAGAAGCTGGCTGACCCAAAATCCGACACTC TGGGGGCAGCAACTTAACAAAATGTTTGTACGAGGGTTTAGAGCCCCACGGTGCTTCTTTG GATAAGTATG
MBD113	T7	T3	Chromatin	BG316880
MEG1	TGCTGCTCATGCGCATGGGGCTG	TTAGAAGCAHGCATGWCTACACACTSAGCC	Endosperm Marker	TGCTGCTCATGCGCATGGGGCTGAAGAAGGAATTTTGCGAGAAAAAAGAGCACAATGCGCT CAAGGGTTTCTTCCATGCAAAGATAACAAGTGCTACTGTTGCATTGGGGGCCGAACTCATGA TTGCTACTATACGATGGCTCAGTGTAGTCATGCATGCTTCTAA

ID	Forward Primer (5'-3')	Reverse Primer (5'-3')	GeneClass	Sequence /GenBank ID
MRP	GTAACAACCAGGTAGCTACATGG ATTAC	GGTAAAATTTCATTGTATTGCTCAG GGGC	Endosperm Marker	GTAACAACCAGGTAGCTACATGGATTACAGACCAGCAGGCAACTGCAAGTTCTTCTATAGCT CCTCCAGCGACGGAGGAGAGCCAGATATATAACCGATAATATATATAATGGTCATCAGCAGC TGGGAGAGGCTTTCTTCATATATAATCAATAGGTAGATAGATATGGACAACGTCCATTGACTA GTTTAATTTCTATCTATATGTTTTGTATCCAATGATGCATGTAAAACCTAGTTGGTTGTTAAAG GTCATTAGTACCATACTATATATGGGCTAGAAACAGTTTCATTGAAATTTGCCCCTGAGCAAT ACAATGAAATTTTACC
MYB			Endosperm Marker	CACAGAGAAACAACGCTTTAGCATCAACGATGTTGGCCTCTACGACACACAGCCATGGGTG CGGCAGAACAACTCCTCTAGCAGCTGGGAGGCGCTCACCTTCACTGCTGGCCGTGCGTAC AATAATACAAACTACTGTGCCTTTAACAGCCTCCCGTATGCCAGCAGCCAGGCAAGTAACAA CCAGGTAGCTACATGGATTACAGACCAGCAGGCAACTGCAAGTTCTTCTATAGCTCCTCCAG CGACGGAGGAGAGCCAGATATATAACCGATAATATATATAATGGTCATCAGCAGCTGGGAG AGGCTTTCTTCATATATAATCAATAGGTAGATAGATATGGACAACGTCCATTGACTAGTTTAA TTTCTATCTATATGTTTTGTATCCAATGATGCATGTAAAACCTAGTTGGTTGTTAAAGGTCATT AGTACCATACTATATATGGGCTAGAAACAGTTTCATTGAAATTTGCCCCTGAGCAATACAATG AAATTTTACCAATGTGTTATTTATATATTAATGTGTC
NFA101	TGGCCTGTTGATGTTGGTTG	TAATATCATCCGATTGCACG	Chromatin	AW067612
NFA103	ACGATGAAGGCGATGATTTG	ACCGAGATTATGATATAAGC	Chromatin	AI941570
NFA104	TGATGACAATGACGATGATG	CAAACTGCGTTGATAGATATACC	Chromatin	TGATGACAATGACGATGATGGGGACCAGGAGAACTGAGCCCTCGCGTTTAGGCGGGGAATT ATGT
NFB101	CCAACCTCTTATTATATGCAACC	CACCTGTACAAATTCTGGCAGG	Chromatin	AI881422
NFC101	CATGTGAAAACTCTGAGTAGG	GAACCGGTTGTTGATAATATCC	Chromatin	CATGTGAAAACTCTGAGTAGGTTATGTAGTTTCCTGGCCCTAAGTAGTTTCCTGGCCATAAG TATGGCCCCAATTGACATTGGATCTGTCTTCTCATGCTCCTAGAGCTGTACCCTTCGAGGTG CTTGTGCCTTTGTGAGCTTGTGTTGAACCATGTTGTATGACTAAATAGTCCGCAAACGAGTC AGCTGGATGGCTAATCTGTAGGCCAACTATTGGGATATTGTTGGATATTATCAACAACCGGT TC
NFC102	AGTCATGTGGGTTACCTGCG	TCCGGCTCTGACGCGTTTGC	Chromatin	AGTCATGTGGGTTACCTGCGGAATGCAATCCTGTGGAAACTGAGTAGGTGATGTAGTTTCAT GGCCATAAGTATGAGCCCCAGTTGACATTGAATCTGTCCTCTGATGCTCCTAGAGCTGTACC CTTGAGGGGCTTGTGCTTTTGTGAGCTTGTGTTAAACCTGCTGTACGAGTAAACTAGTATGT CCGCAAACGCGTCAGAGCCGGA
NFC103	GGAATCTTAACTCCTATCGAC	CACAACTGCGACGAACAGATC	Chromatin	AW438124
NFC104	TCGTGCCTCGAAGCATAAAGG	GATTACTCAAGTCAATAACTGAC	Chromatin	TCGTGCCTCGAAGCATAAAGGTGAGTTACCTGACCAATGATGCAAGCGTATCTGGAGGGGT TATATACCTGTATGATTTGAGAGATTTCAATTTCGCTATTTTTGTTCACTGCAGCCAGAGTGG GCATCTATAAGAACTCTTTAAATCCGTGACATGCATTAGGCTATCTCCAGCAGATCTCCTATC ACATCCTTTATGTTATCCCCTATTTTCAAACTTCTCTCTGCAAAGAGTCAGTTATTGACTTGAG TAATC
NFC108	AGCAACACAAAGCAGAAGAG	GCATCATCGACTACATTTACTC	Chromatin	AGCAACACAAAGCAGAAGAGATGCTTGTTACATATACTACGTACCGAGAGATTTCAGAATGA TGATGCAGTCCCCTGATGACTCTCATATATATAGATAGGTATGGGCCAATATGAGCCTTAAC ACACCCCTTCAAACTCAAGTCGGATCAGAGACATTGAGTTTGAGTAAATGTAGTCGATGATG
NFD101	GTATGACGGATGAAACTCTG	AAAACAGGCACCATAGATGC	Chromatin	GTATGACGGATGAAACTCTGTCCAAGTGTTTGCAGTCAGGGAAAAAATCTCGTCGTTGTCGG CCTAACCGGGTTTTCTCGAAGGACTTTCTAGATGCAGTAGATTAGTGCCGTCTTCTGAGTAG TTTCCTTGCAAGTGTCGTCCCGTCCCTGGAGTGAGGCAGTTTAGTTTGTAGTTTCTTCGCTT CATCGGCACTCGACCGTTGTTGGATTGGGATGTTTAGCCTGCAGTCTGTACTTACCCCAAGC ATCTATGGTGCCTGTTTT
NFD102	ATAGTACGATGGGGAACAGC	TACAAATTAGTGAAGCAATCC	Chromatin	AW017485

ID	Forward Primer (5'-3')	Reverse Primer (5'-3')	GeneClass	Sequence /GenBank ID
NFD103	TGGAGCTCCTCGAGACAATGG	CAGTGGCGCCAACGAGG	Chromatin	TGGAGCTCCTCGAGACAATGGACCGTGCTTCATCCAACAATGGAGCGGCTACACAAGGCCC CGTGGCGATCACAAAAAAGGAGCCTATATCCATGTACTAGAATTATTCAGTTTCACTCCACAT CGTGATGTTTTATTTTTACTTTTGTCGTGCTATAACGGATAGCGCTCCTCGTTGGCGCCACTG
NFD104	CAGTCTCCAAGGTCACGAGG	CACTAGATTAATCAATGAAGC	Chromatin	AI619085
NFD105	TGACGAGGAAAACGACGACG	TAATAGCTAGCAGCATTCGG	Chromatin	TGACGAGGAAAACGACGACGACGAGTAAATAGTATCTCCGCTGTCGTGAGGTAGCAAATCG ATTGTTAGGTCCCCAATTAGCTCACTGGTTTGCTATCTTTAATGTTAAGTTGTTTGGAACAGG TGTTTATGAAGGGGATCTATTAGTTTCGTGAGGTTGTTAGTTTAGTGCATTAGCAAAGAAAAA GCCGAATGCTGCTAGCTATTA
NFD106	GACCAGGAAGACGAGGATG	CAGACAGAACTGCATCATAG	Chromatin	AW267641
NFD107	GGATGACGGGTGAAACTCTG	GTATAAGTGCCGAAGAAGCG	Chromatin	GGATGACGGGTGAAACTCTGGGCTGCCCAAGTGCTTGCAGTCAGCAGGGAAATCTCATCGT CGTTGGGCCTAACCGGGCTTTGCTGCTCCGTAGCGTCTGGAAGCTGTCTGTCTTTTGAGAT GCTGTAGACTAGTAGTGCCTGCTCTGAGTTAGTAATTGAAAGTGTTGTCCCTGAAGTGATGC TGTTTAGTTGGTAGTGTGTTCGCTTCTTCGGCACTTATAC
NFD108	M13_For	M13_Rev	Chromatin	AI586679
NFD109	FULL INSERT		Chromatin	AI467690
NFD110	T7	605R	Chromatin	AI677100
OPIE	CCCCCAACTCAAGACGATGAAC	CCTCGAAATCCAAACCTGCGAC	Retrotransposon	CCCCCAACTCAAGACGATGAACAGGTTCATCAACAGGAGGTGTGTGATCAAGGGGGAGCAC AAGATGATCATGTGTTGGAGGAAGAAGCGCAACCGGCACCTCCAACCCAAGTTCGTGCGAT GATTCAAAGAGATCATCCCGTCGACCAAATTTTGGGTGACATCAGCAAGGGAGTAACTACTC GATCTCGATTAGTTAATTTTTGTGAGCATAACTCCTTTGTCTCTTCTATTGAGCCTTTCAGGG TAGAAGAGGCCTTGCTAGATCCGGACTGGGTGTTGGCCATGCAAGAGGAGCTCAACAACTT CAAGAGGAATGAAGTTTGGACACTGGTGCCTCGTCCCAAGCAGAATGTTGTGGGAACCAAG TGGGTGTTCCGCAACAAACAGGACGAGCGCGGGGTGGTAACGAGAAACAAGGCTCGACTT GTGGCAAAAGGTTATGCCCAAGTCGCAGGTTTGGATTTCGAG
PP1_Elongation factor 1A	605F	605R	Control (positive)	
PP10_Tubulin ?	605F	605R	Control (positive)	
PP11_Tubulin ?3	605F	605R	Control (positive)	
PP3_Histone H4	T7	T3	Control (positive)	605004D02
PP4_High Mobility Group1	T7	T3	Control (positive)	606045H01
PP5_MNBB	605F	605R	Control (positive)	
PP6_Photosystem II	T7	T3	Control (positive)	603025F08

ID	Forward Primer (5'-3')	Reverse Primer (5'-3')	GeneClass	Sequence /GenBank ID
PP7_Rubisco Large Subunit	T7	T3	Control (positive)	687006H08
PP8_Rubisco Small Subunit	T7	T3	Control (positive)	606002C02
PP9_Tubulin ?1/2	605F	605R	Control (positive)	
PREM1	CATTGCATCGCAGGTGATGTTG	AAAGTTGACTGGGAGACGACGG	Retrotransposon	CATTGCATCGCAGGTGATGTTGTTACTTTTGTTCTACTACTTAATTGGGTTGGTATTTACTTAT ACTTAGTAACTGCTAATAAAATTTTGACCAACTTTAAAAGCAATGCTCAGCTCTAACCATCCT CTTTGGTAAGCCTTACACTTCACGTGAGCTCCCACCTTTGGCGAGTTCATGCACATTATTCC CCACAACTTGTTGAGCGATGAACGTATGTGAGCTCACTCTTGCTGTCTCACACCCCCCCCCC CACAGGTCAAGAACAGGTACCGCAGGATGAGGCGCATGGAGGATGCTGCGATGAGTTCGT GAGAGGTCTAGGCCGTCGTCTCCCAGTCAACTT
PREM2	TGATGAGTGCCAACTTGCTTCC	ACTCGACAGCGGATGCACAAAC	Retrotransposon	TGATGAGTGCCAACTTGCTTCCCTGCTTGGCATGCGCTACAAATCCTGTCTTTCTCAAAATG AACATTGGTTAGTCCTAAAATGTGCTCTCCCTTTAGAAGCTTATGAAGATTCTTCATCCCAAC ATGGGCTAGTCGGCGGTGCCAGAGCCAACCCATGTTAGTCTTAGCAATTAAACATGTGTCG AGCTCAGCTCTATCAAAATCTACTAAGTATAGCTGACCTTCTAACACACCCTTAAATGCTATT GAATCATCACTTCTTCTAAAGACAGTGACACCTATATCAGTAAAAAGACAGTTGTAGCCCATT TGACATAATTGGGAAACAGAAAGCAAGTTGTAATCTAATGAATCAACAAGAAAAACATTGGAA ATGGAATGGTCAGGTGAAATAGCAATTTTACCCAAACCTTTGACCAAACCTTGATTTCCATCC CCGAATGTGATAGCTCGTTGGGGATCCTGGTTTTTCTCATATGAGGAGAACATCCTTTTCTC CCCGGTCATGTGGTTTGTGCATCCGCTGTCGAG
REINA	TCGTAGTGGAAACCGACGCAAG	CCTCATAGCTGTTAGTCACCTC	Retrotransposon	TCGTAGTGGAAACCGACGCAAGTGGATCGGGTATTGGGGCAGTTTTGCAGCAAAAAGGCCA TCCTATAGCTTACATCAGTAAAGCTTTAGGTCCTAAGAACTTGGGATTGTCCACCTACGAAAA GGAATATCTAGCCATTTTATTTGCAGTTGATCACTGGCGCCCTTACCTCCAACATGGAGAATT TTTTTATCAAGACTGATCAGCAAAGCCTAACCCATCTAGAAGACCAAAAGTTGTCGACTATCT GGCAACAAAAAGCCATAACCAAGCTATTGGGTCTCCAATTCCGCATTATATATAAGAAAGGG GTGGAAAACCGTGTGGCAGATGCTCTTTCTAGACGGCCAAACCTATGCACTGATGGTTCCA AACTGCAACTCAACGTGTTAGCAACTTCGACAGTGGTACCAAGTTGGCTCACCGAGGTGAC TAACAGCTATGAG
RSI_Rough Sheath1	M13_For	M13_Rev	Control (positive)	Gift from G Marcos
SDG101	AGCCCTATAAGGAATATATTGG	ACTGCATGACAGATAATTTGTGC	Chromatin	AW066331
SDG102	CGTTGCCTCCACTCCACTAC	GCGTTTGACATGATTCCAC	Chromatin	BE345442
SDG103	TTGCTCCCAGGCCATGGAAG	TCACAGTTCACACCTTGGACACG	Chromatin	TTGCTCCCAGGCCATGGAAGATGTGCGCTTCACTATAGATGCATCTGTGTATGGAAACATCG GAAGGTTCATCAACCACAGCTGTTCACCAAATCTTCAAGCACAAAATGTTCTCCGGGACCAT GGTGACAAAAGGATGCCACATATTATGTTCTTTGCTGCAGAGACAATTCCACCACTACAGGA ACTGACTTATGACTATAATAACTCTGAAATAGATCGTGTCCAAGGTG
SDG104	GTATTACGTTTAGAATCTGC	CAAATTCACTTGAACTTATTGC	Chromatin	GTATTACGTTTAGAATCTGCAGTTTTTTTAAGCATAACTGAACTTGAACTTGTTTCATTTTGCC TTCATCTAGATCATGTTGTGACAAAATGTAAAATAAGACATCAGTGTTTTTAGTTGCCAATGG

ID	Forward Primer (5'-3')	Reverse Primer (5'-3')	GeneClass	Sequence /GenBank ID
SDG105	TAAGAGCCTTGCTCTTTGTGG	ACAAGTTTACACAAACCGGTGC	Chromatin	TAAGAGCCTTGCTCTTTGTGGTTAATCATGCAAACCACTTTCAGCTAGATCCTTTCTTAGTCT CCGGATGATATGAGTATGACTGATGATTCCAATTGCTGCCATGTGTACTGAAGCCACCAAGT TGTTGTCAAACTGTAGAATTTCTGTTGCCCTCGCACGGCATGGTGCCCGTTCTCCCACCATT TGTTCAAATATTTATGGGGGTGTGATTGATGTATTTCTTCTCGCTAATCTTGTCCACCAGCTC GCACCGCGTTTGTGTAAACTTGTTATGATCCCTGTTTGCTCTGGACGCCTGGTTGTTTCACC GTCACTCCTTGGATCAGAGTCATGTAAATTATCCTGGTTTAGG
SDG106	TGTAGATCAAGATACTGCAGG	CAGTGTCACATGAACAGCTG	Chromatin	TGTAGATCAAGATACTGCAGGCGGTACCTCAATTAGTTTCCGGTGGGCGTAACATGTCTATA TATAACTTATATATCTATATGTATATCTTATATATTTACAAATTACAATTGCGATTCATGCCTGC TGTATATGTAGGAACTTGCATCTAGAACTGCCTGTTTATCTTTTGTGAAGTCAAATGAGTTAT CAGCTGTTCATGTGACACTG
SDG107	AGTGTTGACTGAAAGCAGCG	TGGCTTAGGGTTTAGGAGTAGC	Chromatin	AGTGTTGACTGAAAGCAGCGGTGCATGGTTGGTTTGGGAAAGCGAACGGACTAATCAGTTT TAGAATGCTTTGTTAACCTAGGTAGCTATGCATTCATAAAGTGCCCGCTGTGCTTGACTGCA GAACGTGAAATTCAATTAGACAGACTGAAATGCAATGTACGTACGCTACTCCTAAACCCTAA GCCA
SDG108	TGAGAAGAATGATGCTGAGG	GGCCTTCAGATCACTAAGTACC	Chromatin	TGAGAAGAATGATGCTGAGGTTCAGGTAAAGGGAGTACTAAATTCAAGATTACAGCAAATAG TGCATACACTTGACAAGGTGAGATATGTTATGAGATGCATATTTGGAGACCCAAAGAATGCT CCTCCTCCTCTTGTAAGGTTGTCTGGGAAAAGTCTGGTATCTGCCATCTGGAAAGGGGACA GTTCGATAGTTGCTGAACTCATTCAGTCGATGGAACCTCATGTTGAAGAAGAGGTGCTGCGT GATGAATTGCGAACTCTTTCATGCACTTACAAGTGCCGTCATGATGCTGCTGCAGATTTGAT TCATTTGTATGCTTATACAAAGTGCTTCTTCCGAGTCCGGGATTATAAGACAGTAAAATCTCC ACCAGTTCATATCAGTCCGCT
SDG109	GATTGTAGGCAGATAAGATGG	ATTCCTATGGGAAAATTCAC	Chromatin	GATTGTAGGCAGATAAGATGGTATTTGGTTGTTTCTCCCTTTTGGAGATTTTAGTCTGCAGGA TTTTTTTTTCCAGGCTCACACCTTGGCTACCTGATCACAGGGAGTTTAGCTTATAGATAGCAG GGTAACAGTCACCACCCCCCATCTGTACATTCCCCAGTCCCTACATACCTTGCGGAGATGTA TCTGGTTAGTAAATGATAGATTTGTCATTGATTTAATGTAGTGAATTTTCCCATAGGAAT
SDG111	TGATGGGTTAAGATCCTTGC	AAGAAATACATCAATCACAC	Chromatin	AW216196
SDG112	T7	pAD-GAL	Chromatin	AW282247
SDG113	GGTCTCCTAGAGCAAAGAAC	GATTGACATTAATTAATTTAATGG	Chromatin	GGTCTCCTAGAGCAAAGAACTGCCTTTGTGAATCCTCAAATTGTCGGGGATTCTTTATCTGA CCGTGGAAAATGCATATTCTATCTATCAACATTCAACCCTACCAGCTAGATCCTTTTTGGTGG CACCTGAAGGAAAAGTAGGACTGTTCGAACAGCTTGCATCACTCTGTCGTTGTGAACCATTA AATTAATTAATGTCAATC
SDG115	CACCAAGGGAAGAGAGGGATC	CAAGGACCATTTCTCCTTCTTG	Chromatin	CACCAAGGGAAGAGAGGGATCCAGAATCTTTTGTACTTTCAAAGAGCGTTTGTGTTATTTAC AGGTTCGTGTGTCTAGTGCTTTCAAGTTTCTGTAGGACCTTCTGGTTTCTGTGCTAGTTTGAT ATTATGATATATTTGTCTTCCTTTTTTCAGAAAACAGAGCATGGTAGAGTATGCTTTGGTAGAT CAGGGATTCATAGGTGGGGTCTTTTTGCTCGGAGAGACATTCAAGAAGGAGAAATGGTCCT TG
SDG116	GAACTGATGCCATCACCTGG	ATGTCCAGCCAAACCAAATC	Chromatin	GAACTGATGCCATCACCTGGGATGTCTTACAAATGAGTTTGATTCTATTTCTGACTTCGCAAT AACATTGGCTTCATGAGGAATTGCCTGGCCCATTCTTTTAGCCATACTGTAATGTGTGCAAC AAACAAAAAAACAAGCTGTCAGTTTTTGGTGAAGAACTGGTCATGTTTGTAATGTGATTTGGT TTGGCTGGACA
SDG117	GCAGGAATACACTAACTGTTC	GAATCAGTTTCATTTCAGGCTGC	Chromatin	GCAGGAATACACTAACTGTTCAATGCTCGGTTGCAGAACTCCAGAGATTTCAACAGTTGTGT AGGCTTGCTGAGCGATTTAGCATTGCAATTTTCCTACAGCTTGAGTGCAGGGCTAGGATCTC ATTGTTAATGCTACAAATAGTGATTGTGATATAGGAATACAGGCAAGCATTAGGCAATAACAT AGTGTCGATTTTAGTCATGGTACCATCATGTTTAACTGATTATTTGTTAGAAGCAGCCTGAAA TGAAACTGATTC

ID	Forward Primer (5'-3')	Reverse Primer (5'-3')	GeneClass	Sequence /GenBank ID
SDG118	TTGGGATCAGTCGCAGTATG	AGCTTTGATTATAGGTGAGCAG	Chromatin	TTGGGATCAGTCGCAGTATGCCGATGAGTATGATCTGAACCAGTTCCTGGCGAAGCTGAAC GGCTGAGCCACCCCCGGCTGGCTCGTGTACAGTGTATAAAGCAATTTGGGGCAGGTTAATT TTGTTGTTGCTGTTGTTTGATAGAGGAAGGAAAAGATTAGATGCTAACAGATATAGGTGGTG GTTGGGGTTACGATCTCAATTTTTTTGGTTTCCATCTTCCTCAATTTTTGTTTTGAGGAGTTGT ATATACTCTGCTCACCTATAATCAAAGCT
SDG119	GAGGGGCTGAAGTCAGTGG	AGCAATGTTTAACCTTCTCCTC	Chromatin	GAGGGGCTGAAGTCAGTGGTTGGCCTAGGATCTGCTACTTCGTCTTCTGAGACCATGGAAG GCTTCACAATAGATGCATCTGAGTGTGGCAATGTGGGAAGATTTATCAACCATAGTTGTTCA CCAAACCTCTATGCCCAAAATGTTCTCTGGGACCATGATGATATGAGGATGCCACATGTTAT GTTTTTTGCTGTTGAGAATATCCCGCCACTACAGGAGCTGACCTACCACTATAACTATAAGAT AGGGGAGGTCTACATAAATGGCGAGGAGAAGGTTAAACATTGC
SDG120	GCCATGTGTACTGTGAAGACAG	GCAACATGATCCTTAACCAGG	Chromatin	GCCATGTGTACTGTGAAGACAGCAAGTTGTTGTCAAACTGTAGAATTTCTGTTGCCCTGCAC GGCATGGTGGCCGTTCTCCCACCATTTGTTCAAATATTTATGGGGGTGTGGATTGATGTATT TCTTCTGCTAATCTTGTCCACCAGCTGCACCGGTCTGTGTAAACTTGTTATGATCCCTGTTTG CTCTGGACGCCTGGTTGTTTAACCGTGACTCCTTGGATCTGGTCATGTAAATTATCCTGGTT TAGGATTATGTTGCTTTTGACTCGA
SDG122	CTTGCGTGCGTCATTCTGGAG	GATCTCTCATTGCGTGCCTC	Chromatin	CTTGCGTGCGTCATTCTGGAGTGCATACTCCGCGTTGTATGATTCTGATAAGCTGATGAAGA AATGGGGAAGACGAATTCCAAGCNAGGCCCTTGTAGCAGAGAGTGTTGCTGATGCAGTCGG CGGAAATGAGTGCGTGCTGAGAGCAACGCTGAGGGGTTCCAGGGACGGCAATGGCTATGG CAATCGGCTAGAGGTGGAGGACAAGGTGGTGAGGATTGGGAGGGCAACCTATGGCAAGTT GGTGAAGAGGCACGCAATGAGAGATC
SDG123	GTACCATCAGAATCTTAAGTTGC	TCCGGTATTCAAATGATCCGAG	Chromatin	GTACCATCAGAATCTTAAGTTGCAATCTTCTGTACACTAGTTACAGCCACGCCAATTGAGTTG TGGTTAAGCATTACTTTGCCTTTTGCACTTCCCTACCACGACATGCTAGAAATGTTGTGGTGT AAATGTTGTAGCCGTTTTTGTTTAGCTTCTTGTTCAGTCCTTAGTGTCAGAAATAAAATTTCGA TAATACTCGGATCATTTGAATACCGGA
SDG123	GTACCATCAGAATCTTAAGTTGC	TCCGGTATTCAAATGATCCGAG	Chromatin	GTACCATCAGAATCTTAAGTTGCAATCTTCTGTACACTAGTTACAGCCACGCCAATTGAGTTG TGGTTAAGCATTACTTTGCCTTTTGCACTTCCCTACCACGACATGCTAGAAATGTTGTGGTGT AAATGTTGTAGCCGTTTTTGTTTAGCTTCTTGTTCAGTCCTTAGTGTCAGAAATAAAATTTCGA TAATACTCGGATCATTTGAATACCGGA
SDG124	CCCCAATTTATTTGAACTCGTC	ATGGCTTGAGGAACTAGAGAC	Chromatin	CCCCAATTTATTTGAACTCGTCATTGTAACTCGTATGTGCTTGAATTCTCCATGGCATCTGGT CCTGCCATCCGTAGAGTTAGGTCCCGTTTGTTTTGAGGAACTAAAAATTAATCCCTCTATTTT AGTCACATTGAGTCTTAGATTGTTAAACGGCGGGACTAAAACAAAAGACTAAACTATTTGTCT CTAGTACCTCAAGCCATGACTAAAAGGGAATAAATCATATAAATTTTATTTTTATCCTTCCTTT AAAAAAA
SDG125	TTCAGTCCTAGTTCATTGGTAG	ACAAACAAGGCTATAGTCGCAG	Chromatin	TTCAGTCCTAGTTCATTGGTAGATAAACATGCTATATACTATCCATTAGTAAATAAACTCTCAT TCATCGAGTTGGAGAATAAATGCGTATAAACATATGTGGACCTCAGGTCGGGAAGGTGGCA ACCTTGTTAGTTTGAGCACCAACAGGTTCTCAAACTTGAGTGGCTATTGCTAGAGTATCAAAT AATGGCTGCGACTATAGCCTTGTTTGTATATTTTCTTGGTGAGATGAAATAATTTGTCAAATG TACACT
SDG126	CATCAAGTTGGGGAATAAGTG	CCAGCAACAATGTGTACATTTG	Chromatin	CATCAAGTTGGGGAATAAGTGCATATAAATATTTTTTACCTTAGGTCGGAAGGTGACAAACTG ACAACCTTGTTAGGTTGAGTAAAGTTAGTGCTTACAGGTTTTCAAACTGGTGTGGCTATTGCT AGAGTATCAAATAGCTGAGATGGCTACAACTATAGCCTTATTTGTATTTTTTTTTGGCAAGAG GGATGGAATTTATTTGCCAAATGTACACATTGTTGCTGG
SDG129	TCAGGCCGTGTCCACCATTG	AGCATTGTCACCCCTAGGTG	Chromatin	TCAGGCCGTGTCCACCATTGCAGCTCCAGCTCCACAGCTGCCAAAATTTCGTGTGATGTGG GGACCTGTTTGAATCGGCTTTGGCTAGTAACTAGTTGGAGCGGAAAGTTAATCTATATTAGC TGGCCTTCTTGTCTGGACAGATTAGAGTGGCTAACTAAAATGTCGATTCTTGCATAACAGTC AGGGCAGATGAAGCAATCTAGATATTCATCAGCACCTAGGGGTGACAATGCT

ID	Forward Primer (5'-3')	Reverse Primer (5'-3')	GeneClass	Sequence /GenBank ID
SGB101	CGGATGAAGCATATCTACAG	CAATAACAATAGATGACAAGG	Chromatin	CGGATGAAGCATATCTACAGAGAGGAGTCACCAAGCCCATTGGACTTCCGATACGCAGCTT TTTCTAATAGAATGGGAAACGAATTGTACTGTAATTGTTTGTGTACTTTACTTTGGCTCAATC GTCCGACCGTATCTTTGTAATCAGTTAGAATGGCAAGCAAATGTC
SGB102	CTAACATGCCAGCGATAAGG	ACAACAATACACTCACTACCAG	Chromatin	CTAACATGCCAGCGATAAGGATGTATGAAAAGCTTGGTTATGTGGTTTATCGAAGGGTGCTC CGGTACTACTCAGGGGAAGAAGATGGCCTCGATATGAGAAAAGCATTATCACAAGATGTTGA GAAGAAGTCCATCATACCACTCAAGAGGCCAATTACGCCTGATGAACTTGAATATGATTGAT GTTAGCTACCTGGTGCTTACGTTGGATGCTCCGGTGGGGAATTTGACAGGCTTTGAAAATCT GGTAGTGAGTGTATTGTTGT
SGB103	CACCAAGCCAATTGTACTCATC	TCAAACTTGCATTATATATTCCG	Chromatin	CACCAAGCCAATTGTACTCATCTTTCCTCTCACTGGGTATACAATATAGTAGTTGTAGCGGTC TTTCCGTAGTAGAACGGAGAAAGACCTGTGCAGTGAGTATTTGTGTACTTGGCCGATGGCTC GAATCGTCTGACCTAATCAGCGAGAATGGCAACCAAATATGCTTCGTACTTGAAACCTGTGT TGTTTTTAGACATGTGCGAATTTTCACAGTTCATTTGATTGGTGATGCCTGACAATTTGGTCA GATCTGATACCGCGGAATATATAATGCAAGTTTGA
SP1_B-cell receptor	M13_For	M13_Rev	Control (spike)	AF126021
SP2_Myosin Heavy Chain	M13_For	M13_Rev	Control (spike)	X13988
SP3_Myosin regular light chain 2	M13_For	M13_Rev	Control (spike)	M21812
SP4_Insulin -like growth factor	M13_For	M13_Rev	Control (spike)	X07868
SPM	GGATGGATTCGTCTGGTAGG	AATCCCCAACTCTGTGAGG	Transposon	GGATGGATTCGTCTGGTAGGAGATCCAGATCACGAAGGTCGAGGGGTTCCAGTGGTGCTC CTAACATGTTTGAGGTACTTGTTTTTTTCTTAGTTTAGTTACTGGGGCCAGCTAGCAATCTAA AATTGTTTACATTTCAGGGCACCACAACAAGCAGAAGCAGGCAAGAGCAGTTGCTTGCCTCC CTAGAGCAGATGCGGGGTAGCAGCGGACCCTCCAATACAGAAGGGACTACTTCGCGTGCA GCTGATTTAGTGGCACCTACTATGGCACCTACTGCGGAGGCAGCTGTGGACGCGGAGGCA GCTGTGGACGCGGAGGCTGAGGAAGCTGCCGCAGAACTGGACGATGGAGAGGAGACTTC GGGAGCAGACGCTTCTACAGAGGAGGCTGCCACACAGGCTCCGCCACGGCGTGCCATCAG GTATTCTTATTGTTGAGTTGAAGTTTTATTGTCAAGTTGAAGTTTTAAGTAAATGATTTTCTTT GTTAAATAACAAACTTTCATTTGTAGGTATCGTCGGTCCCTCACATTGAAGCCTTCTAAACCC TTCGATCAACGGAGGGTTATTGAACCGAAAGGCACCAGGGCGTGGAAGGAGGTTAGTTGG GACGGAACAGGGCATCGCACACCAATCCTCACAGAG
STONOR	AATCGACACTTTGAGAGG	ATTCTACTGAGGAGTATAGG	Retrotransposon	AATCGACACTTTGAGAGGAACCGATCCCCTAGGCGCTCGGAGTGACTGGAATCGCTGTGAC TATGTCTCAGATCCCCGTTATCCCATCGGGATCTGAGATTCCCCCTGCAGGCAACATTGGAG GAGCTGATCCACCGCCGATCAGCGACGGGAACCAATCTCCTCGACACCAGCAGGGAGATC GTCAAGAAGGAGGAGGCGGGTCTTCGCTGCAAGACACCGCCGCAATGATGGCTAGGAACA ACAACGACGGTATGTTAACCTATACTCCTCAGTAGAA

ID	Forward Primer (5'-3')	Reverse Primer (5'-3')	GeneClass	Sequence /GenBank ID
TEKAY	GTGGGTAAAGCTGTACCACCTC	CTCCATCAAAGTCTCCAGCATC	Retrotransposon	GTGGGTAAAGCTGTACCACCTCTGCAGAGTGTAAAACTGATATATCAGTCGTGCTCCCGGTT AAGAGCAGCCTGGACTCCTCACATTATAATTGAACTTAAAGATGGACAATAAATCGAGATCC GATGATCTGCCAATGGTTTGCTTAAGTGGTTTCACTTAAGTGTTTTTGATATACTCATATTCCT TTGTTTAAATGGGATACTTGGGATTCACACTTATTAGTAAGACATGTGTTGTTAATAAAATGTT GACCAACTAAAATGCTTACTGCTCAACCTTAGCCTCACCTTGTTATACCTGCATTAATTCCCC CCCCCCACTTGCTGAGCCCCGACCATAAGTGAGCTCACCCTTGCTAATATTTTGATCAGAGG GTGATGCTGGAGACTTTGATGGA
TG1_CRY1aC	M13_For	M13_Rev	Control (transgene)	U89872
TG2_Bar Gene	M13_For	M13_Rev	Control (transgene)	X17220
TG3_Green Fluorescent Protein	M13_For	M13_Rev	Control (transgene)	AF078810
TG4_Globin	M13_For	M13_Rev	Control (transgene)	NM_000518
TG5_GUS	M13_For	M13_Rev	Control (transgene)	A00196
TG6_Hygromycin B phosphotransferase	M13_For	M13_Rev	Control (transgene)	K01193
TG7_Luciferase	M13_For	M13_Rev	Control (transgene)	X65316
VC1_anthocyaninless1	T7	T3	Control (positive)	VC1
VC11_b1 colored plant	T7	T3	Control (positive)	VC11
VC12_C2 DNA	T7	T3	Control (positive)	VC12
VC14_MuDR	T7	T3	Control (positive)	VC14
VC15_MudrA	T7	T3	Control (positive)	VC15
VC16_MuRA	T7	T3	Control (positive)	VC16
VC17_MuRA	T7	T3	Control (positive)	VC17
VC18_Ubiquitin	T7	T3	Control (positive)	VC18

ID	Forward Primer (5'-3')	Reverse Primer (5'-3')	GeneClass	Sequence /GenBank ID
VC19_MuRA	T7	T3	Control (positive)	VC19
VC2_anthocyaninless1	T7	T3	Control (positive)	VC2
VC20_coloured aleurone1	T7	T3	Control (positive)	VC20
VC3_anthocyaninless1	T7	T3	Control (positive)	VC3
VC4_anthocyaninless1	T7	T3	Control (positive)	VC4
VC6_Bronze1	T7	T3	Control (positive)	VC6
VC7_colourless2	T7	T3	Control (positive)	VC7
VC9_Ubiquitin	T7	T3	Control (positive)	VC9
XILON	CCAAGAGGCCTTTGAAGC	GATTTCGCGCTCCAAG	Retrotransposon	CCAAGAGGCCTTTGAAGCGCTGAAAGAGAAGTTGACAACAGCGCCTGTCCTAGTCTTGCCT GATGTGCACAAGCCCTTCTCGGTGTATTGTGATGCTTGTTACACAGGTTTGGGATGTGTGTT GATGCAAGAGGGAAGAGTTGTGGCTTACTCGTCCCGACAGCTGAAGGTTCATGAGAAGAAC TACCCAATCCATGATCTAGAGTTGGCAGCAGTGGTTCACGCACTGAAGACATGGAGGCACT ATCTGTATGGACAGAAATGCGATGTTTACACAGATCATAAGAGTCTGAAGTACATATTCACTC AGTCAGAATTGAACATGAGGCAGCGAAGATGGTTAGAATTGATCAAAGACTATGAGTTGGAG ATTCATTACCATCCAGGCAAAGCAAACGTAGTGGCAGATGCTTTGAGCAGAAAGAGTCAAGT CAATCTGATGGTTGCTCGCCCGATGCCTTATGAGTTGGCCAAGGAGTTTGACAGGTTGAGT CTCGGATTTCTGAACAATTCGCGAGGAGTCACAGTTGAATTGGAACCTACCTTGGAGCGCG AAAT
ZEON	CCCATGTACGACGGTCATACAG	CTCCAAAGTTTGTGGTGGCTTC	Retrotransposon	CCCATGTACGACGGTCATACAGACCCGAAGCAATTCTTGATGAGCTATGAAGCAACCATATC TTCGTATGGTGGCAATACTGCAGTCATGGCAAAGTCTTTTGTCATGGCCGTCAAGAGTGTCG CTCAGACCTGGTACTCCTCTCTCCGGCCAGGGACAATCACCTCTTGGCAGAAGCTGAAGGA TATGTTGATAACCAGCTTCCAAGGGTTTCAGACGAAGCCAGTTACTGCTCAAGCTTTATTCC AGTGTACTCAGGACCACGAAGAATACCTCCAGGCGTATGTCCGAAGGTTTCTGCGTCTGAG GGCACAAGCGCCAACAGTGCCCAATGAAATTGTCATTGAGGCCATGATTAAGGGGCTTCGG TCGGGACATTCAGCGCAATACTTTGCCAGGAAGCCACCACAAACTTTGGA