Epigenetic regulation and inheritance of autonomous seed development in apomictic Hieracium

Julio Carlyle M Rodrigues

Thesis submitted for the degree of

Doctor of Philosophy

University of Adelaide,
School of Agricultural Food and Wine
Discipline of Wine & Horticulture

In collaboration with
CSIRO Plant Industry, Adelaide

December, 2007
Table of Contents

Abstract

Declaration

Abbreviations

Acknowledgements

Chapter 1. Introduction...........................................................................................................1

Chromatin modifications and gene expression.................................................................7

Epigenetics and plant development.................................................................................. 12

DNA methylation and plant development.........................................................................12
  DNA methyltransferases.................................................................................................12
  DNA methylation and genomic imprinting.....................................................................14
  DNA methylation and transposon silencing..................................................................16

Histone modifications and plant development............................................................... 17
  Histone acetylation........................................................................................................17
  Histone methylation.....................................................................................................18

FIS-class genes and seed development.............................................................................20
  Function of the FIS complex.........................................................................................20
  Imprinting mechanisms of FIS genes..........................................................................24
  FIS-class genes and DNA methylation.........................................................................28
  Diversification and redundancy of PcG complexes in plants.....................................29

Developmental, genetic and molecular aspects of apomictic development............. 32
  Apomixis is a dominant trait in grasses....................................................................32
  Multiple loci control apomixis in the Asteraceae family..........................................33
  Epigenetic aspects of apomictic development...........................................................34
Isolation of genes differentially expressed in apomictic plants.................36
The FIS complex and autonomous development........................................39

Context of thesis and project aims..........................................................42

Chapter 2. Hieracium Fertilization Independent Endosperm (HFIE) is essential for apomictic and sexual seed development..........................................44

Acknowledgements and clarification of contributions.................................45
Abstract......................................................................................................46
Introduction.................................................................................................47

Results

FIS complex genes are expressed in ovaries of sexual and apomictic Hieracium..........................................................50
HFIE inhibits central cell proliferation in the Arabidopsis fie-2 mutant........52
HFIE gene expression in Hieracium...........................................................53
Down-regulation of HFIE in the embryo sac of sexual Hieracium does not lead to autonomous seed initiation..................................................53
Down-regulation of HFIE in ovaries inhibits autonomous seed initiation and leads to early seed abortion in apomictic Hieracium.........................56
Expression of HFIE in vegetative and inflorescence tissues is essential for embryo sac formation and autonomous seed initiation.......................61
HFIE does not interact with FIS complex proteins in yeast 2-hybrid assays..62
AtFIE and HFIE proteins differ in surface protein loops and charge which may alter protein function…………………………………………………………………………………64

Discussion

HFIE is essential for viable seed formation in sexual and apomictic plants…70
Parental genome ratios, PcG gene function and seed viability in sexual and apomictic plants…………………………………………………………………72
Models for seed activation in sexual and apomictic plants……………………74

Materials and Methods…………………………………………………………………75

Chapter 3. Segregation of autonomous embryo and endosperm components in Apomictic Hieracium………………………………………………………………………84

Acknowledgements and clarification of contributions……………………….85
Abstract………………………………………………………………………………………86
Introduction…………………………………………………………………………………………87

Results

Mutant 115 has deletions in LOA and LOP……………………………………...89
Cytological analysis of loa/lop mutant 115………………………………………90
lop mutant 179 only produces viable progeny when cross-pollinated……….92
Cytology of lop mutant 179 confirms origin of viable progeny…………………93
Seed viability is affected in progeny of loa mutant 134…………………………93
Ploidy analysis of mutant 134 progeny confirms reproductive origin………..94
Mutant 134 progeny form irregular embryos……………………………………100

Discussion

Function of LOA might explain segregation distortion…………………………104
LOA contains a suppressor of megasporogenesis..........................105

Deletions in LOA and LOP restore the sexual pathway in apomictic

*Hieracium caespitosum*.................................................................107

Autonomous embryo and endosperm are closely linked in *Hieracium*.......107

**Materials and Methods**..............................................................110

**Chapter 4.** Characterisation of the MULTICOPY SUPPRESSOR OF IRA1

(MSI1) in apomictic and sexual *Hieracium*........................................112

**Abstract**......................................................................................113

**Introduction**................................................................................114

**Results**

Isolation and sequence analysis of *Hieracium* MSI1.........................115

*HMSI1* expression analysis.............................................................120

*RBR* isolation and expression in *Hieracium*....................................123

HMSI1-HRBR interaction is conserved in *Hieracium*.........................129

*HMSI1* expression is altered in a *lop* mutant 179............................129

**Discussion**

*HMSI1* expression in conserved in apomictic and sexual plants.........135

Unspliced *HRBR* transcript in embryo sacs could indicate an early
Gametophytic function.....................................................................135

HMSI1/HRBR interaction is conserved in apomictic and sexual

*Hieracium*.......................................................................................137
HMSI1 and HFIE might have similar functions during autonomous seed development in Hieracium

Materials and Methods

Chapter 5. Expression analysis of two classes of DNA methyltransferases in Hieracium apomictic and sexual ovaries

Introduction

Results

Isolation and sequence analysis of HMET

HMET organization in the genome of apomictic and sexual plants

Isolation of a partial cDNA for HCMT

Isolation of partial cDNA for HDME

HMET and HDME expression in apomictic and sexual ovary development

HCMT expression analysis

DNA methyltransferase expression analysis in apomictic mutants

Analysis of methylated cytosines in ovaries of apomictic and sexual plants

Yeast-two hybrid interaction of HMET1 with HRB and AtMEA

Discussion

Apomictic D3 HMET gene structure is distinct from sexual P4 HMET

HMET and HDME expression are coordinated in Hieracium ovary development
HCMT expression is associated with late stage seeds derived from unreduced gametophytes.................................180

Materials and Methods..................................................................................182

Chapter 6. Discussion......................................................................................186

The FIS genes and apomixis hypothesis......................................................187
HFIE has a post-fertilization function in seed formation.........................188
HFIE and HMSI1 are unlikely to be LOP determinants..........................189
Autonomous embryo is linked to autonomous endosperm in Hieracium....190
HMSI1 function in Hieracium.................................................................192
Role of HCMT expression in late stage seed development derived from unreduced embryo sacs.................................194
Function of LOA....................................................................................196
Conclusion..............................................................................................197

Appendix 1. Supplementary Data Chapter 2..............................................199
Appendix 2. Primers List Chapter 5..........................................................201

References..................................................................................................202
Abstract

Apomixis is an intriguing and agronomically valuable asexual reproductive pathway resulting in seeds that give rise to plants that are identical in genotype to the female parent. Apomixis is absent in agriculturally important seed crops and our work has focused on the analysis of apomixis in the daisy-like genus *Hieracium* which contains sexual and apomictic species. Prior studies have shown that apomixis in *Hieracium* is controlled by two dominant loci. The LOA controls the avoidance of meiosis during female gametophyte (embryo sac) formation and the LOP locus is required for fertilization independent embryo formation during seed initiation. The genes conferring apomixis are unknown. In this study we focused on the events of autonomous seed initiation.

Cytological examination of apomictic mutants that have lost LOA or LOP and analysis of their progeny enabled us to characterize developmental aspects associated with the function of these loci. Upon removal of LOA meiosis occurs normally and LOP segregates with a 1:1 ratio in the progeny, characterizing maternal gametophytic control. We also show that autonomous embryo formation segregates with autonomous endosperm formation, suggesting that these two loci are closely linked. However, upon meiotic division, embryo lethal components arise and embryo development in apomeiosis mutants was generally defective and seed set was low. Similarly, upon removal of LOP, apomixis initiation occurs normally and unreduced embryo sacs can only form seeds if pollinated.

Autonomous seed initiation is actively repressed in the sexual model plant Arabidopsis by the action of a chromatin remodelling complex encoded by the FERTILIZATION INDEPENDENT SEED (FIS)-class genes. These genes are homologues of the Drosophila PcG complex that also repress gene expression throughout Drosophila development. Mutations in the FIS-class genes lead to elements of apomixis, such as autonomous endosperm, and in one
particular mutant, autonomous egg cell development. Given the similarity in apomictic and FIS-
class gene mutant phenotype we isolated three homologues from sexual and apomictic Hieracium
plants: **FERTILIZATION INDEPENDENT ENDOSPERM (FIE), MULTICOPY SUPPRESSOR
OF IRA1 (MSII) AND RETINOBLASTOMA (RBR)**. FIS-class genes from sexual and apomictic
*Hieracium* and examined their expression, interaction and function during seed initiation. The
isolated *Hieracium* FIS-class genes were highly conserved in sexual and apomictic plants in
terms of gene sequence and temporal and spatial expression pattern. Analysis of protein
interactions by yeast-two hybrid showed that the HFIE gene from sexual and apomictic plants
does not interact with other complex members in the same manner found in Arabidopsis. Protein
modelling uncovered structural differences between the Arabidopsis and *Hieracium* FIE proteins.
RNAi-mediated down-regulation of HFIE in sexual *Hieracium* did not lead to autonomous seed
initiation indicating HFIE was not part of a repressive complex. Down-regulation of HFIE in
sexual and apomictic plants revealed the gene was essential for embryo growth and viability.
Therefore, FIS-complex genes interact differently in Arabidopsis and *Hieracium* and have
different developmental roles.

In summary, the results presented here suggest that the FIS-genes are not mutated in
apomictic *Hieracium* plants, but they interact differently relative to the Arabidopsis counterparts
and play a fundamental role in embryogenesis. Thus, engineering autonomous seed into crops
will not depend on mutating these genes but rather in uncovering the molecular signal that
triggers apomictic development.
Declaration

This work contains no material that has been accepted for the award of any other degree or diploma in any university or any other tertiary institution. To the best of my knowledge and belief this thesis is original and contains no material previously written or published by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University library, being available for loan and photocopying.

Julio Carlyle Macedo Rodrigues

December, 2007

Abbreviations

35S – Cauliflower Mosaic Virus 35S promoter
A1 – Aposporous initial cell
At – Arabidopsis thaliana
bp – base pairs
BMM – butyl methyl methacrylate
CC – Central Cell
cDNA – complementary DNA
CMT – Chromomethylase
DAPI – 4’,6’-diamidino-2-phenylindole
DIG - digoxygenin
DNA – deoxyribonucleic acid
dNTP – 2’-deoxynucleotide triphosphates
EC – Egg Cell
ES – Embryo Sac
FAA – formaldehyde acetic acid
FIE – Fertilization Independent Endosperm
FIS – Fertilization Independent Seed
FPN – Fused Polar Nucleus
GUS – β-D-glucoronidase
H3K – Lysine residue of histone 3
HCl – hydrochloric acid
hp – hairpin
Kb – Kilobase pairs
LOA – Loss of Apomeiosis
LOP – Loss of Parthenogenesis
M - Molar
MEA - MEDEA
MET – DNA methyltransferase, MET1-class
MMC – Megasporocyte Mother Cell
MSI1 – Multicopy Suppressor of Ira1
mRNA – messenger RNA
ORF – Open Reading Frame
PBS – phosphate buffered saline
PcG – Polycomb Group
PCR – Polymerase Chain Reaction

qPCR – Quantitative PCR

RACE – Rapid Amplification of cDNA Ends

RBR – Retinoblastoma-related Protein

RNA – ribonucleic acid

RT-PCR – Reverse Transcription PCR

SSC – standard saline citrate

UB – ubiquitin

UTR – Untranslated Region

WT – Wild Type
CHAPTER 1

INTRODUCTION
Introduction

In angiosperms, the decision to reproduce is marked by a transition from vegetative to reproductive growth, leading to the formation of a flower, in which the gametes are formed. The multi-cellular male and female gametophytes of flowering plants are embedded in sporophytic tissue and are formed after meiotic reduction and subsequent mitosis. Sexual reproduction is marked by a double fertilisation event, when the male gametophyte pollinates the female flower, leading to the discharge of two male gametes into the female gametophyte.

The female gametophyte is formed within the ovule located in the carpel. The most common form of female gametophyte in Angiosperms (Polygonum-type) begins with the differentiation of an archesporial cell near the tip of the ovule primordium and it enlarges to form the megaspore mother cell (MMC) (Figure 1). The MMC undergoes meiotic division, resulting in the formation of four megaspores. Three of these degenerate and the nucleus of the selected megaspore divides mitotically and after three mitotic divisions, nuclear migration and cellularisation a mature 7-celled embryo sac (ES) is formed. During the double fertilisation event that is characteristic of sexual plant reproduction, one male sperm cell fuses to the homodiploid central cell nucleus, to give rise to the endosperm, while sperm cell fusion with the egg cell initiates embryo formation. Both endosperm and embryo develop in a coordinated way to form a mature seed.

Some plants form seed asexually by apomixis (Nogler 1984; Asker and Jerling 1992). In gametophytic apomixis, two major events characterise this mode of reproduction: (1) the avoidance of meiotic reduction during embryo sac formation and; (2) fertilisation independent embryo development (Koltunow and Grossniklaus
Apomixis results in seeds containing an embryo that is genetically identical to the female parent. Apomixis does not occur to the mutual exclusion of sexual reproduction as the capacity to form a percentage of seeds via sexual reproduction is retained in most apomicts. The avoidance of meiosis results in the formation of an unreduced embryo sac, and this can occur by two different pathways (Asker and Jerling 1992). In displospory, the MMC does not complete meiosis and enters the gametophytic program; whereas in apospory, a somatic nucelar cell enlarges and differentiates to form an aposporous initial (AI) cell which is the progenitor of the unreduced embryo sac (apospory) (Figure 1). In both cases, the egg cell develops without fertilization. Endosperm development can occur without fertilization (autonomous) and this is restricted to the Asteraceae family (Koltunow et al., 1998a). In the majority of apomicts, such as the grass species, fertilisation of the central cell (pseudogamy) is required for endosperm formation and viable seed development.

The production of many agronomic crops depends largely on sexually derived hybrids, whose beneficial traits including yield and hybrid vigour can be lost in successive generations as a result of recombination and segregation events that define sexual reproduction. Apomixis is largely absent in agricultural crops but harnessing apomixis in crops could be an important tool to fix genetically desired traits. This would allow the production of clonal seeds with embryos that are genetically identical to the maternal parent, enabling the maintenance of an elite genotype through seed (Bicknell and Koltunow 2004). An essential step towards the effective utilization of apomixis in crops is an understanding of the molecular mechanisms that control
Figure 1. Schematic representation of sexual and apomictic reproduction. Sexual reproduction is characterized by the production of a female gametophyte after meiotic reduction and a double fertilization event to produce embryo and endosperm. In gametophytic apomixis these two events of sexual reproduction are bypassed. Firstly, female gametophyte formation occurs without meiotic reduction. Secondly, the unreduced egg cell develops without fertilization to form an embryo genetically identical to the mother plant. Female gametophyte formation can occur through apospory, where a somatic cell (aposporous initial – AI) in the nucelus differentiates and initiates gametophytic development; or diplospory (see text for details). In most apomicts the central cell needs to be fertilised to produce endosperm (pseudogamy), whereas in other plants the central cell develops autonomously to produce the endosperm.
apomixis. Genetic studies show that apomixis is generally a dominant trait (Savidan 2000; Bicknell and Koltunow 2004) however, genes controlling apomictic development are unknown. Isolation of genes involved in apomixis has been approached in a number ways: comparative screening of differentially expressed genes in related sexual and natural apomictic populations; the analysis of induced mutants that have lost apomixis and the mutagenesis of sexual species to induce components of apomixis. In Arabidopsis, mutants that display a component of apomixis, the initiation of fertilisation-independent endosperm development were isolated. Such genes are related to the Drosophila chromatin modifying Polycomb group genes. However, their role during apomictic development is still unclear.

Recent advances in understanding the molecular relationship between sexual and apomictic reproduction were made in the model plant Hieracium. Hieracium belongs to the Asteraceae family of daisy-like plants. The developmental, cytological and genetic aspects of apomictic development have been extensively described (Koltunow et al., 1998a; Bicknell et al., 2000; Koltunow et al., 2000). Apomixis initiation is typically aposporic and the differentiation of somatic cells occurs after the differentiation of the megaspore mother cell (MMC). There is variation in the timing and penetrance in which apospory occurs depending on the species, however, apomixis still behaves as a dominant genetic trait. The embryo sac is formed after three rounds of mitotic divisions of the aposporous initial cell, resulting in 8 nuclei, similar to the Polygonum-type embryo sac structure of most sexually reproducing plants. As with the other apomictic species of the Asteraceae family, apomictic seed development is completely autonomous in Hieracium, as both the formation of the embryo and endosperm are independent of fertilization. Apomixis is facultative in Hieracium and low levels of sexuality still occur (Bicknell et al 2000). An elegant
marker study showed that the aposporous initial cell does not share the identity of the MMC but acquires gametophytic identity soon after mitotic divisions of AI nucleus occurs, thereby demonstrating that sexual and apomictic reproduction share molecular components in *Hieracium* (Tucker et al., 2003). Although a temporal and spatial shift in the expression of one of the markers was observed during early ovule initiation in apomictic *Hieracium*, the identity of the AI cell remains elusive (Tucker et al., 2003).

Given the demonstrated relationship between sexual and apomictic pathways, two models have been proposed to explain the molecular manifestation of apomixis. A genetic mutation model predicts that a key factor regulating sexual reproduction is mutated or has an altered function, thereby leading to the deregulation of the sexual pathway. The mutation is not fully penetrant, thus the sexual process can also occur. The epimutation model predicts that reversible changes in chromatin configuration might alter the timing of expression of key regulatory genes in both time and space, possibly triggering the apomictic pathway at different developmental time points or cell types (Koltunow and Grossniklaus 2003). These models are not mutually exclusive considering that a mutation in an epigenetic regulator could lead to altered heterochromatin formation. Some evidence supporting the view that apomixis is epigenetically regulated includes the observation that in most apomicts the developmental program is not tightly regulated and differences in the timing of initiation and variation in structures formed occurs in an individual plant in response to different environmental conditions and stresses. Such flexibilities are characteristic of an epigenetic process. Epimutations are also more frequent, can be reversible and transmitted to progeny, thus making it more likely to generate multiple changes in gene expression to allow a complex developmental trait such as apomixis to co-evolve.
and occur simultaneously with sexual reproduction in an individual plant (Koltunow and Grossniklaus 2003).

The following review will provide a brief explanation of epigenetic regulation, its mechanisms and the role of histone and DNA modifications in the formation of heterochromatin. Therein, an overview of the developmental processes in plants that have been shown to be controlled epigenetically will be given, with particular focus to the epigenetic regulation of seed development. This will then lead to the formulation of the hypothesis of the current thesis, to assess whether epigenetic regulators are involved in the control of apomictic development, in which specific objective will be outlined.

**Chromatin modifications and gene expression**

Plant and animal growth and development depends on the coordinated expression of specific subsets of the total complement of genes found in the genome. The molecular events that integrate developmental signals and cell differentiation can be set early in development, defining specific gene expression patterns that must be precisely remembered throughout the subsequent divisions of cells comprising a particular tissue. These set states of gene expression programs can be established and maintained by controlling chromatin configurations in specific regions of the DNA and thus the expression of the genetic information of the DNA sequence itself.

In eukaryotes, DNA is packed into the nucleus through an association with histones, to form chromatin. The basic organization unit of chromatin is a nucleosome, consisting of 146 bp of DNA wrapped around an octamer of histones. Histones are highly conserved basic proteins with a globular domain that interacts
with DNA and a flexible amino-terminal region that remains somewhat protruded from the nucleosome (Jenuwein and Allis 2001). Chromatin can exist in a more compact, condensed form (heterochromatin), inaccessible to transcription factors and RNA polymerase and thus transcriptionally silent; or in a less condensed state which is associated with transcriptional activity (euchromatin). The extent to which chromatin is condensed depends on DNA and histone modifications and is of great biological significance. Chromatin condensation can be targeted to specific regions in the genome, leading to the formation of heterochromatin in telomeres, centromeres and transposable elements, ensuring genomic stability; or can be targeted to specific genes, to control gene expression involved in different aspects of plant and animal development (Chan et al., 2005; Steimer et al., 2004; Reik et al., 2001). In this context, cell differentiation is largely determined by what transcription factors are available, the accessibility of chromatin to these transcription factors and how this interplay is controlled throughout a developmental process.

The condensation state of chromatin depends on chemical modifications that occur to both DNA and histones. The amino-terminal region of histones can be the target of several post-translational modifications, including phosphorylation, methylation, acetylation and ubiquitination. To date, the most described modifications that control chromatin condensation are histone methylation and acetylation (Jenuwein and Allis 2001). DNA can also be chemically modified by methylation of cytosine residues. Together, these modifications interact to reinstate epigenetic marks, adding another layer of regulation on top of the information encoded in DNA itself, known as the histone code or epigenome (Figure 2). According to the histone code hypothesis, histone modifications are interdependent and can be synergistic or antagonistic (Strahl and Allis 2000; Jenuwein and Allis 2001). Histone methylation is
usually associated with DNA methylation and deacetylated histones. This particular combination of modifications trigger chromatin condensation associated with gene silencing. On the other hand, histone acetylation and histone methylation co-exist to determine a more relaxed chromatin state, compatible with transcriptional activity. The apparent contradiction of histone methylation being associated with both chromatin states can be explained by the fact that different amino acid residues are methylated in each case. Typically, condensed chromatin is marked by methylation of lysines residues 9 and 27 on histone 3 (H3K9 and H3K27), whereas active chromatin is associated with H3K4 (Jenuwein and Allis 2001).

Specific enzymes carry out chromatin modifications. Histone methyltransferases transfer a methyl group to histones, whereas histones acetylases transfer an acetyl group to histones. Histone acetylation is counteracted by histones deacetylases, which remove acetyl units from histones. DNA modifications are made by cytosine DNA methyltransferases that transfer a methyl group to certain cytosine residues. DNA can also be demethylated by DNA glycosylases (Penterman et al., 2007). Although the sequence of chromatin modifications is important in any given promoter sequence, it can vary according to the signalling process and the target sequence. These specific enzymatic activities are often associated with chromatin remodelling complexes that have the ability to control gene expression and cell fate during development. These protein complexes may contain transcription factors that account for specificity of a chromatin modifying complex. For instance, proteins of the Polycomb Group (PcG) can form complexes that are capable of maintaining chromatin states established early in development, defining a cellular memory of transcriptionally active or inactive genes (Pirrota et al., 2003). In Drosophila, PcG complexes recognize and bind to upstream responsive elements of homeotic genes to
establish a predominantly repressive state and define segment identity (Rongrose and Paro, 2007). Trithorax complexes have an antagonistic action, preventing the establishment of repressed states on specific promoter sequences (Rongrose and Paro, 2007; Pirrota et al., 2003). The PcG complex comprised of the proteins EXTRA SEX COMBS (ESC) and ENHANCER OF ZESTE (E(z)) of Drosophila has histone methyltransferase activity conferred by the SET domain of E(z) that is only active in the context of the protein complex (Ringrose and Paro, 2007). This complex also interacts with a DNA methyltransferase and a histone deacetylase, suggesting that histone modifications are important to establish and/or maintain a repressed state (Viré et al., 2006; Simon and Tamkun 2002).

Small interfering RNA (siRNA) can also target chromatin modifications to regions of complementarity. Heterochromatin formation in centromeres and transposable elements is established and maintained by siRNA-mediated silencing (Schramke and Allshire 2003; Lippman et al., 2004). Furthermore, transcription factors can also be targeted by siRNA, leading to chromatin modifications in homologous regions and transcriptional repression that possibly have important functions in controlling plant and animal development (Steimer et al., 2004). Recent findings also point to a role of microRNAs in directing DNA methylation in specific regions to control the expression of genes involved in leaf morphology (Bao et al., 2004).
Figure 2. Schematic diagram showing the dynamic chromatin modifications that make up a histone code, or epigenome. DNA is wrapped around histones to form nucleosomes, which are the basic structural unit of chromatin. Chromatin modifications can reinstate condensed chromatin resulting in transcriptional silencing, or relax chromatin ultimately resulting in transcriptional activity. The enzymes responsible for histone and DNA modifications indicated can interact with different protein complexes to control gene expression.
Epigenetics and Plant Development

In plants, epigenetic regulation has been shown to be important for normal development. Epimutations, aberrant DNA methylation patterns and defects in chromatin remodelling genes all have been associated with different kinds of developmental abnormalities, such as homeotic transformations, sterility, transposon activation and defects in flowering response pathways (Finnegan et al., 1996; Ronemus et al., 1996; Soppe et al., 2000; Lippman et al., 2004).

Chromatin condensation states can also be controlled via the displacement from specific loci, this activity is mediated by protein complexes of the SWI/SNF family (Sarnowski et al., 2005). This review will focus on the biological roles of enzymes and protein complexes that promote histone and DNA modifications.

DNA methylation and plant development

DNA methyltransferases

The plant family of DNA methyltransferases has been separated into classes that differ in sequence, domain content and substrate specificity (Bestor 2000; Finnegan and Kovac 2000). Class I, the MET-class, which is closely related to the mammalian Dnmt1; chromodomain methyltransferases (chromomethylases - CMT) which are specific to plants and finally, domains rearranged methyltransferases (DRM) which are related to mammalian Dnmt3 (Bestor 2000; Finnegan and Kovac 2000). These classes are involved in methylation of cytosine residues in symmetrical
sites (CpG, CpNpG) and de novo (CpNpN) methylation, respectively (Henikoff and Comai 1998; Genger et al., 1999; Finnegan and Kovac 2000; Cao and Jacobsen 2002a,b). The different classes of DNA methyltransferases in both mammals and plants probably operate cooperatively and redundantly to regulate DNA methylation in different regions (Bestor 2000; Morel et al., 2000; Xiao et al., 2003). Previously established methylation marks can be copied to the daughter strand during DNA replication, and thus maintained indefinitely.

Methylated DNA is usually associated with chromatin condensation and transcriptional repression. A clear role for DNA methylation has been established in silencing repetitive DNA, transgenes and viral sequences, suggesting that it might have evolved as a mechanism to ensure genomic stability in larger genomes by controlling the activity of parasitic sequences (Matzke et al., 1999; Walsh and Bestor 1999; Ehrlich 2003). However, recent mapping of methylated regions in the Arabidopsis genome also has shown that endogenous genes with important developmental function are also targeted by DNA methylation (REF).

Several lines of evidence show that DNA methylation is important for normal plant development, and more specifically, plant reproduction. DNA methylation is involved in the control of genomic imprinting (allele-specific gene expression) an important aspect of seed development (Julien et al., 2006; Adams et al., 2000). Ectopic down-regulation of a class-I DNA methyltransferase leads to genome hypomethylation and alterations in leaf morphology, flowering time and defects in gametophyte and seed development (Finnegan et al., 1996; Ronemus et al., 1996). Furthermore, reduced DNA methylation caused by a loss-of-function mutation in a chromatin remodelling gene of the SWI/SNF family (DDM – decreased DNA methylation) also leads to pleiotropic developmental defects that are initially weak,
but become more severe in subsequent generations (Jeddeloh et al., 1998; Jeddeloh et al., 1999).

The altered flowering time phenotype in \textit{fwa} mutants has been shown to be due to locus-specific changes in DNA methylation patterns in flowering repressors, such as the early flowering locus \textit{FWA} (Soppe et al., 2000). Furthermore, hypomethylated plants also display altered flowering time and this has been shown to be correlated with a decrease in the expression of \textit{FLOWERING LOCUS C (FLC)} (Genger et al., 1999). These findings provide evidence that DNA methylation has a potential to control the expression of specific genes involved in plant development. However, direct evidence of the relevance of these changes to actual control of flowering in a wild type situation is still lacking.

\textit{DNA methylation and genomic imprinting}

It has been shown genetically that DNA methyltransferase activity is necessary during gamete formation during plant reproduction (Saze et al., 2003), but little is known of how this epigenetic mark is recruited to specific regions of the genome to mediate gene expression. Nevertheless, DNA methylation has been shown to be involved in other aspects of seed development, such as genomic imprinting, parental-specific gene expression. In mammals inheritance of aberrant DNA methylation marks in imprinted loci lead to abnormal development and imprinting marks is mainly maintained by the class-I DNA methyltransferase DNMT1 (Ferguson-Smith and Surani 2001). Similarly, DNA methylation is also involved in mediating parent-of-origin effects, in which a phenotype is only transmitted through the female gamete regardless of a wild type paternal allele. In \textit{Arabidopsis} distinct
phenotypes are obtained in reciprocal crosses between hypomethylated and wild type plants (Adams et al., 2000). These crosses phenocopy the effects of interploidy crosses (Scott et al., 1998). Furthermore, hypomethylated pollen is able to rescue maternal gametophytic mutations that affect seed development (Vielle-Calzada et al., 1999; Luo et al., 2000; Vielle-Calzada et al., 2000; Guitteny et al., 2004). The need for the correct balance of maternal and paternal genomes is probably due to the combination of male and female contributions for correct gene expression necessary for normal development. In some plants, a 2:1 maternal to paternal genome ratio is essential for normal endosperm development and if it is altered it affects embryo viability (Haig and Westoby, 1991; Adams et al., 2000). In plants, genomic imprinting seems to be restricted to the central cell of the female gametophyte, thus associated with endosperm development (Gehring et al., 2004). Activation of the maternally expressed genes MEA and FWA occurs during female gametogenesis by the activity of a DNA glycosylase, DEMETER (DME) (Choi et al., 2002; Xiao et al., 2003; Kinoshita et al., 2004). DME is activated in the central cell before fertilisation. After fertilisation it is down regulated, but both MEA and FWA expression continues even in the absence of DME. The absence of DME expression in stamens suggests that expression of MEA and FWA in the female gametophyte is largely due to the activation of DME, its absence in the male gametophyte would result in silencing of MEA and FWA (Gehring et al., 2004). This model suggests that DNA methylation is a default condition of imprinted genes, a pattern overcome by an epigenetic modification caused by DME (Xiao et al., 2003; Gehring et al., 2004). In fact, MET1 activity has been shown to be required to silence the expression of imprinted genes in the male gamete and vegetative tissues (Jullien et al., 2006).
Taken together, these results indicate that DNA methylation is differentially established during male and female gametogenesis, and when the two gametes combine during sexual reproduction, these differences interact to control gene expression programs for proper seed development. Therefore, DNA methylation is an epigenetic mark that mediates parent-of-origin effects and ensures that correct expression in each parental gamete is established. Alterations in this epigenetic program can have drastic effects on gametophyte development and progeny survival.

**DNA methylation and transposon silencing**

DNA methylation has also been shown to be involved in silencing the activity of transposable elements in both plants and animals. Due to a widespread distribution of transposons in eukaryotic genomes and their potential to promote genomic rearrangements if activated, this function of DNA methylation has been attributed to maintaining genomic stability. In plants, transposon silencing can be reversed under certain types of stress, such as genomic stress imposed after hybridisation between different genomes (Kashkush et al., 2003) or physiological stress such as reduced temperature, in which an association with DNA methylation was established (Hashida et al., 2003). Additionally, transposon sequences can also attract epigenetic modifications that silence gene expression when inserted in the proximity of genes (Lippman et al., 2004). In this scenario, transposons can be a source of adaptive variability in response to different types of stress, mediating alterations in gene expression through epigenetic mechanisms. Furthermore, transposable elements can function as regulatory elements for host gene expression. Recent findings in mouse oocytes have shown that retrotransposons are expressed abundantly and that this
expression is developmentally regulated, apparently due to transcription sites within
the retrotransposon sequence (Peaston et al., 2004). This leads to the expression of
chimeric transcripts with altered exon composition, possibly encoding proteins with
alternative functions (Peaston et al., 2004). However, the developmental significance
of this control has yet to be demonstrated.

Interestingly, a surge of transposon activity at the time of apomictic initiation
is observed in Hieracium, which is decreased in mutant plants that have lost the
capacity for apomixis (Bicknell, personal communication). It remains unclear whether
these differences are of functional significance for apomictic development or what
role transposon activity might have in regulating apomixis. In light of the information
described above, it is tempting to speculate that transposons might be inserted in key
epigenetic regulator(s) of sexual reproduction in apomictic plans, which would
deregulate expression of the sexual developmental program in space and time.
Alternatively, if a defect in the RNAi silencing machinery deregulates the expression
of genes controlling sexual reproduction in apomictic plants, then the transposon
activity observed would be mostly an effect rather than a cause of apomixis.

Histone Modifications and Plant Development

Histone acetylation

The most common site for histone acetylation are lysine residues and the
overall effect of acetylation on local nucleosome structure is probably to reduce
interaction to the negatively charged DNA molecule, making the DNA more
accessible to the transcriptional machinery (Strahl and Allis 2000). Conversely,
Histone deacetylation, promoted by histones deacetylases (HDs) would make chromatin more condensed (Strahl and Allis 2000). Histone deacetylation is linked to DNA methylation. In mammalian cells, methylated DNA is recognized by methyl DNA-binding proteins (MBP), which in turn interact with HD activity (Nan et al., 1998).

There are several classes of HDs in both plants and animals, suggesting divergence in function and/or substrate specificity (Pandey et al., 2002). In plants, functional analysis of a class I-type RPD3/HDA1 enzyme by an anti-sense strategy in Arabidopsis showed various developmental defects, including flower defects, male and female sterility (Tian and Chen 2001). This class is common to both plant and animals, suggesting that this family evolved as an important general regulatory mechanism in development. Regulation of flowering time is mediated by changes in histone acetylation of the flowering repressor FLC (He et al., 2003; Kim et al., 2004).

Plants have a unique family of histone deacetylases, HD2, and expression analysis suggests a role in seed development (Wu et al., 2000; Pandey et al., 2002; Lagace et al., 2003; Zhou et al., 2004). Consistent with the expression pattern, down-regulation of the HD2 family member HD2a by an anti-sense strategy leads to seed abortion (Wu et al., 2000). Furthermore, somatic and zygotic embryo development induces the expression of HD2a, Hd2b and HD2c and over-expression of HD2a leads to the repression of seed-specific genes (Lagace et al., 2003; Zhou et al., 2004). These data suggest that the HD2 family might have functionally evolved in plants to control gene expression in seed development.

**Histone methylation**
As for acetylation, the major sites for histone methylation are lysine residues. Histone methyltransferase activity has been largely associated with the presence with a conserved SET domain, named after the *Drosophila* PcG proteins SUPPRESSOR OF ZESTE3-9, ENHANCER OF ZESTE (E(Z)) and TRITHORAX (TRX), all of which mediate epigenetic processes during *Drosophila* development (Muller et al., 2002; Francis and Kingston 2001). In *Arabidopsis* there are 29 expressed SET domain-containing proteins divided into four classes according to their *Drosophila* counterparts, E(z), Trx, Ash1 and Suz 3-9, suggesting functional diversity and preference for certain lysine residues (Baumbusch et al., 2001).

As in other organisms, a close relationship between histone methylation and DNA methylation has been reported in plants. Mutations in *KRYPTONITE* (*KYP*), a gene of the Suz3-9 family, lead to a loss in DNA methylation at CpNpG sites, consistent with the activity of CHROMOMETHYLASE (CMT3) (Johnson et al., 2002). It was further shown that the chromodomain of CMT3 binds to histone 3 methylated at lysines 9 and 27, providing a mechanism by which histone methylation directs DNA methylation (Johnson et al., 2002; Jackson et al., 2004; Lindroth et al., 2004).

The large variety of histone methyltransferases in *Arabidopsis* suggests that they are involved in various aspects of plant development. *KYP* was originally identified as a suppressor of the clark kent alleles of *SUPERMAN*, a gene involved in defining flower homeotic gene expression boundaries (Sakai et al., 1995). These alleles are silenced by DNA methylation leading to the production of more stamens and unfused carpels (Jacobsen and Meyerowitz 1997). There are three E(Z) homologs in *Arabidopsis*, CURLY LEAF (CLF), SWINGER (SWN) AND MEDEA (MEA), all of which are involved in repressing gene expression of different processes of plant
development. CLF and SWN are involved in epigenetic gene regulation of flower development, whereas MEA is involved in repressing endosperm development without fertilisation (reviewed in Hsieh et al., 2003). Another PcG protein, the Su(z)12 homolog VERNALIZATION2, mediates FLC repression by promoting H3K27 methylation (Bastow et al., 2004; Sung and Amasino 2004).

Repression of gene expression is imposed at specific targets by promoting chromatin modifications such as histone methylation and possibly DNA methylation. The identification of MEA suggests that chromatin modification is an important molecular mechanism in the control of plant reproduction. In *Drosophila*, E(Z) forms a complex with EXTRA SEX COMBS (ESC), a WD-40 protein required for histone methylation activity. In *Arabidopsis*, FERTILISATION INDEPENDENT ENDOSPERM (FIE) is an homologue of ESC and *fie* mutants also share the endosperm phenotype of *mea* (Chaudhury et al., 1997; Luo et al., 1999; Ohad et al., 1999). Like their *Drosophila* and human homologues, MEA and FIE interact (Luo et al., 2000; Spillane et al., 2000), suggesting that this functional relationship between chromatin remodelling genes has been conserved and is of fundamental importance to maintain proper gene expression programs in seed development. Defects or untimely expression of these epigenetic regulators could lead to mis-expression of target genes that control seed development in response to fertilisation.

**FIS-class genes and seed development**

*Function of the FIS-complex*
The importance of chromatin modifications in the control of gene expression during plant reproduction has been shown by the isolation of several mutants that have abnormal seed development phenotypes. Mutations in a group of chromatin remodelling genes encoding homologues of the Polycomb Group-like proteins called \textit{FERTILISATION INDEPENDENT SEED (FIS)} in \textit{Arabidopsis} lead to the initiation of and endosperm development without fertilisation. If fertilisation occurs, there is an arrest in seed development characterised by an over-proliferated endosperm and an arrested heart-stage embryo (Ohad et al., 1996; Chaudhury et al., 1997; Grossniklaus et al., 1998; Luo et al., 2000). The expression patterns for these genes in the female gametophyte are similar, starting before fertilisation in the central cell and continuing after fertilisation in embryo and endosperm until cellularisation (Grossniklaus et al., 1998; Kinoshita et al., 1999; Ohad et al., 1999; Vielle-Calzada et al., 1999; Luo et al., 2000). The \textit{FIS}-class genes encode homologues of the Polycomb Group (PcG) proteins from \textit{Drosophila}. FIS1 or MEA is closely related to the SET-domain protein E(Z), FIS2 is a zinc finger protein related to \textit{SUPPRESSOR OF ZESTE12} (SU(Z)12), and FIS3 (or \textit{FERTILISATION INDEPENDENT ENDOSPERM - FIE}) is a WD40 protein related to \textit{EXTRA SEX COMBS} (ESC) (Grossniklaus and Schneitz 1998; Luo et al., 1999; Ohad et al., 1999). Other mutants with \textit{fis}-like phenotypes have been recently identified \textit{medicis} and \textit{borgia} (Guitton et al., 2004). MEDICIS was shown to be an orthologue of the yeast \textit{MULTICOPY SUPPRESSOR OF IRA} (MSI1) and the histone binding protein p55 of \textit{Drosophila}. p55 interacts with ESC and E(Z) and a histone deacetylase, thus providing a direct link between PcG function and chromatin modifications (Tie et al., 2001). This interaction has been conserved in Arabidosis, as MSI1 was shown to interact with FIE and MEA (Kohler et al., 2003). The molecular identity of \textit{BORGIA} is unknown (Kohler et al., 2003; Guitton et al., 2004).
The *Drosophila* and human orthologues of PcG genes encode proteins that interact with and are controlled by Retinoblastoma protein (Rb), a tumor suppressor involved in the G1/S transition of the cell cycle progression (Pasini et al., 2004). Plant homologues of Retinoblastoma-related proteins (RBR) also control cell proliferation in the female gametophyte and mutants show *fis*-related phenotypes, suggesting that Retinoblastoma-mediated control of PcG activity also occurs in plants (Ebel et al., 2004). Interestingly, the autonomous proliferation of female gametophytic cells in *rbr* mutants seem to be distinct to that of *fis* mutants as they do not express endosperm-specific markers (Ingouff et al., 2006). Similarly, mutation in *msi1* results in a distinct phenotype relative to *mea*, *fie* and *fis2* in that the egg cells, in addition to the central cell, also display autonomous proliferation, producing a parthenogenetic non-viable embryo (Guitton and Berger 2005). MSI1 interacts with RBR as well as FIE (Ach et al 1997; Kohler et al., 2003) therefore, a complex containing the PRC2 related MEA/FIE/FIS2 along with MSI1/RB is likely to be responsible for cell cycle arrest in the central cell of the female gametophyte (Guitton and Berger 2005). Additionally, the distinct phenotypes of *msi1* and *rbr* suggests that MSI1 and RBR could perform other gametophytic functions independent of FIS complex, associated with egg cell arrest and establishment of female gametophytic cell identity (Guitton and Berger 2005). In support of this, it has been shown that RBR plays a central role in the establishment and maintenance of stem cell identity in the root meristem (Wildwater et al., 2005).

Besides the repressive roles demonstrated for the FIS complex in the female gametophyte, *fis* mutants also display a post-zygotic phenotype. Endosperm of WT Arabidopsis plants typically go through an initial syncytial phase of cell division characterized by mitotic division without cytokinesis (Olsen, 2005). During this
phase, nuclear migration occurs and distinct domains are formed in the anterior and posterior poles of the endosperm (Olsen, 2005; Ingouff et al., 2005). Cellularization occurs soon after the formation of these mitotic domains. Pollinated fis seeds produce defective endosperm without the characteristic cellularization stage and as the result of this, embryos abort at the heart stage (Ingouf et al., 2006; Sorensen et al., 2000). Detailed analysis of fis mutant endosperm development showed that they are defective in mitotic control, nuclear migration and the failure to establish some aspects of antero-posterior polarity during the syncytial phase (Guitton et al., 2004; Sorensen et al., 2001; Ingouff et al., 2005). The biological function of the FIS genes and the proposed participation of cell-cycle regulators MSI1/RBR in the FIS-complex is consistent with them playing a role in the proliferation and patterning defects observed in fis endosperm (Ingouff et al., 2005). It is possible that these proteins form a complex conserved throughout evolution due to their important roles in controlling development.

The proposed mechanism of action of the FIS complex is related to chromatin modifications that alter gene expression. This biochemical property has been well characterized in the homologues of the FIS complex in Drosophila and humans. The Drosophila PRC2 complex targets genes which contain conserved sequences in their promoters, the Polycomb Repressive Element (PRE) (Sengupta et al., 2004). These target genes are essentially Homeobox (Hox) genes that are involved in the control of pattern formation (Francis and Kingston 2001; Ringrose and Paro, 2007). Thus the PRC2 complex in Drosophila establishes a cellular memory of the tissue by silencing a particular set of genes not involved in tissue function. Gene repression is generally achieved through the additional of the tri-methyl mark on lysine 27 of histone 3 (H3K27), an epigenetic mark associated with transcriptional repression. Central to
this function is the SET domain protein E(z) whose histone methyltransferase activity is only detected in the presence of the other members of the PRC2 complex ESC, SUZ12 and p55 (Cao et al., 2004). Similar to Drosophila, the Arabidopsis complex has been shown to contain similar members, with FIE a homologue of ESC, FIS2 a homologue of SUZ12, MEA a homologue of E(z) and MSI1 of p55. However, the targets of the Arabidopsis MEA/FIS2/FIE/MSI1 complex have been less well characterized. A transcriptomics approach identified the type I MADS box gene \textit{PHERES1} (\textit{PHE1}) as a target of MEA/FIE and this correlated with the presence of H3K27 marks and transcriptional repression of this gene (Figure 3) in the central cell and early developing seed (Köhler et al., 2005; Makarevich et al., 2006). Therefore, the FIS complex controls seed development and is formed by proteins that are conserved in animals and have a similar biochemical function of regulating gene expression through chromatin modifications.

\textit{Imprinting mechanisms of FIS genes}

Mutations in the FIS genes all show parent-of-origin effects such that only when the mutated allele is transmitted through the maternal parent the effects on seed development can be observed (Ohad et al., 1996; Grossniklaus et al., 1998; Luo et al., 2000). The nature of the parental effect has been shown to be due, at least in part, to genomic imprinting. Imprinting refers to parent-of-origin specific gene expression. At least two members of the FIS complex have been shown to be imprinted: FIS2 and MEA. The mechanism involved in establishing the imprints are distinct in these two genes. \textit{MEA} expression is restricted to the egg and central cell before fertilization, after fertilization, only the maternal allele is expressed in the endosperm and biallelic
expression occurs in the embryo (Kinoshita et al., 1999; Vielle-Calzada et al., 1999). Therefore, MEA imprinting is specific to the endosperm, as is the case for FIS2 and FWA. The MEA/FIE complex deposited in the central cell is responsible for silencing the expression of the paternal MEA allele in a negative feedback mechanism (Gehring et al., 2006; Jullien et al., 2006; Baroux et al 2006). Establishment of paternal MEA silencing occurs via binding of MEA protein to regulatory sequences of MEA and the addition of H3K27 marks (Baroux et al., 2006; Gehring et al., 2006; Jullien et al., 2006). DNA methylation has no effect on MEA imprinting, as paternal alleles are not activated in hypomethylated pollen (Jullien et al., 2006; Gehring et al., 2006).

Imprinting of FIS2, along with the homeobox gene FWA, also occurs in the endosperm. However, in contrast to the imprinting of MEA, both FIS2 and FWA are dependent on the activity of MET1, as paternal alleles are activated when derived from a hypomethylated genome and maternal expression correlated with hypomethylation at the FIS2 and FWA locus (Kinoshita et al., 2004; Jullien et al., 2006). The MEA/FIE complex also regulates imprinting of the target gene PHE1. PHE1 is preferentially expressed from the paternal allele right after fertilization and in the absence of MEA activity, PHE1 expression from the maternal allele was strongly up-regulated (Köhler et al., 2005). It is interesting to note that imprinting mechanism do not always reflect complete silencing of the imprinted allele, as PHE1 maternal expression is still detected at low levels (Köhler et al., 2005). Therefore, MEA is a paternally imprinted gene that establishes maternal imprints on target genes.

Interestingly, activation of FIS2 and MEA expression in the female gametophyte before fertilization occurs via the activity of a DNA glycosylase, DEMETER (Jullien et al., 2006; Choi et al., 2002). DEMETER has been shown to
Figure 3. Schematic representation of the FIS complex. FIS proteins interact to form a chromatin remodeling complex that promotes de methylation of lysine 27 of histone 3 (H3K27) on target genes such as the MADS-box gene *PHERES1* leading to transcriptional repression. H3K27 methyltransferase activity is conferred by the SET domain of FIS1/MEA
actively remove 5’-methyl groups from cytosines from the MEA promoter in the central cell and early endosperm (Gehring et al., 2006), while FIS2 expression was significantly reduced in ovules carrying a mutant dme allele (Jullien et al., 2006). As with other fis mutants, mutations in DME have a maternal gametophytic effect on seed development causing embryo and endosperm abortion; however, dme mutants do not show the fis phenotype of autonomous central cell proliferation (Guitton et al., 2004; Choi et al., 2002). Seed abortion in dme is at least partly due to the failure to activate MEA expression in the female gametophyte (Choi et al., 2002). Mutations in MET1 are able to suppress dme seed abortion, and this occurs via the restoration of MEA expression (Xiao et al., 2003). Although met1 mutants do not directly affect MEA imprinting (Jullien et al., 2006), the restoration of MEA expression in dme and met1 double mutants suggests an indirect effect of DNA methylation on MEA activation (Xiao et al., 2003; Gehring et al 2006).

The activation of FIS2 and MEA expression in the female gametophyte and seed development relies on the establishment of silencing patterns once vegetative growth resumes. These epigenetic silencing mechanisms in vegetative tissues are also established by PcG complexes that involve the FIS2 homologue EMBRYONIC FLOWER 2 (EMF2), the MEA homologues SWINGER (SWN), CURLY LEAF (CLF) and also FIE (Jullien et al 2006). Therefore, FIE seems to be the central player in recruiting PcG complexes during different stages of the plant life cycle. Consistent with this observation are the vegetative defects and flowering time phenotypes observed in FIE co-suppressed lines and the ability of FIE to bind to CLF, SWN and EMF2 (Kinoshita et al., 2001; Katz et al., 2004; Wang et al., 2006; Chanvivattana et al., 2004).
**FIS-class genes and DNA methylation**

The parent-of-origin effects of the *fis* mutations together with the observation that DNA methylation has also been shown to mediate parent-of-origin effects on seed development in *Arabidopsis* suggests that there is an association between the mechanisms that control genomic imprinting, DNA methylation, and the FIS-class genes. Indeed, a genetic interaction between the FIS genes and DNA methylation has been shown by the observation that pollen from a hypomethylated plant harbouring a MET1 anti-sense gene can rescue the *fis* phenotype, independent of FIS function (Vielle-Calzada et al., 1999; Luo et al., 2000; Vinkenoog et al., 2000). Contrary to other *fis* mutant backgrounds, hypomethylated pollen does not restore the expression of an endosperm-specific marker in *fie* and *msi1* mutants, although seed rescue still occurs, indicating that in these backgrounds seed rescue occurs through a alternative pathway (Guitton et al., 2004). These results suggest that DNA methylation controls paternally imprinted genes and that zygotic expression of these imprinted genes, that would otherwise be silent, is sufficient to rescue seed development. Additionally, by combining a hypomethylated maternal genomic background and the *fie* mutant background, endosperm development was significantly altered, showing aspects of cellularization and regional differentiation, normally absent in a normally methylated maternal background (Vinkenoog et al., 2000).

This maternal effect of DNA methylation on gene expression was also demonstrated in the MEA/FIE target gene *PHE1*. In *mea* mutants, *PHE1* expression is upregulated and extended to later stages of seed development, contrasting to wild type plants, when *PHE1* expression is restricted to a brief period after fertilization. Mutations in DECREASE IN DNA METHYLATION 1 (DDM1), a SWN/SNF
chromatin remodelling protein, also result in a global decrease in genomic DNA methylation (Kakutani et al., 1999). A maternal *ddm1* hypomethylated genome was able to restore *PHE1* expression in the absence of functional MEA activity and this correlated with rescue of seed abortion (Köhler et al., 2003). Together these findings indicate that DNA methylation interacts with PcG complex function to control seed development. Rescue of *fis* phenotype even in the absence of FIS activity further indicates that DNA methylation targets a different set of yet unidentified genes in both the female and male gametophyte.

**Diversification and redundancy of PcG complexes in plants**

Plant PcG complexes are homologous to PRC2 complexes in *Drosophila* and contain a basic protein composition consisting of a WD-40 protein, a SET-domain protein and a VEFS/Zn-finger domain protein. In the Arabidopsis PcG complex that controls seed development these proteins are FIE, MEA and FIS2, respectively. FIE is a single copy gene in Arabidopsis; however, MEA and FIS2 contain other family members. The SET-domain is diverse, with 29 genes containing this domain in the Arabidopsis genome (Baumbusch et al., 2001); but aside from MEA, there are 2 other members that are related to the PcG protein E(z): CURLY LEAF (CLF) AND SWINGER (SWN OR EZA1). Similarly, the Arabidopsis genome also contains two other Suz(12) homologues aside from FIS2: EMBRYONIC FLOWER2 (EMF2) AND VERNALIZATION2 (VRN2). Recent genetic analysis and protein interaction studies have shown that these different PcG proteins have the potential to interact with each other and thus form distinct PcG complexes. For instance, FIE can interact with all three E(z)-related SET domain proteins in Arabidopsis (Wang et al., 2006); whereas
the MEA, SWN and CLF can all form interactions with the three VEFS-domain proteins (Chanvivattana et al., 2004). Biologically, this has enabled plants to use similar PcG-based epigenetic mechanisms to control different processes in the plant life cycle. The MEA/FIE/FIS2 complex is involved in controlling endosperm development (Luo et al., 2000), while FIE/CLF/EMF2 mediates the transition from vegetative to floral meristem identity (Goodrich et al., 1997; Chanvivattana et al., 2004; Katz et al., 2004). FIE/VRN2/CLF/SWN also seem to form a complex to mediate flowering response to cold temperatures, vernalization (Wood et al., 2006).

Central to the formation of these distinct PcG complexes is FIE, whose WD-40 protein structure serves as a platform that mediates protein-protein interactions (Ng et al., 1997). Although functional diversification has occurred in the plant PcG complexes, genetic analysis has shown that these complexes might have overlapping functions, characterized by genetic redundancy. The double mutant swn/swn; mea/+ show an enhanced phenotype of autonomous endosperm development compared to single mea/+ mutants (Wang et al., 2006). A mutation in SWN also enhanced the phenotype of clf single mutants (Chanvivattana et al., 2004). Consistent with redundancy in PcG function is the ability of distinct PcG complexes to alter chromatin structure and regulate the expression of the common target genes PHE1 and FUSCA3 (Makarevich et al., 2006). Therefore, unlike Drosophila in which there is only one member of each of the PRC2 complex, in Arabidopsis PcG proteins have duplicated and diverged to control several aspects of plant development.

Functional and biochemical characterization of plant PcG complexes are mostly concentrated on the sexual dicot plant model species Arabidopsis. The degree of conservation and similarity in complex components and their function has yet to be determined in great detail in other plant species. A phylogenetic study on the
evolution of MEA has shown that it is specific to the Brassicaceae family, arising from a duplication event of a chromosome segment that contains SWN (Spillane et al., 2007). This implies that in other plants PcG function might have evolved differently. Characterization of PcG homologues has been carried out in maize, with the focus on determining whether imprinting mechanisms have been conserved in monocot species. As with Arabidopsis, maize contains 3 E(z)-like sequences all of which are expressed in developing seeds (Haun et al., 2007). Of the three maize E(z) genes, only Mez1 is imprinted in the endosperm, while Mez2 and 3 show bi-allelic expression (Haun et al., 2007). Different copies of E(z) genes in plants most likely reflect the functional diversification of these complexes during the plant life cycle. Contrary to Arabidopsis, maize contains two FIE genes, ZmFIE1 and 2. These two genes have diverged to have distinct spatial and temporal expression patterns and only ZmFIE1 is imprinted in the endosperm, whereas ZmFIE2 is expressed in the gametophyte before fertilization and mostly in the developing embryo (Hermon et al., 2007; Danilevskaya et al., 2003). The paternal silencing of both Mez1 and ZmFIE1 has been strongly correlated to hypermethylation in differentially methylated regions (Gutierrez-Marcos et al., 2006; Haun et al 2007; Hermon et al., 2007). Therefore, although there are differences in the phylogenetic relationships between Arabidopsis and maize PcG genes, a subset of these genes have been selected for imprinting in the endosperm. This suggests that imprinting of PcG genes is likely to be an important aspect of endosperm development in flowering plants.
Developmental, genetic and molecular aspects of apomictic development

Apomixis is a dominant trait in monocots

In all of the apomictic species studied to date, apomixis has been shown to be inherited as a dominant trait. In particular, the grass species have received a lot of attention due to the prevalence of apomixis in this group, their economical importance in South American and North American countries and mainly because of the possibility of transferring apomixis into cereal food crops. With the exception of the maize relative *Tripsacum*, these studies have largely relied on using apomicts as pollen parents, as reciprocal crosses are not feasible, and progeny are generally scored based on the presence of a *Panicum* type embryo sac, i.e. the presence of a 4-nucleate embryo sac in contrast to the 8-nucleate *Polygonum*-type embryo sac of sexual reproduction and in some cases, the use of molecular markers associated with apomixis. *Tripsacum* is diplosporous and has an eight nucleate embryo sac. For the aposporous grass genus *Paspalum, Pennisetum, Poa, Brachiaria* and *Panicum*, apomixis seems to be inherited as a single dominant trait with the Aaaa genotype. However, in some studies, apomixis has been shown to be controlled by multiple loci with modifiers. In *Poa pratensis*, a detailed analysis using flow cytometry of single seeds (FCSS) suggests a 5 loci model, in which not only the formation of an unreduced embryo sac during apomixis initiation was controlled by one locus, but also the ability for parthenogenetic development was shown to be an independent locus (Matzk et al., 2005). Besides these two loci, genetic elements that act in the sexual pathway to prevent apomixis are also suggested to be in a recessive state for proper expression of the trait (Matzk et al., 2005). These analyses were largely based
on the variability of the penetrance of individual components of apomixis, apomeiosis and parthenogenesis in progeny.

**Multiple loci control apomixis in the Asteraceae family**

Other studies performed in dicot species of the Asteraceae family have also highlighted that apomixis is controlled by at least two independent loci. This is the case for apomictic species of the genus *Hieracium, Erigeron* and *Taraxacum*. While *Taraxacum* is diplosporous, *Hieracium* and *Erigeron* are aposporous. However, in all three genera, seed initiation is completely autonomous, and both embryo and endosperm development are independent of fertilization. This suggests that the evolution of an autonomous endosperm has occurred independently of the mode of apomixis initiation, as in both diplosporic and aposporic species endosperm can develop without fertilization.

The segregation of autonomous endosperm formation has not been completely elucidated. Genetic evidence in *Erigeron annuus* has shown that progeny from a cross between an apomictic pollen donor and a sexual maternal parent can yield equal proportions of four distinct progeny classes representing the different combinations of diplospory/meiosis and autonomous seed development/fertilization derived seed (Noyes et al., 2007). These results show that in this species, apomixis initiation and parthenogenesis are controlled by two distinct loci and the autonomous seed component representing both autonomous embryo and endosperm are tightly linked. A three-component model has also been proposed for *Taraxacum*, as in this system, parthenogenetic embryo development was observed without the formation of an autonomous endosperm, suggesting that these two components are genetically
unlinked (van Dijk et al., 2003). In *Hieracium*, a deletion mutagenesis strategy was used to generate markers associated with apomixis and it was observed that apomeiosis and parthenogenesis are independently controlled (Catanach et al., 2006). While this approach has the benefit of creating mutants that are in an isogenic background relative to the wild parent apomict, the screening methods used did not allow a characterization of the autonomous endosperm components and its genetic linkage to parthenogenetic embryo formation in this group.

**Epigenetic aspects of apomictic development**

There are several lines of evidence that suggests a relationship between epigenetic regulation and apomictic development. Apomixis is not as tightly regulated as the sexual process. Variability in developmental timing and diversity exists even within a vegetatively propagated plant. The diploid apomictic *Hieracium piloselloides* shows an array of developmental variations, producing an additional form of apomixis, adventitious embryony. Germinated seedlings vary widely in morphology and are often abnormal yet as they grow they regain normal growth pattern (Koltunow et al., 2000). Similarly, autonomous endosperm development in the apomictic *Hieracium piloselloides* initiates with irregular cell division patterns, differing from the fertilisation-derived endosperm of the sexual plant. This alteration in cell division and patterning in early stages of autonomous endosperm development resembles that seen in plants with defects in chromatin remodelling genes. In later stages, autonomous endosperm development recovers to a more regular and defined pattern of cellularisation similar to the sexual seed (Tucker, 2003). Developmental variation in a single genetic background and abnormal growth patterns that revert to normal are
often hallmarks of epigenetically regulated processes where epimutations deregulate timing of expression of developmental genes.

Some clues as to how chromatin metabolism might be involved in apomixis regulation have come from the characterization of the locus controlling apomixis in the grass species *Pennisetum* and *Cenchrus*. In these species, this locus, termed apospory-specific genomic region (ASGR), is strongly associated with repetitive sequences and heterochromatin formation (Ozias-Akins et al., 1998; Goel et al., 2003). This condensed chromatin configuration of the ASGR has evolved as either the result of or the consequence of the lack of recombination observed for this region. One possibility is that suppression of recombination resulted in sequence divergence, which ultimately leads to the hemizygosity. Interestingly, this hemizygous state for the locus associated with apomixis seems to be prevalent in grasses, as it has also been suggested for *Tripsacum dactyloides* (Grimanelli et al., 1998) and *Paspalum simplex* (Calderini et al., 2006). These regions are not necessarily associated with heterochromatin in all of the apomictic grass species, as cytogenetic analysis of the APO locus in *Paspalum* mapped this region to a heterochromatin-poor region (Pupilli et al., 2001; Calderini et al., 2006). However, comparative analysis of the apo locus with syntenic regions in the rice genome seems to suggest chromosomal re-arrangements, among other mechanisms, by transposons (Calderini et al., 2006). In one particular case, this transposon-induced re-arrangement has resulted in local interruption of the coding capacity of a coding sequence. It is reasonable to suppose that genes associated with regions that control apomixis in the grass family, either are associated with, or influenced by the presence of repetitive DNA. Under this assumption, it is conceivable that repetitive sequences in the vicinity of expressed genes may attract epigenetic changes with the potential to deregulate gene expression.
and trigger elements of sexual development in different cell-types, a characteristic of apomixis.

Epigenetic control of sexual seed development by the FIS-class genes, as explained in the above sections, and the elements of autonomous seed development that occur upon mutations in these genes have also highlighted the possibility of an epigenetic control for apomixis (Rodrigues and Koltunow 2005). The fis phenotypes have emphasized the need to elucidate the function of the FIS chromatin remodelling complex in apomixis, whether apomictic plants express components of this complex in reproductive structures, if their expression are conserved both spatially and temporally throughout ovary development in apomictic plants relative to sexual plants and if the individual components are conserved for similar complexes in apomictic plants are all questions that remain to be addressed.

Isolation of genes differentially expressed in apomictic species

The occurrence of sexual and apomictic accessions in closely related species or even within the same species as agamic complexes has stimulated the search for genes that are specifically expressed in apomictic plants by using several different methods of gene expression comparisons. This approach has been used successfully in several grass species such as Poa pratensis, Pennisetum ciliare, Paspalum notatum, Panicum maximum and Brachiaria brizantha. In Pennisetum ciliare, suppressive subtractive hybridization (SSH) was used between leaf and ovary tissue to isolate sequences that were specific to apomictic ovaries and differentially expressed in ovaries of sexual plants. The authors show that two genes, Pca21 and Pca24, are
specifically expressed in apomictic ovaries when compared to ovaries of sexual plants (Singh et al., 2007). While sequence analysis showed that these genes contain conserved protein domains, such as coiled-coil-helix-coiled-coil in Pca24, they do not clearly show homology to previously characterized genes that allow for the assignment of putative functions. Interestingly, a well characterized gene, Somatic Embryogenesis Receptor Kinase (SERK) was identified in a cDNA-AFLP screen for differentially expressed genes in Poa pratensis (Albertini et al., 2005). SERK was previously identified as a marker for the competency to initiate somatic embryogenesis in cell culture (Hecht et al., 2001). Expression analysis in apomictic ovaries showed discrete differences in timing and spatial distribution of PpSERK mRNA in apomictic ovaries. Markedly, PpSERK transcripts are present in the megaspore mother cell of sexual ovules, but was absent from this cell in apomictic ovules, where it was shown to be in neighbouring nucelar cells (Albertini et al., 2005). The authors hypothesize that the presence of SERK in the nucelar cells of apomictic plants might mark competency of these cells to become apomictic embryo sacs (Albertini et al., 2005).

Comparisons of differentially expressed sequences in other species show interesting similarities that could be related to apomictic development. A cDNA-AFLP mRNA profiling method was employed in Poa pratensis and a Differential Display RT-PCR method in Brachiaria brizantha both detected similar sequences based on homology searches (Albertini et al., 2004; Rodrigues et al., 2003). Apomictic ovaries of P. pratensis and B. brizantha in early developmental stages both express genes that encode enzymes involved in DNA recombination/repair and cytoskeletal microtubule formation. A gene also encoding a cytoskeletal module protein was also detected in apomictic ovaries in Paspalum notatum (Pessino et al.,
The lack of meiotic recombination is a distinct feature of apomictic development. The presence of enzymes related to these functions in apomictic ovaries might be related to the impairment of this pathway during apomixis initiation. Furthermore, the presence proteins involved in microtubule organization might indicate that re-arrangements and nuclei positioning during embryo sac development might be an important developmental cue for apomixis. Notably, the grass species from which these genes were isolated produce a *Panicum*-type embryo sac, which differs by the presence of only 4 nuclei compared to the 8-nuclei present in the typical *Polygonum*-type embryo sac of the majority of sexual plants.

Differential screening approaches also identified genes that are expressed in sexual plants but down-regulated in apomictic plants. APOSTART was identified in *Poa pratensis* and its expression seems to be related to sporogenesis, possibly in program cell death of non-functional megaspores (Albertini et al., 2005). It is not clear how the silencing of APOSTART in apomictic plants might be the cause or consequence of apomixis initiation. Similarly, the expression of the MADS-box gene *DEFICIENS* is spatially altered in apomictic *Hieracium* plants during early ovule formation and is down-regulated in a specific region where apomixis initiation occurs compared to sexual development (Guerin et al. 2000). Therefore, differential screening of gene expression in different apomictic organisms indicate that in order for apomixis to initiate, developmental programs that are established in the sexual pathway might have to be “turned off” in specific cell types. Although the differential screening strategies used in these studies to identify apomixis-specific genes were efficient in detecting differentially expressed genes, ultimately, a model system in which to carry out functional studies in a natural apomictic species is still required for the confirmation of functional annotations.
**FIS complex and autonomous development**

The parent-of origin effect of the *fis* mutations suggests there is maternal control of early seed development. The putative MEA/FIS2/FIE complex functions to repress endosperm development in the absence of fertilisation, probably through preventing transcription of target genes, such as *PHERES1*. The molecular cue by which fertilisation alters the function of the complex, perhaps by relieving transcriptional repression and triggering seed development, is still unknown. The observation that cellularized, differentiated endosperm can develop in the absence of fertilisation when combining DNA hypomethylation and *fie* mutant background (Vinkenoog et al., 2000) leads to the speculation that DNA methylation might be an important part of autonomous endosperm development and possibly parthenogenesis. An attractive model can be envisaged in which aberrant DNA methylation marks in key regulatory genes would lead to complete maternal control over reproduction, or autonomous development. Due to the association of *fis* phenotypes with apomictic development, candidates for such key regulatory genes are chromatin-remodelling proteins. Deregulation of the FIS complex for example might lead to altered timing of its repressive functions, therefore allowing the initiation of seed development without fertilisation.

The autonomous endosperm phenotype of the *fis* mutants and the association of FIS proteins with cell cycle control retinoblastoma-like proteins suggest a role for this developmental pathway in apomixis. Expression analysis of *Arabidopsis* FIS promoter GUS fusions in sexual and apomictic *Hieracium* plants showed a conserved pattern in both plants suggesting that sexual and apomictic reproduction are closely related (Tucker et al., 2003). Interestingly, FIS2::GUS expression pattern occurred
early in ovule development in *Hieracium* compared to *Arabidopsis*. Furthermore, the pattern was also slightly shifted in two different apomictic species when compared to sexual plant. In sexual plants, FIS2::GUS expression was restricted to the three megaspores destined to degenerate and was absent in the selected spore and surrounding nucellar epidermis. In apomicts at the time of aposporous initial enlargement, FIS2::GUS expression was found in all 4 megaspores and in the surrounding nucellar epidermis enveloping them. These are cells that are destined to degenerate. Expression was absent in the aposporous initials as well as in the functional megaspore in the sexual plant until their nuclei divided (Tucker et al., 2003). The earlier expression of FIS2::GUS in *Hieracium* may reflect a different function for this gene in *Hieracium* relative to that in *Arabidopsis* that may have enabled the development of apomixis (Eckhardt, 2003). The pattern shift in AtFIS2::GUS expression at aposporous initial cell differentiation may be caused by signals from the AI or to consequences resulting from AI differentiation. To investigate this further, it will be necessary to identify the cis-regulatory elements required for the spatial shift of AtFIS2::GUS expression and to identify factors interacting with these sequences. Furthermore, isolation of a *Hieracium* FIS2 orthologue is critical to functionally dissect its role in apomictic development.

If epigenetic modifications are involved in apomixis, how might altered epigenetic marks such as DNA methylation have arisen and be maintained? The genetic and molecular characteristics of DNA hypomethylation mutants suggests that disturbing the maintenance methylation machinery leads to the accumulation of altered methylation states that are not reset even after the genetic components are restored (Kakutani et al., 1996; Jedelloh et al., 1998; Saze et al 2003). In these mutants, hypomethylated states at specific loci could co-exist with methylated loci
thus generating epigenetic variation (Saze et al., 2003). Apomixis could have arisen through the perturbation of the DNA methylation machinery, such as observed in the *ddm* mutants (Jeddeloh et al., 1999) which would lead to the accumulation of epimutations, causing the mis-expression of key developmental genes. Alternatively, naturally occurring epimutations due to environmental stress or hybridisation would cause altered expression of genes that control the sexual reproductive pathway leading to apomixis. These epimutations would be maintained through DNA replication by maintenance methyltransferases. The recent association of transposable elements to a locus that confers apomixis in the grass species *Pennisetum* sp (Labombarda et al., 2002; Akiyama et al., 2004) also suggests that altered DNA methylation could cause the expression of transposable elements which can lead to the inappropriate expression of neighbouring genes (Lippman et al., 2004; Peaston et al., 2004).

However, there is no molecular evidence that directly shows a significant association between epigenetic regulation and apomixis is lacking. This complex issue can be approached in many different ways. The isolation of *FIS* orthologs in apomictic plants would be a first step to elucidating their expression and function in sexual and apomictic reproductive pathways. Furthermore, if DNA methylation is indeed involved in controlling the expression of genes involved in apomictic development, functional analysis of DNA methyltransferases during key stages of apomictic reproduction could help identify the stages in which this epigenetic mark is important.
Context of thesis project and specific aims

The *Hieracium* model system is a powerful tool to address fundamental questions regarding the developmental aspects of apomictic reproduction. The ability to use marker genes of the sexual reproduction pathway in an apomictic system clarified an important issue concerning the relationship between the sexual and apomictic pathways, demonstrating that apomixis is related to and uses the molecular machinery of the sexual program (Tucker et al., 2003). This is an important conclusion and suggests that apomixis most likely occurs via a deregulation of the sexual pathway. One mechanism in which the sexual program can be deregulated is through the modifications of chromatin remodelling factors. The hypothesis specifically addressed in this project is that FIS complex function is altered in apomictic plants relative to their sexual counterparts. This could entail different primary sequences of complex members that could potentially disrupt complex assembly, altered spatial and/or temporal expression patterns of genes that encode complex members and/or different composition of complex members. Previous studies in *Hieracium* showed that it encodes and expresses at least one of the FIS genes, *HFIE*, and that similar to Arabidopsis, it functions in vegetative development in apomictic plants (Tucker, 2003).

One of the specific aims of this project is to expand on this initial study to determine what role HFIE plays in reproductive development. This will be addressed in Chapter 2 by analysing FIE interactions with other PcG proteins and down-regulating *HFIE* expression in ovules. Results from this study opened the question of linkage of the genetic components of autonomous embryo and endosperm in apomictic *Hieracium*.
Chapter 3 will contain cytological analysis of apomictic mutants that were made available from Crop & Food Research (New Zealand). These mutants have been generated by gamma radiation and have deletions that have impaired the two genetic loci for apomixis. By analysing the cytology in combination of seed viability of progeny derived from these mutants, the relationship between the apomictic and sexual reproductive programs was explored. These mutants also provided an opportunity to explore the segregation of the autonomous endosperm component relative to autonomous embryo formation in apomictic *Hieracium*.

In Chapter 4, the function of MSI1 was addressed. In Arabidopsis mutations in this gene elicits parthenogenetic development of the egg cell. Expression analysis and localization of *HMSI1* transcripts, along with expression in mutant background suggested that *HMSI1* might have a similar function in embryo and endosperm formation as *HFIE*.

The FIS complex also interacts with another epigenetic mark, DNA methylation, in controlling the expression of imprinted FIS genes and by altering the expression of the *fis* phenotypes. Finally, due to its importance in gametophytic and sexual seed development, in Chapter 5 we will also address the role DNA methyltransferases might play in apomictic development by analysing the expression of genes that encode different classes of DNA methyltransferases. Gene expression studies in the apomictic mutants will be a valuable tool to understand the role of *FIS* genes and DNA methyltransferases in apomictic development.
CHAPTER 2

SEXUAL AND APOMICTIC (ASEXUAL) SEED FORMATION REQUIRES THE PLANT POLYCOMB-GROUP GENE

FERTILIZATION INDEPENDENT ENDOSPERM (FIE)
Acknowledgements and clarification of contributions made to this chapter

This chapter was prepared as a draft manuscript to submit for publication. It contains data from the PhD dissertation of Matt Tucker, who initiated the work on FIE function. Matt Tucker’s data include: the isolation of the FIE homologue from sexual and apomictic *Hieracium, HFIE* expression in different tissues, 35S::hpHFIE lines and the Arabidopsis *fie3-2* complementation work that are part of this manuscript. This data is included to give a better context of the work produced in this PhD thesis and is cited accordingly. The protein modelling work was done by Maria Hrmova of the Australian Centre for Plant Functional Genomics. We thank Fritz Matzk for the flow cytometry of single seed analysis.
Abstract

Double fertilization is required during sexual reproduction in plants to activate embryo and endosperm formation in seed development. A Polycomb-Group (PcG) complex, Fertilization Independent Seed (FIS), represses endosperm development in Arabidopsis (At) until fertilization occurs. At\textit{fis} mutants exhibit fertilization-independent endosperm cell proliferation with subsequent seed abortion. Apomictic \textit{Hieracium} plants form seeds asexually without fertilization and here three FIS homologues were isolated from \textit{Hieracium} to investigate FIS function during apomictic seed formation. The role of \textit{FERTILIZATION INDEPENDENT ENDOSPERM (HFIE)} was examined because AtFIE binds many AtFIS proteins. HFIE down-regulation in ovaries of sexual \textit{Hieracium} did not induce autonomous endosperm proliferation. Seed abortion occurred after cross-pollination. HFIE down-regulation in apomictic \textit{Hieracium} significantly inhibited autonomous embryo and endosperm initiation and autonomous seeds displayed defective embryo and endosperm growth. HFIE is essential for both apomictic and fertilization-induced seed initiation in \textit{Hieracium} but does not play a role in repressing autonomous seed formation. Fertilized and autonomous seeds derived from At\textit{fie} embryo sacs also abort. The primary role of FIE in seed development may be to control embryo and endosperm cell proliferation after the activation of seed initiation with FIE being recruited in some sexual species to participate in repressive complexes prior to seed initiation.
Introduction

Seed initiation in sexually reproducing flowering plants requires signals arising from a double fertilization event that occurs in embryo sac found in the ovule of the flower. The egg cell and the central cell of the multicellular embryo sac are the progenitors of the embryo and endosperm in the seed, respectively. One sperm cell fuses with the egg to generate the zygote and initiation of embryo development. The other fuses with the central cell to initiate endosperm development. The central cell contains two polar nuclei and these nuclei either fuse prior to fertilization, as in Arabidopsis, or fuse together with the sperm nucleus during fertilization, as in maize. Fertilization of the central cell in diploid plants produces triploid endosperm containing two maternal and one paternal genome (2m:1p) in contrast to the 1m:1p genome ratio observed in the embryo. In the absence of double fertilization, the egg and central cells remain in a quiescent state and degrade during senescence of the flower suggesting signalling processes are required to activate development of the fertilization products.

Chromatin remodelling Polycomb group (Pc-G) complexes mediate the arrest of the egg and central cell in the Arabidopsis (At) embryo sac. A Pc-G complex called FERTILIZATION INDEPENDENT SEED (FIS) restricts premature proliferation of the central cell (Berger et al., 2006, for review). This complex contains the WD40 domain protein FERTILIZATION INDEPENDENT ENDOSPERM (FIE, Ohad et al., 1999) which interacts with the SET domain containing protein FIS1 (Luo et al., 2000) and the WD40 protein MULTICOPY SUPPRESSOR OF IRA1 (MSI1; Kohler et al., 2003). The VEFS domain protein FIS2 interacts with FIS1 in the complex (Wang et
al., 2006). Loss-of-function mutations in these FIS genes result in central cell proliferation and seed coat development in the absence of fertilization forming seed-like structures that eventually abort. FIS genes are under maternal gametophytic control and fis phenotypes are only observed when the mutation is in the maternal allele. Fertilized seeds derived from fis embryo sacs abort showing defects in embryo and endosperm growth suggesting FIS genes are also essential in the control of cell proliferation (Sorenson et al., 2001).

The Arabidopsis homologue of the mammalian tumour suppressor protein Retinoblastoma, RBR also forms part of the FIS complex and it interacts with both MSI1 and FIE (Katz et al., 2004; Ach et al., 1997). Mutations in RBR lead to cell proliferation in the embryo sac but cellular identity is ambiguous (Ingouff et al., 2006; Ebel et al., 2004). Mutations in MSI1 also show autonomous division of the egg cell that triggers non-viable parthenogenetic embryo formation (Guitton and Berger 2005). Thus, while cell cycle arrest of the central cell in Arabidopsis is controlled by the FIS pathway, egg cell arrest appears to be controlled by MSI1 and other factors independent of the FIS pathway (Berger et al., 2006).

Double fertilization is thought to initiate signal transduction cascades that disassemble repressive complexes in the egg and central cell and activate transcription of downstream seed development genes thereby maintaining a requirement for a paternal contribution for seed formation. Other PcG complexes regulate different developmental pathways during the plant life cycle and FIE participates in a number of these. Different AtPcG complexes can also control common target genes in different stages of the plant life cycle suggesting functional redundancy in PcG complexes containing different protein members (Makarevich et al., 2006; Wang et al., 2006; Chanvivattana et al., 2004; Katz et al., 2004; Kinoshita et al., 2001).
Plants can also form seeds asexually by apomixis, a remarkable mode of reproduction that omits critical events observed in the sexual pathway. Meiosis is avoided prior to embryo sac formation (apomeiosis), embryo development is autonomous and endosperm formation may or may not require fertilization of the central cell (pseudogamy; Koltunow and Grossniklaus, 2003). Fully autonomous seed initiation although rare amongst apomicts is common in the daisy-like Compositae including the genus *Hieracium* which has been developed as a model system for the analysis of apomeiosis and autonomous seed formation (Bicknell and Koltunow, 2004). Two dominant and independent loci confer apomeiosis and autonomous seed initiation in *Hieracium* (Catanach et al., 2006). Yet apomixis is not a distinct pathway because studies using molecular markers in sexual and apomictic *Hieracium* plants have shown that gene expression programs are shared in both pathways once the cell initiating apomixis undergoes its first nuclear division (Tucker et al., 2003). As there is no paternal contribution to either embryo or endosperm development maternally derived information drives the entire process of seed initiation in *Hieracium*. The egg and central cells of apomicts lack the quiescent arrest observed in sexual plants, therefore, one hypothesis is that *Hieracium* homologues of the FIS complex may not be expressed or the FIS complex might be mis-regulated in apomicts facilitating autonomous seed initiation (Ohad et al., 1996; Koltunow and Grossniklaus, 2003).

Partial translational GUS fusions of AtFIS genes are co-expressed during embryo sac and seed development in both sexual and apomictic *Hieracium* indicating that the mechanisms and regulators that direct spatial expression of these Arabidopsis genes to reproductive tissues are conserved in *Hieracium* (Tucker et al, 2003).

The role that FIS complex members play in seed development in species outside Arabidopsis is lacking and has not been functionally addressed in apomictic
plants. In this paper, autonomous seed initiation was examined and three *Hieracium* FIS complex homologues, *HMSII*, *HRBR* and *HFIE* were isolated and found to be expressed in ovaries and developing seeds of sexual and apomictic plants. HFIE function in sexual and apomictic plants was determined by down-regulation of the gene in transgenic plants and partnering interactions between the *Hieracium* and Arabidopsis FIS proteins were determined in yeast two hybrid experiments. AtFIE binds to many proteins in the FIS complex and participates in a number of PcG complexes during plant development. HFIE is required for embryo and endosperm growth and proliferation in both sexual and apomictic *Hieracium* and is not part of a complex that represses central cell proliferation.

**Results**

**FIS complex genes are expressed in ovaries of sexual and apomictic *Hieracium***

Full-length cDNAs for *FIE*, *MSII* and *RBR* were isolated from ovaries of sexual and apomictic *Hieracium* containing mature gametophytes. We failed to isolate a *Hieracium* FIS2 cDNA and the only sequences that resemble FIS2 in *Hieracium* ovaries are similar to At*EMBRYONIC FLOWER 2* (Chaudhury et al., 2001). The putative *HFIE*, *HMSII* and *HRBR* protein sequences are highly conserved between sexual and apomictic plants (99-97%) and share high conservation with Arabidopsis, rice and maize (Table 1, Appendix 1). AtFIE is known to interact with AtFIS1, AtMSI1 and AtRBR in the AtFIS complex therefore we focused on an examination of *HFIE* function during seed initiation in sexual and apomictic *Hieracium*. Only one expressed *HFIE* cDNA was isolated from the ovaries of apomictic and sexual plants. *HFIE* from the apomict differed from the sexual (*HFIE*) in 32 bp over the total of
1110 bp resulting in only one amino acid substitution (glutamic acid in HFIEa to glutamate in HFIEs; Appendix 1). The HFIE proteins share significant conservation with AtFIE (73% identity), with the majority of amino acid differences found outside of the seven WD-40 protein domains (Appendix 1). The Hieracium plants used in this study are polyploids with the sexual being tetraploid and the apomict triploid. HFIE genomic sequences were isolated and compared with the expressed cDNAs to determine gene copy number (Appendix 1). The genomes of apomictic and sexual plants contain four and two non-coding alleles, respectively, with in-frame stop codons and cDNA sequences matching these alleles were not detected. Only one pair of highly conserved alleles matched the expressed cDNA sequences in each plant (Tucker M., 2003) which is consistent with the observation that plant PcG proteins are encoded by small gene families (Makarevich et al., 2006).

**Table 1.** Amino acid sequence identity of FIS-related proteins isolated from apomictic (APO) and sexual (SEX) Hieracium with Arabidopsis thaliana (At), Oryza sativa (Os) and Zea mays (Zm) homologues.

<table>
<thead>
<tr>
<th></th>
<th>FIE</th>
<th></th>
<th>MSII</th>
<th></th>
<th>RBR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SEX</td>
<td>At</td>
<td>Zm</td>
<td>SEX</td>
<td>At</td>
</tr>
<tr>
<td>APO</td>
<td>99.7%</td>
<td>73.7%</td>
<td>74.2%</td>
<td>99.7%</td>
<td>88.4%</td>
</tr>
<tr>
<td>SEX</td>
<td>-</td>
<td>73.5%</td>
<td>74.9%</td>
<td>-</td>
<td>88.6%</td>
</tr>
<tr>
<td>At</td>
<td>-</td>
<td>-</td>
<td>67.0%</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Accession numbers of the sequences used in alignments were as follows: AtFIE – AAD23584; ZmFIE2 – AAO26660; AtMSI1 – NP200631; OsMSI1 – ABF97823; AtRBR - AF245395; ZmRBR3 - DQ124423
**HFIE inhibits central cell proliferation in the Arabidopsis fie-2 mutant**

To examine if HFIE is a functional orthologue of AtFIE, complementation of the Atfie-2 mutant was attempted. Homozygous Atfie-2 lines cannot be generated due to embryo lethality. In emasculated fie-2 heterozygous mutants, 50% of ovules initiate central cell proliferation (Ohad et al., 1996). Expression of functional HFIE protein in emasculated complementation lines should halve the number of ovules exhibiting autonomous central cell proliferation. The AtFIS1 promoter (Luo et al 2000) was used to drive the expression of a full length HFIE cDNA isolated from the apomict in transformed WT Arabidopsis. Pollen from the fie-2/+ mutant was crossed into WT control and plants containing the AtFIS1:HFIE transgene. Two F2 lines containing the fie-2 mutant and the AtFIS1:HFIE transgene were identified, emasculated and observed 3 days later for autonomous central cell development and compared to emasculated fie-2 control lines without the transgene. In control lines, 30 flowers were emasculated and an average of 21% (SD ± 4.8%) of the ovules analysed contained proliferating cells reflecting previous reports for this allele (Ohad et al., 1996). In the complementation lines, an average of 11% (SD ± 2.4%) of ovules from 40 emasculated flowers showed autonomous proliferation of the central cell. This represents the expected 50% reduction in the number of ovules containing autonomous endosperm. However, HFIE could not correspondingly restore embryo viability in pollinated complementation lines and the expected 50% reduction in embryo lethality was not observed. In pollinated lines, the proportion of aborted seeds per silique was similar in both control (49.2% ± 1.4) and complementation lines (49.7% ± 2.2).
HFIE gene expression in *Hieracium*

*HFIE* transcripts are detected by RT-PCR in leaves, roots, sepals, petals, stamens and in ovaries throughout embryo sac and early seed development in sexual and apomictic plants (Figure 1A). *In situ* hybridization was used to examine the spatial localization of *HFIE* transcripts in ovaries and in developing seeds. *HFIE* mRNA was present in the central cell and embryo in apomictic ovules and in surrounding maternal sporohytic tissues (Figure 1B). Similarly, *HFIE* mRNA was detected in sexual embryo and cellular endosperm and generally throughout the ovary (Figure 1D). Signals were not detected when a sense *HFIE* probe was used in serial sections (Figures 1C and E). This contrasts with the expression of *AtFIE* where transcripts are restricted to cells in the embryo sac and absent in maternal tissues of the developing Arabidopsis seed (Spillane et al., 2000). Collectively, the data shows that *HFIE* expression is similar in apomictic and sexual plants and consistent with a role in seed development.

**Down-regulation of HFIE in the embryo sac of sexual Hieracium does not lead to autonomous seed initiation**

To test the function of HFIE in sexual seed initiation in *Hieracium*, primary transgenic lines were generated containing a construct to specifically target RNAi-mediated down-regulation of *HFIE* in the embryo sac and in early seed structures in ovules. We hypothesized that if HFIE is part of a complex repressing seed initiation, an autonomous central cell proliferation phenotype would be observed like that found
Figure 1. Expression analysis of HFIE. (A) RT-PCR analysis showing HFIE expression throughout ovary development, leaf, and other floral tissues (OFT) of sexual (Sex) and apomictic (Apo) Hieracium. Equal loading of cDNA samples was verified according to expression of β-tubulin (β-TUB). Data compiled from Tucker, M (2003). Events of ovule and seed development are schematically represented. (B-E) Analysis of the spatial distribution of HFIE transcripts by in situ hybridization. (B) apomictic stage 6 ovary section were hybridized with an anti-sense (A/S) FIE probe showing signals in the embryo (em), central cell (cc) and throughout maternal tissue; (C) serial section of (B) hybridised with control sense (S) probes; (D) stage 10 ovary of cross-pollinated sexual plant hybridized with A/S probe showing expression in embryo (em) and endosperm nuclei (ne); (E) serial section of (D) hybridized with S probe. Bar = 50um.
in AtFIS complex mutants. Previously described *AtFIS1* promoter sequences were used to drive the expression of a hairpin construct, AtMEA::hpHFIE, during the mitotic events of embryo sac formation through to the heart stage of embryogenesis in *Hieracium* (Tucker et al., 2003).

Three independent transgenic tetraploid sexual *H. pilosella* lines containing one or two copies of AtMEA::hpHFIE that also displayed significantly reduced *HFIE* mRNA expression in ovules were selected for examination. Figure 2A shows that *HFIE* was not completely silenced in ovules because the promoter was selected to down regulate *HFIE* in gametophytic structures and not *HFIE* mRNA found in maternal sporophytic ovule tissues. *HFIE* was not down-regulated in leaves of these transgenic plants and they resembled untransformed control plants with respect to vegetative and floral growth (data not shown) but seed set and germination was defective (Table 2; Figure 2B).

Multiple florets in *Hieracum* are located together in a capitulum and embryo and endosperm initiation and growth is marked by ovary enlargement and external pigmentation to produce a filled black (FB) achene (fruit) that generally contain viable seeds (Koltunow et al., 1998a). The terms achene and seed are used here interchangeably. Seed set in sexual *H. pilosella* is generally dependent upon cross-pollination but can be partially self-compatible under warm growth conditions and this was evident in control plants. Table 2 shows that down-regulation of *HFIE* in ovaries of transgenic lines did not increase the percentage of FB seeds as would be expected from the induction of autonomous seed formation. The percentage of FB seeds only increased following cross pollination in control and transgenic plants indicating that fertilization was required for seed set. Seed viability was however drastically reduced in unpollinated and cross pollinated transgenic lines (Table 2).
Cytological analysis indicated that embryo sac formation was normal. Observation of seed initiation in cleared ovules of control and transgenic plants at stage 10 of capitulum development enabled comparison of the percentage of arrested embryo sacs and initiated seeds (Figures 2 C-F; 2K). In control sexual *H. pilosella* plants, 50% of unpollinated florets contained fused polar nuclei positioned close to the egg cell in ovules (Figures 2C; 2K) and the frequency of seeds with embryos and cellular endosperm increased following cross pollination (Figures 2D; 2K) In transgenic sexual lines containing AtFIS1:hpHFIE lines, the ovules of unpollinated florets also contained an egg and central cell (Figure 2K), however, in approximately half of these the central cell had failed to migrate near the egg cell (compare Figures 2C and E). In cross-pollinated florets, embryos aborted development at the globular stage and were irregular in structure. Nuclear endosperm was evident but failed to cellularise (Figure 2F, 2K). Collectively these data show that down regulation of HFIE does not lead to fertilization independent embryo and endosperm seed initiation in sexual *Hieracium*. HFIE is required for fertilization-induced endosperm cellularization and embryo growth.

**Down-regulation of HFIE in ovaries inhibits autonomous seed initiation and leads to early seed abortion in apomictic Hieracium**

Unlike sexual *Hieracium*, egg and polar nuclei do not enter into a quiescent state at embryo sac maturity in ovaries of the apomict. Embryo and endosperm initiated together but in a stochastic manner amongst florets in the capitulum of the self-incompatible apomict, typically before the flower opens. Fusion of polar nuclei was a pre-requisite for autonomous endosperm initiation from cytological
observations (Figure 2G) and flow cytometry of individual FB seeds (n=178) which confirmed that the residual endosperm located in the aleurone of the seed had double the ploidy of the embryo (data not shown). Differences were observed in the spatial patterning of the early nuclear endosperm divisions in the apomict relative to those in the sexual plant. Endosperm nuclei and associated cytoplasm clumped together in the apomict with irregular spacing between nuclei, however this normalized with increasing nuclear divisions and at cellularisation the endosperm of the apomict resembled that of the sexual plant (data not shown).

Four independent transgenic plants with reduced HFIE expression in ovules (Figure 2A) and normal vegetative and floral growth habit were analysed. They all showed significantly reduced capacity to autonomously develop FB seeds and seed viability was very low compared to untransformed control plants (Table 2; Figure 2B). Cytological examination of the transgenic plants showed that the initiation of apomixis and the events of apomeiosis and embryo sac formation occurred normally with mature embryo sacs containing egg cell and central cell (Figure 2G), thus HFIE is not essential for these early events in the apomictic pathway. Autonomous embryo and endosperm formation were severely inhibited. Figures 2H shows that in control plants heart stage embryos with cellular endosperm are evident in ovaries from stage 10 capitula but in transgenic lines an average of 40% of stage 10 ovules contained egg cells and a fused polar nucleus indicating that autonomous seed initiation was arrested (Figures 2I and 2L). The other 40% of ovules contained irregular embryos that had aborted at an early stage that was always accompanied by nuclear endosperm that failed to cellularize (Figures 2I and 2L). Pollination of the transgenic lines with a tetraploid apomict plant induced endosperm cellularization but the embryo remained arrested at an irregular globular stage (Figures 2J, 2L). Therefore, HFIE appears to be
Figure 2. Analysis of AtMEA::hpHFIE lines. (A) HFIE expression in ovules of transgenic lines of sexual (SEX) and apomictic (APO) Hieracium plants, HFIE expression levels in control (C) plants was used as reference; (B) Representative test of seed viability of AtMEA::hpHFIE lines; picture of seeds from an apomictic control (CON) plant (left panel) and transgenic line 18 (right panel) are shown; (C-F) Cytological analysis of sexual Hieracium plants, all ovules were analysed at stage10; (C) unpollinated control (SEX UP) plants do not form embryos or endosperm and mature embryo sacs contain egg cells (EC) and fused polar nucleus (FPN) that eventually senesce; (D) upon cross-pollination (SEX CP), control ovules generally contain heart stage embryo (HE) and cellular endosperm (CE); (E) unpollinated stage 10 ovules from sexual AtMEA::hpHFIE lines (hp) contain mature gametophytes with egg cell and fused polar nucleus that do not initiate seed development; often the fused polar nucleus has not migrated towards the egg cell as in control plants (compare to C), no autonomous seed initiation was observed in these lines; (F) cross-pollinated ovules from sexual AtMEA::hpHFIE lines contain irregular embryos (IE) that abort and nuclear endosperm (NE) that fail to cellularized; (G-J) Cytological analysis of apomictic AtMEA::hpHFIE lines, with the exception of (G), all ovules were analysed at stage 10. (G) stage 6 ovules of control apomictic Hieracium plants generally contain an embryo sac with an egg cell (EC) and fused polar nucleus (FPN); (H) unpollinated apomictic control ovules produce autonomous heart stage embryos and cellularized endosperm similar to the cross-pollinated sexual plant (compare to D); (I) stage 10 ovules of apomictic AtMEA::hpHFIE lines contain mature gametophytes that arrest with an egg cell (EC) and fused polar nucleus (FPN); or (J) autonomous seed development initiates but embryo formation is irregular (IE) and abort early with nuclear endosperm (NE) that fails to cellularize; (K) frequency of phenotypes in sexual Hieracium untransformed control plants and AtMEA::hpHFIE lines. CC - central cell, EM – globular or heart stage embryo. Other legends are as above. Frequencies are shown as the average of three transgenic lines and error bars represent standard deviation of the respective phenotypes; (L) frequency of phenotypes in apomictic Hieracium untransformed control plants and AtMEA::hpHFIE lines. Legends are as above. Bars (C-J) = 50μM
<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>3%</td>
<td>4%</td>
<td>6%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31%</td>
<td>54%</td>
<td>65%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>62%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 2. Seed analysis of sexual (SEX) and apomictic (APO) Hieracium plants transformed with constitutive (35S) or embryo sac-specific (AtMEA::hpHFIE) hairpin constructs for the down-regulation of HFIE expression (hpHFIE).**

- % Germination: Number of germinated seedlings per capitula in parentheses. Germination scores per total number of seeds divided by 100.
- % Mean and standard deviation of individual capitula: a minimum of 4 capitula were analysed for each line.
- Number of germinated seedlings per capitula, in parentheses germinated seeds per total number of seeds divided by 100.

**APO - 35S::hpHFIE**

<table>
<thead>
<tr>
<th>(%/2)</th>
<th>(%/3%)</th>
<th>%</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>3.8</td>
<td>8.6</td>
<td>12</td>
<td>45%</td>
<td>78%</td>
</tr>
<tr>
<td>1.5</td>
<td>9.2</td>
<td>2.2</td>
<td>2.4</td>
<td></td>
<td>10%</td>
</tr>
<tr>
<td>3.8</td>
<td>19.6</td>
<td>2.4</td>
<td>4.3</td>
<td></td>
<td>31%</td>
</tr>
<tr>
<td>4.0</td>
<td>28.6</td>
<td>3.0</td>
<td>4.3</td>
<td></td>
<td>38%</td>
</tr>
<tr>
<td>25.7</td>
<td>60.9</td>
<td>0.2</td>
<td>0.3</td>
<td></td>
<td>0%</td>
</tr>
</tbody>
</table>

**APO - AtMEA::hpHFIE**

<table>
<thead>
<tr>
<th>(%/2)</th>
<th>(%/3%)</th>
<th>%</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>3.8</td>
<td>8.6</td>
<td>12</td>
<td>45%</td>
<td>78%</td>
</tr>
<tr>
<td>1.5</td>
<td>9.2</td>
<td>2.2</td>
<td>2.4</td>
<td></td>
<td>10%</td>
</tr>
<tr>
<td>3.8</td>
<td>19.6</td>
<td>2.4</td>
<td>4.3</td>
<td></td>
<td>31%</td>
</tr>
<tr>
<td>4.0</td>
<td>28.6</td>
<td>3.0</td>
<td>4.3</td>
<td></td>
<td>38%</td>
</tr>
<tr>
<td>25.7</td>
<td>60.9</td>
<td>0.2</td>
<td>0.3</td>
<td></td>
<td>0%</td>
</tr>
</tbody>
</table>

**SEX - AtMEA::hpHFIE**

<table>
<thead>
<tr>
<th>(%/2)</th>
<th>(%/3%)</th>
<th>%</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>3.8</td>
<td>8.6</td>
<td>12</td>
<td>45%</td>
<td>78%</td>
</tr>
<tr>
<td>1.5</td>
<td>9.2</td>
<td>2.2</td>
<td>2.4</td>
<td></td>
<td>10%</td>
</tr>
<tr>
<td>3.8</td>
<td>19.6</td>
<td>2.4</td>
<td>4.3</td>
<td></td>
<td>31%</td>
</tr>
<tr>
<td>4.0</td>
<td>28.6</td>
<td>3.0</td>
<td>4.3</td>
<td></td>
<td>38%</td>
</tr>
<tr>
<td>25.7</td>
<td>60.9</td>
<td>0.2</td>
<td>0.3</td>
<td></td>
<td>0%</td>
</tr>
</tbody>
</table>

**HFIE Transgenic Plants**

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Seeds</td>
<td>% Germination</td>
<td>Field Black (FB)</td>
<td>Hfle</td>
<td>Copy no.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>3%</td>
<td>4%</td>
<td>6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31%</td>
<td>54%</td>
<td>65%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>62%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
required for the initiation and progression of embryo and endosperm development in apomorphic Hieracium.

Expression of HFIE in vegetative and inflorescence tissues is essential for embryo sac formation and autonomous seed initiation

HFIE is widely expressed in tissues of Hieracium plants (Figure 1). To examine if HFIE plays a role in vegetative and floral growth and development in Hieracium, the CaMV35S promoter was used to down-regulate HFIE expression in apomorphic plants via RNAi-mediated silencing. The CaMV35S promoter is known to be active in vegetative and floral tissues, but it is not active during all stages of ovary and embryo sac development (Koltunow et al., 2001). Three independent 35S::hpHFIE plants were recovered with down-regulated HFIE expression in leaves (Figure 3). The inflorescence internodes failed to elongate resulting in a stunted appearance relative to controls (Figure 3A). Capitulum form and floret number appeared normal, the florets were pollen fertile (Figure 3A). The leaves were generally smaller and curled in growth relative to untransformed control plants (Figure 3B). Curled leaf phenotypes are also observed in Arabidopsis when mutations occur in the SET-domain gene CURLY LEAF (CLF) or expression of its interacting partner AtFIE is reduced in vegetative tissues (Goodrich et al., 1997; Katz et al., 2004). HFIE from sexual and apomorphic Hieracium were also found to interact with AtCLF using yeast 2-hybrid analysis (Figure 4A). Thus the curled leaf phenotype in these transgenic Hieracium plants may relate to reduced levels of HFIE protein available for appropriate interaction with a putative Hieracium CLF orthologue in leaf tissues.
Figure 3C shows representative RT-PCR analysis of one of the transgenic plants and HFIE expression was absent in leaves and also in the floral disc of the receptacle where the florets attach via the ovary. HFIE was not significantly down-regulated in ovaries but the transgenic plants produced low levels of viable seed (Table 2). Cytological analysis of two lines showed that ovule and embryo sac development was defective from an early ovule stage in these transgenic plants (Figures 3D-F). In contrast to the control (Figure 3D), initiation of apomixis was poor and early necrosis was often observed in sexual and apomictic cell types (Figure 3E) with extensive embryo sac collapse in later stages (Figure 3F) that corresponded with poor autonomous seed initiation and high seed sterility (Table 2).

Even though HFIE was not appreciably down-regulated in ovaries general down-regulation of HFIE expression in the floral disk, leaves and inflorescence tissues lead to defects in embryo sac formation and autonomous seed initiation. Appropriate levels of HFIE in tissues subtending the ovule may therefore be required in signalling cascades that convey information or nutrient to the ovule to enable embryo sac formation and autonomous seed initiation in apomictic Hieracium.

HFIE does not interact with FIS complex proteins in yeast 2-hybrid assays

Yeast two-hybrid assays were also used to examine whether HFIE from sexual and apomictic plants was able to participate in the formation of a complex with other FIS members. In control experiments, AtFIE interacted as previously reported with AtMSI1, At RBR and AtFIS1 (Figure 4A). AtFIE also interacted with HRBR from apomictic and sexual plants and HMSI1 from apomictic plants (Figure 4A). By contrast, HFIE from both apomictic and sexual plants failed to interact with any of the
Figure 3. Analysis of 35S:HFIE:RNAi lines. (A) Transgenic lines (right) are stunted, with shorter internode lengths compared with WT apomict Hieracium (left); (B) Leaves of 35S::hpHFIE are generally shorter with curled tips (vertical bar = 1cm); (C) analysis of HFIE expression by RT-PCR showing that expression is silenced in leaves and in the floral disc but not the ovary. The floral disc is a structure that supports the growth of individual florets through attachment of the ovary (oy) to the Hieracium capitula, as shown in the schematic diagram. Equal loading of cDNA samples was verified according to expression of β-tubulin (β-TUB); (D) early ovule formation in D3 WT. Megaspore mother cell is outlined with black dotted line and aposporous initial cell with white dotted line; (E) ovule formation in 35S::hpHFIE transgenic line #7, showing degeneration of megaspore mother cell and no aposporous initials; (F) collapsed endothelium of an aborted embryo sac of 35::hpHFIE line #7. Bars in D-F = 20um.
The lack of HFIE interaction with AtFIS1/MEDEA is not surprising given that AtFIS1 originated from a recent genomic duplication event specific to the Brassicaceae lineage (Spillane et al., 2007). As HFIE can partner with another SET domain protein, AtCLF, further analysis for ovary-expressed SET domain containing proteins for possible interacting partners could be informative. Collectively, these data together with the observed absence of autonomous endosperm proliferation in HFIE down-regulated sexual plants support the conclusion that HFIE does not function to repress central cell proliferation and is unlikely to participate in an AtFIS-like repressive complex.

AtFIE and HFIE proteins differ in surface protein loops and charge which may alter protein function

Introduction of HFIE to Atfie-2 only rescued central cell proliferation and post-fertilization embryo lethality was retained. A possible explanation for this partial complementation given that HFIE fails to interact with other AtFIS proteins in yeast 2-hybrid assays is that HFIE recruits alternative proteins that function to repress central cell proliferation. This hypothesis implies that HFIE and AtFIE differ in protein structure. Significant conservation is evident in the WD-40 protein domains of both proteins with the majority of differences evident in the intervening regions (Figure 5; Supplementary Figure 1). Site directed mutagenesis in the Drosophila
Figure 4. (A) FIE interaction in yeast two hybrid assay. FIE protein sequences from Arabidopsis (AtFIE) Hieracium (HFIE) were used as bait and tested with Arabidopsis and Hieracium RETINOBLASTOMA RELATED PROTEIN (RBR), MULTICOPY SUPPRESSOR OF IRA1 (AtMSII and HMSII), MEDEA (AtMEA) and CURLY LEAF (AtCLF) from Arabidopsis. Interaction is observed by growth in selective medium under inducing conditions without uracil, histidine, tryptophan and leucine (-UHTL). (B) Western blot analysis of yeast protein extracts transformed with AtFIE- (A), HFIE-lexA (H) fusion proteins and empty lexA vector (-) detected with a monoclonal anti-LexA antibody. A 67kD peptide was detected in yeast transformed with AtFIE- and HFIE-lexA vector, corresponding to 43kD of HFIE protein and 25kD of lexA; whereas a 25kD peptide was detected in protein extracts of yeast transformed with empty pLexA vector; (C) Schematic diagram summarizing the results of the yeast-two-hybrid assay. AtFIE can interact with MSII and RBR from Arabidopsis and Hieracium and the Arabidopsis MEA protein, all of which are present in the female gametophyte, whereas HFIE is not capable of forming these interactions. Both AtFIE and HFIE interact with AtCLF, involved in vegetative growth.
orthologue of FIE, EXTRA SEX COMBS (ESC) has shown that these regions play critical roles in ESC protein function (Sathe and Harte, 1995; Ng et al., 1997).

Structural models of the AtFIE and HFIE proteins were produced and compared to examine if there were differences in the tertiary structure that might influence protein-protein interactions. Modelling was conducted with HFIE sequence isolated from the apomict plant because the single amino acid change in the sexual HFIE sequence was conservative with respect to protein structure and resulted from a change in the third (wobble) residue of the codon. The crystal structure of human beta-trcp1-skp1-beta-catenin protein 1p22 (chain A; lp22:A) was used as the template for the molecular modelling of the AtFIE and HFIE proteins.

Figure 5A and 5B show front and side views, respectively of superimposed models of the AtFIE, HFIE and 1p22:A proteins. All proteins have highly conserved β-sheets folded into a seven bladed propeller structure. The major structural differences stem from the position of protein loop elements that connect each blade of the propeller and protrude from the propeller surface, providing interacting sites on the surface of the protein. Two conspicuous differences in loop positions between AtFIE and HFIE are due to substitutions of nucleophilic residues in AtFIE (Asp-Asp) replaced by uncharged polar residues (Gly-Ser) in HFIE which correspond to amino acid residue insertions 11 and 19-22 in AtFIE and HFIE, respectively (Figure 5A and B). Figures 5C-5E show the calculated comparative distribution of surface electrostatic potentials on the surface of the β-propeller of lp22:A, AtFIE and HFIE, respectively. The HFIE model (Figure 5E) displays extensive acidic, positively charged regions on its surface indicated in red. We conclude that the observed differences in protein loop structures and charge on the surface of AtFIE and HFIE are the most likely features to influence protein interactions. The HFIE and AtFIE
Figure 5. Three-dimensional molecular models of AtFIE and HFIE. (A) Superposition of the AtFIE (cyan) or HFIE (magenta) 3D models on the template crystal structure of beta-trcp1-skp1-beta-catenin protein 1p22:A (yellow) over the Cα backbone positions showing a beta-propeller fold with secondary structural elements. The rmsd values were 0.76 Å and 0.64 Å for 286 and 283 amino acid residues, superposed in the Cα carbon positions, of 1p22:A and AtFIE or HFIE, respectively. The left-hand-side and right-hand-side arrows indicate NH2- and COOH-termini, respectively. (B) The side view of the AtFIE or HFIE 3D models shows different loop dispositions in both structures, (the corresponding regions in the primary sequence are shown in alignment in Appendix 1 indicated by asterisk) while the overall geometry of the beta-propellers is conserved in both cases. The models are rotated by approximately 90° degrees through the x-axis with respect to the view in panel A. The left-hand-side and right-hand-side arrows indicate NH2- and COOH-termini, respectively. (C) Charge distributions mapped onto molecular surfaces of the template human 1p22:A 3D structure and of the molecular models of AtFIE (D) or HFIE (E). Red and blue patches indicate electropositive (+5 kT e⁻¹) and electronegative (-5 kT e⁻¹) regions, respectively, and white patches indicate electroneutral regions. The orientations of structures are identical to their orientations in panel A.
models indicate regions and sequences to target in future site-directed mutagenesis to further examine in vivo protein interactions.

Discussion

HFIE is essential for viable seed formation in sexual and apomictic plants

Three AtFIS homologues were readily isolated from the ovaries of sexual and apomictic Hieracium in this study suggesting that PcG complexes similar to those found in Arabidopsis regulating central cell proliferation may have the potential to form in Hieracium and control seed development. HFIE was our primary focus for functional analyses because AtFIE interacts with many AtFIS proteins during seed development and also has roles in other aspects of plant development. HFIE is expressed in many tissues in sexual and apomictic Hieracium and down-regulation of HFIE confirmed that it has a role in the regulation of leaf morphology, and inflorescence growth. HFIE is also required for embryo sac formation in an indirect manner as down-regulation of HFIE in the floral disk subtending the ovary in addition to leaf and inflorescence tissues led to embryo sac abortion and defective seed initiation. HFIE may be involved in signalling cascades that convey information or nutrient from sources outside the ovule to enable embryo sac formation and subsequently seed initiation in Hieracium. Co-suppression of AtFIE in Arabidopsis leads to low seed fertility and floral defects (Katz et al., 2004) Whether this is related to direct or indirect effects of the down regulation of AtFIE in different cell types as occurs in Hieracium needs to be determined.
Sexual *Hieracium*, like Arabidopsis, requires fertilization to activate seed formation. HFIE is unlikely to form part of a repressive complex acting in either the egg or fused polar nuclei of sexual *Hieracium* because we failed to observe proliferation in either cell type when *HFIE* was specifically down regulated in the embryo sac. Repression of egg and central cell proliferation in sexual *Hieracium* is therefore likely to be achieved through an alternative repressive complex or possibly a modified FIS complex independent of HFIE function but involving other substituting FIS-like proteins.

It is apparent from the down regulation of *HFIE* in the embryo sacs of *Hieracium* that it is required early for the successful initiation embryo and endosperm formation in both sexual and apomictic plants. Cross-pollination of *HFIE* down-regulated sexual plants was required to activate seed initiation yet embryo abortion occurred at an early globular stage and endosperm failed to cellularize. Similarly autonomous embryo and endosperm formation was significantly reduced in apomictic AtMEA::hpHFIE lines and the arrested seeds resembled those of cross-pollinated HFIE down-regulated sexual AtMEA::hpHFIE lines. Interestingly endosperm cellularization was observed in seeds of the apomictic down-regulated HFIE plant after cross-pollination with another apomict. Endosperm cellularization does not occur during either fertilization-induced or autonomous central cell proliferation in *Atfie* mutants. Embryo development in fertilized *Atfie* mutants, progresses to the heart stage before embryo arrest (Ohad et al., 1996; Vinkenoog et al., 2000). FIE is therefore likely to be involved in the transition from the syncitial to the cellular phase of endosperm development in *Hieracium* and *Arabidopsis*. Importantly the role of FIE in controlling embryo and endosperm cell proliferation and growth after the activation of seed initiation appears to be conserved in sexual and apomictic *Hieracium*. 
However, HFIE is not recruited to participate in repressive complexes prior to seed initiation in Hieracium.

**Parental genome ratios, PcG gene function and seed viability in sexual and apomictic plants**

Inter-ploidy crosses in maize have shown that functional endosperm only develops with a 2maternal (m):1paternal (p) ratio and all others lead to seed abortion (Lin 1982, 1984). There is no paternal genome contribution to either the embryo or endosperm during seed initiation in apomictic Hieracium. The autonomous embryo has a 2m:0p ratio in the embryo compared to the 1m:1p ratio observed in sexual plants. We observed that the polar nuclei fused prior to autonomous endosperm proliferation in apomictic Hieracium, therefore, autonomous endosperm has an unbalanced 4 maternal(m):0 paternal (p) endosperm ratio that deviates from the normal 2m:1p ratio found in sexual plants and yet viable seeds are produced.

Apomictic Tripsacum dactyloides and Paspalum notatum are pseudogamous and require fertilization of the central cell to initiate endosperm formation. Like Hieracium these plants are insensitive to parental endosperm balance deviations and a surprisingly wide variety of m:p ratios support functional endosperm development in these plants (Grimanelli et al., 1977, Quarin, 1999). In the case of Paspalum, this tolerance for unbalanced endosperm is apomixis-specific and is not observed in sexual species (Quarin 1999). Other pseudogamous apomicts are sensitive to endosperm balance and specific developmental processes such as embryo sac development, pollen development and fertilization have been modified and a 2m:1p ratio is retained in the endosperm for viable seed formation (Koltunow and Grossniklaus, 2003).
In Arabidopsis, the requirement for the 2m:1p genome ratio in the endosperm for seed viability is not rigid. Deviations from the 2m:1p endosperm ratio influence seed size (Scott et al., 1998). If egg cell-specific fertilization occurs in Arabidopsis and the polar nuclei are not fertilized, autonomous endosperm proliferates indicating a positive signal from the fertilized egg can set off endosperm proliferation. The seeds are non-viable (Nowack et al., 2006). Small, viable Arabidopsis seeds form when only the egg cell is fertilized in fis mutant backgrounds where for example, either FIE, FIS2 or FIS1/MEA function is defective (Nowack et al., 2007). These data together with the observed rescue of seed viability when fis mutants are pollinated with pollen from a hypomethylated plant (Luo et al., 2000; Vinkenoog et al., 2000) suggest that the FIS genes are not strictly required for embryo and endosperm growth and viability in Arabidopsis after seed initiation. This contrasts with the observed requirement for functional HFIE for seed development in Hieracium. However, the observation that maternally derived endosperm with 2m:0p endosperm ratio can support fertilization-stimulated embryo growth in Arabidopsis if the FIS complex is non-functional is surprising and mimics in many respects the autonomous seed initiation in Hieracium except that here embryo development in autonomous (Nowack et al., 2007).

The requirement for a particular m:p genome ratio in the endosperm for seed viability appears to vary in both sexual and apomictic species. Moreover, it also appears from the data presented here that FIE is differentially recruited in plant species and has different functional impacts on seed growth and viability at least in Arabidopsis and Hieracium. Flexibility in PcG gene function and/or different combinations of these proteins in complexes might influence the sensitivity of seeds to deviations in m:p genome ratios in the endosperm. Further analysis of the function of these genes during seed development in other sexual species and apomicts is
required. Analysis of FIS function in maize would be instructive given the strict requirements for a 2m:1p genome ratio in the endosperm for seed viability and because a SET-domain homologue similar to FIS1/MEA has been isolated from maize and shown to be regulated by genomic imprinting (Haun et al., 2007).

**Models for seed activation in sexual and apomictic plants**

It has been demonstrated that in Arabidopsis the FIS repressive complex acts on the central cell to repress proliferation and an alternative repressive complex inhibits the proliferation of the egg cell (Berger et al., 2006). The signalling cascades occurring in the ovule upon fertilization are not known, nor are all of the targets of fertilization products that activate seed development. However, double fertilization disables repression and seed development occurs with appropriate paternal genome contribution. Sexual *Hieracium* and *Paspalum* exhibit central cell and egg arrest at embryo sac maturity therefore like Arabidopsis may contain repressive complexes restricting development until fertilization occurs. In sexual *Hieracium* HFIE is not part of a seed repressive complex. In Arabidopsis, a fertilized egg can provide signals that bypass FIS repression of the central cell enabling proliferation in the absence of a paternal genome contribution and formation of a viable seed only if the FIS complex is disabled. FIS genes are not strictly required for seed viability as their requirement can be bypassed presumably by the activation of other targets.

In apomictic *Hieracium* the *LOSS OF PARTHENOGENESIS LOCUS (LOP)* enables the initiation of autonomous embryo and endosperm development (Catanach et al., 2006). Extending some of the observations in Arabidopsis, we propose that LOP might function to stimulate a signal transduction cascade that activates
autonomous embryogenesis which may in turn stimulate endosperm proliferation. Alternatively, LOP function may co-ordinately but individually activate both embryo and endosperm development. Both possibilities are consistent with the relatively coordinated timing of embryo and endosperm formation observed during autonomous seed development.

In pseudogamous apomicts like *Tripsacum*, *Paspalum* and *Brachiaria* embryogenesis is autonomous but central cell proliferation is blocked and requires fertilization to activate proliferation. Fertilization of the central cell might be required to activate endosperm development because the egg signal is non-functional or cannot effectively disable a central cell repressive complex to allow endosperm formation and viable seed development. Functional analysis of FIS genes and complexes formed in pseudogamous apomicts seems a logical next step towards testing these activation hypotheses and also for further examination of internal signalling between egg and central cells in apomicts.

**Methods**

**Plant material and general procedures**

A triploid accession of apomictic *Hieracium piloselloides* Vill. (D3; 3x=2n=27) was used and a tetraploid sexual biotype of *H.pilosella* L (4x=2n=36). These plants were maintained in culture as stocks from individual characterised source plants and the developmental stages of capitula for gametophyte and seed analysis are as described previously (Koltunow et al 1998). Reproductive pathways of sexual and apomictic seeds were determined by ploidy analysis using Flow Cytometry of Single Seeds (FCSS) as described in Matzk et al. (2000). Plant transformation was carried out as
described by Bicknell and Bost (1994) and selected transgenic were also vegetatively propagated. Seed viability was determined by examining the number of filled black (FB) seeds in typically 5 capitula and then by counting the number germinating on agar plates 15 days after plating on the medium (Koltunow et al., 1998a). Results were usually expressed as FB seeds per capitulum or percentage viability per capitulum. DNA and total RNA was isolated from tissues as described in Tucker et al., 2000)

Cytology

To clear ovules, tissue was fixed and treated to remove oxalate crystals as described in Koltunow et al (1998) dehydrated to 100% ethanol and cleared in methylsalicylate as follows ETOH: methylsalicylate (1:1 v/v; 30 min), methylsalicylate:ETOH 3:1 (30 min) and finally in two change of 100% methylsalicylate (30 min). Tissue was also fixed and embedded in Spurrs’ resin (Spurr, 1969) for semi-thin sectioning as described previously (Koltunow et al., 1998a,b). Cleared ovules were observed by Nomarski differential interference contrast (DIC) under Zeiss Axioplan microscope and images were capture with a Spot II camera (Diagnostic Instruments, Stirling Heights, MI). Sections were observed using bright field and/or DIC microscopy after staining in 0.1% toluidine blue in 0.02% sodium carbonate and photographed. In situ hybridisation was carried out as described by Okada et al. (2007).

FIE MSI1 and RBR cDNA isolation
FIE, MSI1 and RBR were isolated using degenerate PCR oligonucleotides to amplify related sequences from stage ovary cDNA of apomictic and sexual Hieracium plants and full length sequences were obtained by RACE (Invitrogen) followed by sequencing. Degenerate primers were as follows (5’-3’):

FIS3F433 (5’-GATGARGATAAGG-3’); FIS3R656-
CATYCCACARICKAACFIS3F459- ACACDSTRAGTTGGGC and FIS3R636-GATTSATYYTTGTTDGCAG; MSI1F1–
TGGAATAATAAYACTCCTTTTTYTCTAYGA; MSI1R1–
GCNYTNARGTGGCCNTCNYTNA; MSI1F2–
TADATRTTYTCNGCCATYTGCCA; MSI1R2–
TCNGGGGNCRCRTCYTCNGCRTC; RBRF1-
TTYTTAARGARYTNCCNCARTT; RBRF2-GGNYTAGTNCTCNRTHTTHHC;
RBRR1 – AGYTCRTTNGANGGGGCTCTG; RBRR2–
CKNARNGGAGANACRTANAC.

**FIE promoter isolation**

Libraries for promoter walking were constructed during this study following the method of Siebert et al., (1995). Five libraries were constructed for both P4 H. pilosella and D3 H. piloselloides from Scal, Dral, PvuII, PmlI and EcoRV gGene specific primers for HFIIE were PrFIE3 (5’-ACCGCGACAGTGGCGAAGACATT3’) for the first reaction, PrFIE2 (5’-AATGGGCGCTTGCCCTCCTGTAAGAGACATT-3’) for the second reaction and PrFIE1 (5’-CCTCCTGTGACCTGCTGGGTGACT-3’) for the third reaction. Primary PCR reactions were conducted in 50 μl volumes containing 1 μl of the library, 200 μM dNTPs, 1 mM MgSO4, 60 mM Tris-SO4 (pH 8.9), 18 mM (NH4)2SO4, 0.4 μM adapter primer AP1 and gene specific primer, and 0.3 μL (~1.5
units) of High Fidelity Platinum Taq DNA Polymerase enzyme mix (Invitrogen). The cycle parameters were as follows: initial denaturation step at 94°C for 1 min, followed by 35 cycles of denaturation at 94°C for 30 s and annealing/extension at 68°C for 6 min, and a final annealing/extension time of 15 min. PCR products were cloned into pGEM T-easy (Promega).

Complementation analysis

Arabidopsis plants were grown in chambers under 16 hr light/8 hr dark photoperiods generated by fluorescent lighting. L.er fis3-2 seeds were provided by A. Chaudhury and M. Luo (CSIRO, Canberra) and contain the same mutation as the fie-2 allele described by Ohad et al (1999). Germinated seedlings were transferred to soil after the four-leaf stage, and after flowering, 50% of the seeds in siliques from some plants had aborted after fertilisation. Heterozygous fie-2/+ plants were identified by screening segregating populations with a CAPS PCR marker. The fie-2 mutation causes an A → G mutation at the border of intron 3 of the Arabidopsis FIE gene, resulting in the loss of a Hinfl site. PCR primers were designed to obtain a 241 bp product spanning this region, which incorporates three Hinfl sites. Digestion of the wild-type PCR product with Hinfl produced three bands of the expected size The loss of the Hinfl site in the mutant fie-2 gene results in only two bands after digestion However, because fie-2 plants are maintained as heterozygotes, plants containing the mutant allele are identified as those containing all four of the 156 bp, 85 bp, 66 bp and 19 bp bands.

A complementation construct was generated containing 2004 bp of the Arabidopsis FIS1 promoter (upstream from the ATG) fused to 1308 bp (1113 coding + 191bp 3'UTR) of D3-cHFIE in the pBINplus vector (van Engelen et al, 1995). This
binary plasmid was mobilised in Agrobacterium strain LBA4404 and was transformed by floral dip into Arabidopsis wild type Col-4 plants. Four lines were obtained, and lines #1 and #2 were utilised in subsequent experiments.

Heterozygous FIE/fie-2 (L.er) plants were crossed as males with female Col-4 WT plants and AtFIS1:HFIEa Col-4 lines #1 and #2. Plants containing both the fie-2 mutation and the AtFIS1:HFIEa construct were selected on Kanamycin and screened for by CAPS PCR analysis. To determine complementation, flowers from F2 plants were emasculated by removing the sepals, petals and anthers prior to anthesis. Siliques were then collected 3 days post emasculation (DPE) and fixed in FAA for 3 - 30 days at 4°C. The siliques were then stained using haemotoxylin and scored for the presence of multi-nucleate central cells (Ohad et al., 1996).

**RNAi constructs for down-regulation of HFIE and RT-PCR analysis**

An internal 442 bp internal fragment of the D3 HFIE cDNA spanning AA 140 to 288 in the HFIE protein sequence was amplified with appropriate adaptor restriction sites for cloning in sense and anti-sense orientation in pKANNIBAL (35S:pN6) (Wesley et al., 2001) to generate 35S:HFIE:RNAi. AtFIS1 promoter fragment (Luo te al, 2000; Tucker et al, 2003) was cloned in SacI and EcoRI sites substituting the 35S promoter to generate AtMEA::hpHFIE. The resulting constructs were cloned into pBIN+ binary vector and transformed into LBA4404 Agrobacterium. HFIE expression was analysed in total RNA extracted from ovules from stage 6 capitula by quantitative RT-PCR using ubiquitin as reference. Single-stranded cDNA was generated from 500 ng of total RNA template, an oligo(dT) primer and the thermoscript RT-PCR system.
(Invitrogen) as per manufacturer’s instructions. PCR reactions were performed in a 20μl total volume containing 1 μl (~25ng) of cDNA, 0.5 μM of each gene specific primer, 1 unit of RedTaq polymerase (Sigma), 2μl of 10× PCR Buffer (containing MgCl2) and 50 μM of each dNTP. Primers used for the RT-PCR reaction were BtubRTfwd (5'-GGGTGCTGGAAACATTGGGCTAA-3') and BtubRTrev (5'-ACTGCTCACTCACGCGCTAA-3'), RTHFIEF3 (5'-CCGTGTCATTGTAGCTGGCAAT-3') and RTHFIER2 (5'-GCAGCTGCATTGTAGCTGGAAATCA-3'). Primers were designed to intron-exon boundaries so as to exclude the possibility of genomic DNA contamination. Samples were equilibrated using the βtub primers as a control, electrophoresed, transferred to nylon membranes and probed with corresponding radioactively labelled cDNA clones to verify PCR-product identity. Down-regulation levels were calculated considering HFIE expression in untransformed control plant as 100%.

**Yeast-two-hybrid assay**

The Clontech yeast-two-hybrid system was used. FIE coding sequences were amplified, cloned into pGEM T-Easy (Promega) and sequenced to confirm ORF was intact. Full length cDNA was cloned in frame with the lexA DNA binding domain to create a fusion protein. RBR, MSI1, FIS1 and CLF cDNA sequences were cloned in frame into the pB42AD vector to create fusion proteins with the activator domain. Primer sequences with restriction sites are shown in Appendix 1. Constructs were tested for auto-activation under inducing conditions according to Clontech instructions. LexA fusion constructs were transformed into yeast strain EGY048 and selected with SD media without uracil and histidine. Positive yeast colonies were then
transformed with constructs containing the AD fusion proteins and selected in SD media without uracil, histidine and tryptophan. Positive clones were then plated onto growth selection SD media (gal/raf) without uracil, histidine, tryptophan and leucine. Growth in selection media indicated positive interaction. Simultaneously, colonies were also plated onto X-gal assay plates (SD gal/raf media without uracil, histidine and tryptophan) where interaction is observed by the production of blue colour (data not shown). Both the LEU marker gene and β-D-galactosidase genes were under the control of LexA operator. Plates were grown at 29oC for 4-5 days and photographed.

**Western Blot**

Yeast protein extracts were prepared from log-phase yeast cultures transformed with the appropriate vectors and grown under selection. Cells were harvested and re-suspended in Y-PER (Pierce) with 1x protease inhibitor cocktail (Roche). One freeze-thaw cycle (1 min in liquid N₂ / 1 min 37°C) was performed and extracts were centrifuged for 5min at 13000g. Supernatants were denatured in 2x loading buffer and separated in a 4-20% acrylamide gel (NuSep). Precision plus pre-stained standard was used (Biorad). Wet transfer to PVDF membrane carried out according to manufacturer’s (NuSep). The dilutions of primary mouse anti-LexA antibody (Dual Systems) and secondary anti mouse IgG conjugated with alkaline phosphatase were 1:2500 and 1:5000, respectively. Colour detection used the NBT-BCIP liquid substrate system (Sigma).

**Protein molecular modelling of AtFIE and HFIE**
Three-dimensional (3D) molecular models of Arabidopsis (AtFIE) and Hieracium FIE (HFIE) proteins were constructed using the Modeller 9v1 program (Sali and Blundell, 1993; Sanchez and Sali, 1998). To identify the most suitable template for the Arabidopsis (AtFIE) and HFIE proteins, searches were performed through the Structure Prediction Meta-Server (Ginalska et al., 2003) MetaPP server (Rost et al., 2004), SeqAlert (Bioinformatics and Biological Computing, Weizmann Institute of Science, Israel), Protein Data Bank (PDB) (Berman et al., 2000) and 3D-PSSM Server (Kelley et al., 2000). The best 3D template for modelling the AtFIE and HFIE sequences, based on amino acid sequence identity, was found to be human beta-trcp1-skp1-beta-catenin complex (chain A), accession number 1p22:A retrieved from the PDB (Berman et al., 2000) and called hereafter 1p22:A. BioManager, version 2.0 (http://www.angis.org.au/) at ANGIS (Australian National Genomic Information Service) was used to align 1p22:A with target sequences AtFIE and HFIE. The secondary structure predictions of AtFIE and HFIE sequences were performed with SAM T06 (Karplus et al., 1998) and with 3D-PSSM Server (Kelley et al., 2000). The positions of secondary structure elements and hydrophobic clusters was manually examined with hydrophobic cluster analysis software (Callebaut et al., 1997). The positional sequence identity and similarity scores were calculated by the Bestfit program from the ANGIS BioManager, version 2.0 (http://www.angis.org.au/), with the implemented gap penalty function and dynamic programming algorithm of Smith and Waterman (Smith et al., 1981). Only 293 amino acid residues of 1p22:A out of the total 435 amino acid residues were relevant for modelling and these covered approximately 90% of the HFIE and AtFIE protein sequences (Supplementary Figure 2). Structurally aligned 1p22:A and AtFIE or D3FIE sequences were used as input parameters to build 3D models on a Linux Red Hat workstation, running a Fedora
Linux Core 4 operating system. The final 3D molecular models of AtFIE or HFIE proteins were selected from 40 models with the lowest value of the Modeller objective function. The stereochemical quality and overall G-factors of the AtFIE or HFIE models were calculated with PROCHECK (Ramachandran et al., 1963; Laskowski et al., 1993). Z-score values for combined energy profiles were evaluated by Prosa2003 (Sippl, 1993); the plots were smoothed using a window size of 50 amino acid residues. Spatial superposition of 1p22:A and AtFIE structures and 1p22:A and HFIE structures were performed in the DeepView molecular browser (Guex and Peitsch, 1997), using a fragment alternate fit routine; the rmsd values were 0.76 Å and 0.64 Å for 286 and 283 amino acid residues, superposed in the Cα carbon positions, for 1p22:A/AtFIE or 1p22:A/HFIE structure combinations, respectively. The electrostatic potentials of the two models were calculated with Poisson-Boltzmann equation using GRASPv1.3.6 (Nicolls et al., 1991) and mapped onto the molecular surfaces generated with a probe radius of 1.4 Å. The molecular graphics were generated with the PyMol (http://www.pymol.org) and GRASP (Nicolls et al., 1991) software packages.
CHAPTER 3

SEGREGATION OF THE AUTONOMOUS SEED COMPONENT
IN APOMICTIC HIERACIUM
Acknowledgements and clarification of contributions made to this chapter

This section arose as part of a collaborative research effort to characterize reproduction in *Hieracium caespitosum* and a suite of gamma deletion mutants derived from it and involved Susan Johnson, Takashi Okada and Anna Koltunow. Ross Bicknell of Crop & Food lab in New Zealand provided the gamma deletion mutants and SCAR markers for the mutant and progeny analysis. Takashi Okada (TO) examined the presence or absence of the markers in the mutants and Susan Johnson (SDJ) and Anna Koltunow (AMK) were involved in the cytological analysis work in characterizing the apomictic parent plant and the mutant phenotypes. In this chapter, I will refer to unpublished data generated by my aforementioned colleagues and focus on the data I carried out to examine the inheritance of the embryo and endosperm components of autonomous seed development in apomictic *H. caespitosum*. All of the data will be combined in a forthcoming publication.
Abstract

*Hieracium* has been established as one model for the genetic and molecular analysis of apomixis in eudicotyledoneous plants. Three key events give rise to seed in Hieracium in that meiosis is avoided during embryo sac formation and the formation of the embryo and endosperm are fertilization-independent or autonomous. Autonomous seed initiation is rare among apomictic species. Recently, a gamma-deletion mapping approach was used to locate the genomic regions controlling apomixis and it was revealed that one major locus, Loss of Apomeiosis (LOA), controls apomixis initiation and another independent locus, Loss of Parthenogenesis (LOP), appears to control autonomous seed development, which was scored as parthenogenesis. However, due to the nature of the screen, it was not possible to determine if autonomous embryo and endosperm are linked or distinct genetic components. In this study, mutants that have lost one or both of the loci were analysed in greater detail. We show that loss of either LOA or LOP resulted in partial reversion to the sexual pathway and plants missing both loci were sexual indicating that apomixis is superimposed upon a default sexual pathway. Furthermore, progeny analysis enabled a better definition of the function of the LOA locus. We also established that autonomous embryo and endosperm development by LOP are either closely linked genetically or controlled by a single gene and are under maternal gametophytic control.
Introduction

Asexual seed production, termed apomixis, typically involves the bypass of meiotic reduction during embryo sac formation and the formation of an embryo independent of fertilization, resulting in a seed that is genetically identical to the maternal parent. In apomictic *Hieracium*, the formation of the female gametophyte occurs through apospory, where a somatic cell close to the MMC (aposporous initial – AI) undergoes mitosis and enters a gametophytic program to form the embryo sac (ES) (Asker and Jerling 1992). The unreduced ES contains an egg cell which develops autonomously to form an embryo and endosperm formation is also independent of fertilization (autonomous).

Autonomous endosperm formation is rare in apomicts but is found in species belonging to the Asteraceae family. Analysis of hybrids generated from crosses between sexual and apomictic Asteraceae species has shown that dominant genetic factors control the avoidance of meiosis and autonomous seed development (Noyes and Rieseberg 2000; van Dijk et al., 2003; Matzk et al., 2005; Noyes et al., 2007). In *Taraxacum* autonomous endosperm was shown to be unlinked to the locus controlling autonomous embryo development (van Dijk et al., 2003), while in *Erigeron*, these two components appear to be linked (Noyes et al., 2007).

*Hieracium* is another apomict in the Asteraceae family that forms embryo and endosperm autonomously but in contrast to *Erigeron* and *Taraxacum* forms ES via apospory (Koltunow et al., 1998a; Bicknell et al., 2000; Koltunow et al., 2000). Deletion mutagenesis and supportive genetic analysis in the tetraploid apomictic *Hieracium caespitosum* (C4d) has shown that apospory and parthenogenesis are controlled by two dominant but independent loci, Loss of Apomeiosis (LOA) and Loss of Parthenogenesis (LOP) (Catanach et al., 2006). LOA confers the ability of a
somatic cell to enlarge and enter a gametophytic program to produce an unreduced ES while LOP confers the ability to from a seed independent of fertilization. Aside from these two key steps in apomictic reproduction that bypass the sexual reproductive program, apomictic and sexual reproduction share gene expression programs and regulatory components, suggesting that apomixis de-regulates the expression of the sexual program in particular developmental stages and cell-types (Tucker, 2003). However, it is not clear if de-regulation caused by LOA and LOP activity leads to the expression of a modified sexual program or is limited to the modification of a trigger of the sexual pathway. In this chapter, we explore the relation between the sexual and apomictic pathway by analysing in greater detail the progeny of loa and lop mutants to determine the effect on the underlying reproductive pathway that persists as the result of the loss of either LOA and LOP.

In the previous chapter we showed that the Hieracium FIS-class gene *FERTILIZATION INDEPENDENT ENDOSPERM (HFIE)* is needed for autonomous seed initiation. Autonomous embryo development initiated and aborted early in HFIE down-regulated lines. Endosperm initiated at the same time but failed to cellularize and viable seeds were not produced. These data suggest that autonomous embryo development is linked to autonomous endosperm formation. We investigated this further by analysing the segregation of embryo and endosperm formation in a loa mutant that undergoes meiosis but retains the ability for autonomous seed initiation. If autonomous embryo and endosperm are two independent components in the LOP locus and not closely linked, segregation of these components might be observed. Here we demonstrate that these components co-segregate in the Hieracium caespitosum loa deletion mutant analysed, indicating that they are closely linked.
Results

Mutant 115 has deletions in LOA and LOP

Apomictic mutants were generated by gamma-irradiation on the tetraploid self-incompatible apomictic *Hieracium caespitosum* (C4d). As with other *Hieracium* apomicts, initiation occurs via the differentiation of an aposporous initial cell, which undergoes mitotic division to form an unreduced embryo sac. Seed initiation is completely autonomous as embryo and endosperm are independent of fertilization. Apomixis is facultative in C4d, however the majority of progeny are of maternal origin. Deletion mutagenesis resulted in deletions that varied in size with larger deletions increasing the possibility of affecting developmentally important genes unrelated to LOA or LOP function (Catanach et al., 2006; SDJ & AMK, unpublished result). For detailed progeny analysis, *loa* and *lop* mutants with the smallest deletions were chosen and this was based on the number of SCAR markers lost (Catanach et al., 2006).

Mutant 115 was initially identified as a *lop* mutant (Catanach et al., 2006); however marker analysis showed loss of markers in both loci (TO, unpublished results), indicating that it was likely to be a sexual plant. Seed analysis showed that the unpollinated mutant 115 produces only sterile brown (B) seeds (Table 1). Fertile seeds were only produced upon fertilization with mutant 115 producing 29% of filled black (FB) seeds and 12% of soft black (SB) seeds and the remainder (59%) were brown and sterile indicative of aborted ovules (Table 1). Most of FB seeds were able to germinate (86.4% of FB seeds; 24.2% total seeds). To confirm the reproductive origin of germinated progeny, seedlings were analysed by flow cytometry. DNA
content for all the progeny analysed was similar to the tetraploid controls C4D WT and the 115 mutant (Table 2). Therefore, mutant 115 only produces tetraploid progeny when cross-pollinated. Maternal progeny is also expected to be tetraploid, however, progeny derived from mutant 115 had characteristics of the pollen donor plant upon flowering (data not shown), therefore indicating that they are hybrids.

**Cytological analysis of loa/lop mutant 115**

Following meiosis in WT apomictic *H. caespitosum* (C4d), all four megaspores degenerate and nucellar cells surrounding the region of the selected spore differentiate, as shown in Figure 2A. At stage 6, the majority of C4d ovules contain developing embryos and endosperm (Figure 2B) and by stage 10, these autonomous embryos are at late globular to heart stage and endosperm cellularization initiates (Figure 2D). Multiple ES are common in C4d WT and these are derived from multiple AI cells (Figure 2C-D). Cytological analysis of mutant 115 showed that the capacity to form an AI cell and produce autonomous seeds was compromised (Figure 2M; Table 3). Unpollinated ovules at late stages of seed development showed only the presence of resting egg cells and central cell and embryos were not observed (Figure 2N). Embryo sac collapse was high in mutant 115 (Table 3) and at stage 10, 44% of ovules had a collapsed endothelium (Figure 2P) and this was similar to the frequency observed for WT C4d at a similar stage (38.5%; Table 3). Interestingly, there was a higher frequency of ovules with immature ES in mutant 115, a phenotype typical of loa mutants (Table 3). Cytological analysis of pollinated progeny showed that the proportion of ovules with egg cells and central cell decreased to 1.5%, whereas embryos were present in 28% of ovules observed.
Collectively, these data confirm that 115 is a sexual plant because LOA and LOP loci are absent (TO, unpublished results), aposporous initials do not form and meiotically-derived ES develop that are dependent on fertilization to form embryo and endosperm.

(Figure 2O), which closely correlated to the proportion of viable seeds (Table 1).

Table 1. Seed analysis based on morphology of mature seeds and seed viability of WT apomictic *H. caespitosum* and *loa* mutant 134, *lop* mutant 179 and *loa/lop* mutant 115.

<table>
<thead>
<tr>
<th>Crosses</th>
<th>Seed morphology</th>
<th>Seed viability$^1$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FB</td>
<td>SB</td>
</tr>
<tr>
<td>WT</td>
<td>70%</td>
<td>9%</td>
</tr>
<tr>
<td>115 UP</td>
<td>0</td>
<td>4.8%</td>
</tr>
<tr>
<td>(n=268; 6shd)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>115 x SEX</td>
<td>29%</td>
<td>12%</td>
</tr>
<tr>
<td>(n=124; 4shd)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>179 UP</td>
<td>0</td>
<td>1%</td>
</tr>
<tr>
<td>(n=277; 6 shd)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>179 x APO</td>
<td>36%</td>
<td>7%</td>
</tr>
<tr>
<td>(n=132; 3 shd)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>179 x SEX</td>
<td>32%</td>
<td>6%</td>
</tr>
<tr>
<td>(n=124; 3shd)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>134 UP</td>
<td>23%</td>
<td>19%</td>
</tr>
<tr>
<td>(n=307; 6shd)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>134 x APO</td>
<td>43%</td>
<td>15%</td>
</tr>
<tr>
<td>(n=303; 6shd)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>134 x SEX</td>
<td>52%</td>
<td>22%</td>
</tr>
<tr>
<td>(n=151; 4shd)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FB – filled black; SB – soft black; B – brown; G – germinated; NG – non-germinated
Seed analyses are presented as % per capitulum or seed head (shd);
$^1$seed viability was determined by plating FB seeds and counting germinated seedlings after 16-17 days, the number of germinated seedlings was divided by the total number of seeds, including SB and B seeds and multiplied by 100.
lop mutant 179 only produces viable progeny when cross-pollinated

lop mutant 179 only produces viable progeny when cross-pollinated (Table 1). This is expected from an apomictic mutant that has lost the LOP locus and thus is unable to produce autonomous seeds. Viable FB seeds were formed when 179 was cross-pollination with the sexual P4 and apomictic A4 plant (Table 1). Germinated seedlings derived from the cross with the apomictic pollen donor were analysed for relative DNA content. Of the 13 plants analysed, 11 were shown to be hexaploid (Table 2). Therefore, the majority of progeny originated from cross-pollinated mutant 179 were derived from the fertilization of unreduced tetraploid egg cell with a diploid male gamete. One of the progeny was tetraploid and was most likely derived from the fertilization of a reduced egg cell. This suggests that the facultative nature of the WT apomictic is retained in 179 and low levels of sexuality still persist. Another plant was heptaploid. The origin of this plant is not clear and it is possible that it might have originated from the fertilization of an unreduced egg with a triploid male gamete, to

<table>
<thead>
<tr>
<th>Ploidy(^1)</th>
<th>134 UP</th>
<th>134 x APO</th>
<th>134 x SEX</th>
<th>115 x SEX</th>
<th>179 x APO</th>
</tr>
</thead>
<tbody>
<tr>
<td>2n</td>
<td>26</td>
<td>12</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4n</td>
<td>0</td>
<td>27</td>
<td>21</td>
<td>21</td>
<td>1</td>
</tr>
<tr>
<td>5n</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>n+1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6n</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>7n</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>26(31)</td>
<td>39(54)</td>
<td>30(47)</td>
<td>21(30)</td>
<td>13(45)</td>
</tr>
</tbody>
</table>

\(^1\)Leaf tissue from viable FB seeds (Table 1) was used for flow cytometry (see Methods). Total number of seedlings analysed is followed by the total number of germinated seedlings (Table 1) in parentheses.

Table 2. Distribution of ploidy classes in progeny derived from cross-pollinated C4d deletion mutants.
increased ploidy levels. Consistent with this, pollen analysis carried out in the A4 paternal parent used in this experiment showed extensive variation in size among pollen grains suggesting possible meiotic defects (SDJ, unpublished results).

Cytology of *lop* mutant 179 confirms origin of viable progeny

Cytological analysis of mutant 179 in early stages of ovule development confirmed the ability to produce AI cells (Figure 2I). Embryo sacs in mutant 179 are able to form an egg and central cell; however, consistent with progeny analysis, autonomous embryo formation was not observed and egg and central cells remain arrested even at later stages of development (Figure 2J; Table 3). Cross-pollinated mutant 179 ovules showed that in 30% of pollinated ovules embryo formation had initiated (Figure 2L). A large proportion of cross-pollinated mutant 179 ovules still contain an egg and central cell and this is most likely due to the manual pollination process, pollen defects or ES irregularities (Figure 2K). The rate of ES collapse was high in 179 mutant ovules and was comparable to that observed in WT ovules (Table 3), Immature ES were not a predominant feature of mutant 179 ovules (Table 3). The presence of AI cells and the lack of autonomous seed initiation observed in mutant 179 is consistent with ploidy analysis and confirms that viable cross-pollinated mutant 179 progeny are derived from fertilized unreduced egg cells.

Seed viability is affected in progeny of *loa* mutant 134

Marker analysis showed that mutant 134 has a deletion in the LOA locus but retained the LOP locus (Catanach et al., 2006). The absence of a functional LOA
locus would result in segregation of LOP in unpollinated mutant 134 progeny. Therefore, a proportion of the progeny would be expected to arise via autonomous seed formation from reduced egg cells and be of reduced ploidy (diploid) compared to the maternal parent plant. On the other hand, ovules in which the LOP component was absent as a result of meiotic recombination would contain resting egg cells that are unable to initiate autonomous development. To confirm this, we analysed seed phenotype, seed viability and ploidy level of germinated seedlings. Analysis of seed morphology in unpollinated mutant 134 showed that 42% (n=307; Table 1) initiated seed development producing black achenes (FB+SB, Table 1), and the remainder were brown and sterile (Table 1). Germination of seeds collected from unpollinated 134 was low suggesting the action of a seed lethality component (Table 1).

Mutant 134 was cross-pollinated with apomictic and sexual plants to observe the effect on seed viability and the proportion of FB seeds increased (Table 1). More FB seeds formed when the sexual plant was used as the pollen donor (43% to 52%; Table 1). This was also reflected in the proportion of viable FB seeds, as there was a 2-fold increase in viable FB seeds resulting from the 134xAPO cross and a 3-fold increase in viable seeds resulting from the 134xSEX cross (Table 1). However, seed viability was significantly lower than the proportion of FB seeds produced and this suggests that seed initiation occurs but aborts resembling the observation in unpollinated 134 plants and suggesting an embryo lethal component might be acting to reduce seed viability in cross-pollinated plants.

**Ploidy analysis of mutant 134 progeny confirms reproductive origin**
Flow cytometry analysis of leaf tissue from germinated seedlings isolated from unpollinated mutant 134 showed that all of the viable progeny had half of the DNA content of the tetraploid mutant 134 the C4d WT plant (Table 2). These values are consistent with the plants originating from the autonomous development of a reduced egg cell. Ten of these plants were transferred to soil and all were clearly smaller in stature and flower size relative to both the C4d WT and mutant 134 plants (Figure 1A). All of these plants flowered normally but none produced fertile seeds (data not shown). Three capitula from each of these diploid plants were pollinated to assess if haploid eggs could be fertilized, however viable seeds were not recovered (data not shown). Diploid plants do not contain LOA and can undergo meiotic reduction to form haploid female gametes. The failure to produce FB seeds in unpollinated and cross-pollinated progeny from these diploid plants is most likely associated to lethality of female gametes in the haploid state.

Flow cytometry was performed on seedlings obtained after cross-pollination of 134 with apomictic and sexual pollen donors. Table 2 shows that the majority of viable seedlings were tetraploid, indicating that they originated from the fertilization of unreduced egg cells (Table 2). The plants were transferred to soil and phenotypic analysis demonstrated that tetraploid progeny had characteristics of the pollen donor
plant, confirming hybridity (Figure 1B-C). This contrasts with the findings of Catanach et al (2006) describing that loa mutants produce progeny derived from...
autonomous development of reduced egg cells (polyhaploid progeny; n + 0), even when cross-pollinated. In this study, diploid and tetraploid hybrid plants were

NOTE: This figure is included on page 97 of the print copy of the thesis held in the University of Adelaide Library.

Figure 2
**Figure 2.** Cytological analysis of *Hieracium caespitosum* C4d WT apomictic plants (A-D), *loa* mutant 134 (E-H), *lop* mutant 179 (I-L) and *loa/lop* mutant 115 (M-P). (A) In WT plants, apomixis initiates with the differentiation of aposporous initial (ai) cell (shaded white) that often occurs alongside the selected spore (shaded grey); (B) autonomous seed development occurs rapidly and by stage 6 early embryos (em) and nuclear endosperm (ne) are observed; (C) WT plants commonly display chalazal embryo sacs (ces) that occur alongside of a micropylar embryo sac that is collapsed (coll); (D) or contains embryos and endosperm. (E) *loa* mutant 134 do not produce ai cells, but the sexual apparatus is still present and produces 4 megaspores (ms) (shaded grey); (F) a common feature of *loa* mutant 134 is the presence of embryo sacs containing undifferentiated nuclei (arrowheads) that do not form embryo sacs even at late stages of development. *loa* mutant 134 ovules either remain arrested, with egg cells (ec) and fused polar nuclei (fpn) (G) or can develop autonomously to form embryo (em) and endosperm (ne) (H); (I) *lop* mutant 179 still retain the capacity for apomixis initiation and ai cells can be frequently observed (shaded white) along with the selected sexual spore (shaded grey) as in the WT apomictic plant; (J) *lop* mutant 179 ovule do not develop autonomously and even at stage 10, remain arrested with egg and central cell; (K) some mutant 179 ovules contain two egg cells; (L) upon cross-pollination (cp), mutant 179 ovules can form embryos and endosperm; (M) *loa/lop* mutant 115, as with *loa* mutant 134, do not have the ability to form ai cells and the sexual megaspores persist; (N) mutant 115 ovules do not develop autonomously and similar to *lop* mutant 179, ovules remain arrested with egg and central cell; (O) upon cross-pollination, mutant ovules can form embryos and endosperm; (P) a common feature of *loa* mutants is the failure of the embryo sac to expand, leading to collapse (coll). Bars = 50 μM in all panels except C, D, J, O and P (100μM). Pictures taken by SDJ.
not observed as described in WT sexual *Hieracium* plants (Figure 2E). In mutant 134, the proportion of ovules containing an ES with undifferentiated nuclei (Figure 2F) was high compared to C4d but then decreased by stage 10, indicating that these
undifferentiated ES nuclei were most likely developing immature ES (Table 3). Mutant 134 ovules also showed a reduced rate of embryo sac collapse relative to WT C4d plant at both stages analysed (Table 3). At stages 6 and 10, the frequency of ovules with egg cells and fused polar nucleus was similar to the proportion of ovules containing developing embryos and endosperm (Figure 2H), suggesting that LOP was segregating at an 1:1 ratio (Table 3). This is consistent of a locus under maternal gametophytic control (Howden et al., 1998). Parthenogenetic embryo formation was always accompanied by autonomous endosperm, further suggesting that when transmitted through the female gametophyte, these two components are closely linked. These cytological observations are in agreement with the recovery of diploid and tetraploid progeny from cross-pollinated 134.

**Mutant 134 progeny form irregular embryos**

Although autonomous seed initiation occurred in 39% of stage 10 unpollinated mutant ovules, embryo and endosperm development were often not normal compared to C4d WT apomictic plants (Figure 3). In C4d WT plants, embryos found in capitula at stage 10 are typically late globular or heart stage with cellularized endosperm (Figure 3A). In unpollinated mutant 134, embryos with wild type appearance were observed (Figure 3B), but the majority of embryos were small and irregular and often contained cellularized endosperm that was cytologically similar to that found in the WT plant at the same stage (Figure 3C). Less frequently, embryos were at a torpedo-like stage, with a stunted appearance and endosperm development irregular with thick cell walls in embryo sacs that do not fully expand to occupy the entire ovule (Figure
3D). These three phenotypic classes of embryos occurred in different frequencies and the relative proportion of normal to irregular embryos was 1:2 (Table 4).

Table 3. Cytological analysis of Hieracium caespitosum WT apomict, unpollinated and developing seeds from cross-pollinated loa mutant 134, lop mutant 179 and loa/lop mutant 115

<table>
<thead>
<tr>
<th>Plant ID</th>
<th>Immature ES</th>
<th>Collapsed ES</th>
<th>EC + CC</th>
<th>EM + NE</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT s6 (n=153)</td>
<td>4.6%</td>
<td>43.7%</td>
<td>11.1%</td>
<td>37.3%</td>
</tr>
<tr>
<td>WT s10 (n=242)</td>
<td>2.9%</td>
<td>38.5%</td>
<td>1.2%</td>
<td>57.4%</td>
</tr>
<tr>
<td>134 UP s6 (n=118)</td>
<td>36%</td>
<td>8%</td>
<td>28%</td>
<td>28%</td>
</tr>
<tr>
<td>134 UP s10 (n=100)</td>
<td>12%</td>
<td>10%</td>
<td>36%</td>
<td>39%</td>
</tr>
<tr>
<td>134 x APO s10 (n=99)</td>
<td>10%</td>
<td>8%</td>
<td>18%</td>
<td>57%</td>
</tr>
<tr>
<td>134 x SEX (n=126)</td>
<td>9%</td>
<td>8%</td>
<td>17%</td>
<td>63%</td>
</tr>
<tr>
<td>179 UP s6 (n=72)</td>
<td>14%</td>
<td>57%</td>
<td>21%</td>
<td>0</td>
</tr>
<tr>
<td>179 UP s10 (n=75)</td>
<td>5%</td>
<td>38.5%</td>
<td>38.5%</td>
<td>0</td>
</tr>
<tr>
<td>179 x APO s10 (n=134)</td>
<td>10%</td>
<td>34%</td>
<td>20%</td>
<td>30%</td>
</tr>
<tr>
<td>179 x SEX s10 (n=120)</td>
<td>5%</td>
<td>51%</td>
<td>15%</td>
<td>30%</td>
</tr>
<tr>
<td>115 UP s6 (n=133)</td>
<td>48.2%</td>
<td>29.3%</td>
<td>21.1%</td>
<td>0</td>
</tr>
<tr>
<td>115 UP s10 (n=162)</td>
<td>23.5%</td>
<td>44.4%</td>
<td>31.5%</td>
<td>0</td>
</tr>
<tr>
<td>115 x SEX s10 (n=133)</td>
<td>16.6%</td>
<td>44.4%</td>
<td>1.5%</td>
<td>27.8%</td>
</tr>
</tbody>
</table>

ES – embryo sac; EC – egg cell; CC – central cell; em – embryo; NE – nuclear endosperm; UP – unpollinated; APO – tetraploid apomictic Hieracium aurantiacum (A4) used as pollen donor; SEX – tetraploid sexual H. pilosella used as pollen donor. Table assembled with cytology data

These observations suggest that in mutant 134, upon meiotic division of the megaspore and formation of an unreduced embryo sac, parthenogenetic embryo and endosperm formation is affected. Similar to unpollinated parthenogenetic embryos, embryos from pollinated mutant ovules also displayed similar abnormalities that also
affected seed viability. The distribution of embryo phenotypes was similar between unpollinated and cross-pollinated mutant progeny, including progeny from APO and SEX pollen donors and the proportion of normal embryos relative to abnormal embryo development was approximately 1:2 (Table 4).

Table 4. Embryo phenotypes of unpollinated (UP) and cross-pollinated (CP) mutant 134 progeny

<table>
<thead>
<tr>
<th>Embryo Phenotype</th>
<th>CP</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UP (n=39)</td>
<td>APO (n=57)</td>
<td>SEX (n=78)</td>
</tr>
<tr>
<td>Normal EM+NE</td>
<td>26%</td>
<td>32%</td>
<td>33%</td>
</tr>
<tr>
<td>Small IE+CE</td>
<td>59%</td>
<td>53%</td>
<td>56%</td>
</tr>
<tr>
<td>Torpedo-like+IEN</td>
<td>15%</td>
<td>15%</td>
<td>11%</td>
</tr>
</tbody>
</table>

Legends are as in Table 3. UP – unpollinated; CP – cross-pollinated with pollen donor from the tetraploid apomictic A4 (APO) plant and tetraploid sexual plant (SEX); IEN – irregular endosperm. Chi-square values for a 2:1 frequency of irregular and normal embryos: UP=2.46; 134xAPO=0.09; 134xSEX=0.015 (0.5<p<0.1)
**Figure 3.** Cytological analysis of stage 10 WT (A) and mutant 134 embryo phenotypes (B-D). (A) WT seeds generally contain late globular embryos (em) and cellularizing endosperm (ce). In mutant 134 unpollinated and cross-pollinated ovules, three embryo phenotypes are observed; (B) embryos of normal appearance, which are similar to WT embryos with cellular endosperm (ce); (C) more frequently, embryos were irregular and small (ie) and endosperm cellularization still occurred (ce); (D) some ovules contained irregular torpedo-like embryos (ie) and these seeds also contained irregular endosperm (ien) with thick cell walls.
The frequency of embryo and endosperm abnormalities in cross-pollinated progeny is agreement with seed viability (Table 1) and suggests that seeds with irregular embryos and endosperm fail to develop properly and germinate. There was a higher than expected proportion of embryos of normal appearance in unpollinated progeny when compared to the proportion of viable seeds. While the scoring of torpedo embryos with irregular endosperm nuclei was unmistakable, it is likely that other normal appearing embryos arrested later. The seed germination data confirms that embryo and endosperm irregularities in mutant 134 affects viability and appears to segregate as a dominant trait.

Discussion

Function of LOA might explain segregation distortion

Apomixis in Hieracium is controlled by two independent loci, LOA and LOP. In this study, analysis of the progeny derived from a loa and a lop mutant was carried out to confirm the progeny types and reproductive pathways used by mutant plants when either loci is lost. The loss of LOA in mutants 134 and 115 was clearly associated with the absence of an AI cell and the formation of an unreduced embryo sac. Furthermore, cytological evidence suggests that there is a high rate of embryo sac collapse that is associated with the presence of LOA. This feature was observed in both the WT apomictic parent Hieracium caespitosum as well as the lop mutant 179, which also contains an intact LOA locus, and absent in loa mutant 134. Embryo sac collapse associated with LOA function might suggest that LOA is associated with and affects genes that are essential for female gametophyte maturation. This feature was
not present in all loa mutants analysed, perhaps due to the larger size of the deletions in these mutants (SDJ & AMK, unpublished results). This effect of LOA on gametophytic genes might help explain the distortion in inheritance of LOA in a segregating Hieracium population (Bicknell et al., 2000; Catanach et al., 2006). As a dominant trait that segregates independently of LOP, LOA would be expected to be present in 50% of a segregating population derived from a cross between H. caespitosum as pollen parent and a sexual plant; however the presence of LOA was significantly reduced in this population (Catanach et al., 2006). The superimposition of LOA function could lead to mis-expression of genes involved in early ovule development. This in turn could result in a high rate of embryo sac failure and lead to reduced transmission of LOA in viable progeny. Although the effect of LOA on male and female gamete lethality is still unclear, it appears likely that LOA affects gamete formation causing segregation distortion. Transmission distortion was also encountered in the genetic element that control apospory in Pennisetum (Roche et al., 2001), Paspalum (Martínez et al., 2001) as well as diplospory in Tripsacum (Grimanelli et al., 1998). In these cases, transmission through female gametes was strongly selected against.

**LOA contains a suppressor of megasporogenesis**

Cytological evidence of mutant 134 and 115 shows that upon the loss LOA, the ability to produce an aposporous initial cell is compromised and meiosis of the megaspore mother cell (MMC) and the formation of a reduced embryo sac occur normally, much like a sexual plant. This suggests that the LOA locus not only induces the differentiation of the AI cell that gives rise to the unreduced embryo sac, but
might also be associated with or contain a suppressor of the sexual process. In support of this, cytological analysis of other loa mutants confirm that the sexual process occurs normally as in mutant 134 (SDJ & AMK, unpublished results). The common phenotype of loa mutants strongly suggests that AI differentiation can directly or indirectly suppress the sexual process, possibly by promoting cell death of all meiotic spores (SDJ & AMK, unpublished results). When LOA is transmitted through pollen in cross-pollinated 134 progeny, it not only restores the ability to produce AI cells, but also to prevent the survival of the sexual megaspore, suggesting that this activity co-segregates with the LOA locus (SDJ, unpublished results). Demise of the sexual process can occur in other aposporic apomicts. As with Hieracium, the differentiation of the AI cell occurs at the time of MMC differentiation, meiosis and megaspore selection and this often associated with the degeneration of all four sexual megaspores (Asker and Jerling 1992). This might suggest that apomixis initiation could depend on the initiation of the sexual process and its subsequent replacement (Bicknell and Koltunow 2004). However, the presence of sexual and aposporous embryo sacs within the same ovule can occur in some aposporic grass species, such as Brachiaria, although aposporic embryo sacs are only formed after the sexual process has initiated. This indicates that while the initiation of the sexual pathway is most likely an essential feature of aposporic apomicts; the demise of the sexual pathway may not be an essential mechanism for aposporous embryo sac formation. The interdependence of the sexual and apomictic processes at the time of initiation is still unclear and requires further investigation.
Deletions in LOA and LOP restore the sexual pathway in apomictic *Hieracium caespitosum*

Deletions in LOA restored normal meiotic capacity and formation of reduced embryo sacs that have the capacity to produce parthenogenetic embryos. Conversely, a deletion in LOP did not affect apomixis initiation and the formation of unreduced embryo sacs, but prevented autonomous seed initiation, and seed formation was dependent upon fertilization. These observations directly show that upon removal of activity conferred by LOA and LOP, the sexual program is restored and proceeds normally. Consistent with this, mutant 115, which has deletions in both LOA and LOP, behaves much like a sexual plant, producing reduced embryo sacs that require fertilization. Mechanistically, this implies that both LOA and LOP act to superimpose the differentiation of an AI cell and autonomous seed initiation, respectively, upon the sexual pathway. The molecular mechanism in which LOA and LOP alters the sexual program is still unclear.

**Autonomous embryo and endosperm are closely linked in *Hieracium***

In mutant 134, the small deletion in LOA created an ideal model to analyze segregating progeny, since it represents an apomictic mutant in a nearly isogenic background. Autonomous seed initiation in unpollinated 134 only occurs in 50% of ovules and autonomous embryo formation was always accompanied by autonomous endosperm. This suggests that these two components of apomixis are linked in *Hieracium* and under maternal gametophytic control in this mutant plant. The maternal gametophytic segregation ratio of autonomous seed initiation was not
observed in the progeny of other loa mutants that were analysed, however, this could 
be due to other ovule defects that were observed in these mutants and that are 
associated with larger deletion sizes (SDJ & AMK, unpublished results).

The linkage of the autonomous seed component (embryo and endosperm) was 
also observed in other species of Asteraceae, Erigeron (Noyes et al., 2007). Contrary 
to Hieracium, Erigeron and Taraxacum are diplosporous but differ regarding the 
linkage of the autonomous endosperm component. This might suggest that there are 
two genes that control autonomous embryo and endosperm, respectively. In support of 
this, a hybrid plant derived from the 134xAPO cross was recovered in which 
autonomous endosperm development occurs in the absence of autonomous embryos 
(SDJ, unpublished results). It may be that in the plant used as pollen donor, H. 
aurantiacum, there are two genes that are not tightly linked and can be separated. The 
presence of autonomous endosperm with only egg cells in this hybrid strongly 
suggests that autonomous endosperm formation can be triggered in the absence of an 
autonomous embryo.

Although the presence of LOP did confer the ability to initiate autonomous 
seed development, viability of loa mutant 134 progeny was severely affected and this 
was also observed in cross-pollinated progeny that did not inherit LOP. This can be 
interpreted as the uncovering of embryo lethal components that function in the 
absence of LOA because meiosis can now occur. Embryo lethality was a dominant 
feature in mutant 134 progeny and this might suggest the accumulation of several 
recessive embryo lethal mutations that are uncovered upon meiosis.

The function of LOP is to override mechanisms that repress autonomous seed 
initiation in sexual plants. Mechanisms of central cell repression are mediated by the 
FERTILIZATION INDEPENDENT SEED (FIS)-class genes in Arabidopsis (Guitton
and Berger 2005). However, functional analysis of a FIS gene, FERTILIZATION INDEPENDENT ENDOSPERM (FIE) in Hieracium sexual plants showed that it is not involved in repressing seed initiation as described in Arabidopsis furthermore it is required for functional seed formation in Hieracium (Chapter 2). Therefore it is unlikely that the FIS genes, or at least HFIE, are part of the LOP locus in Hieracium. If alternative egg and central cell repression mechanisms exist in sexual Hieracium plants, LOP is capable overriding this mechanism to ensure autonomous seed formation. Alternatively, LOP could be dominant negative allele of such a repressive mechanism, resulting in a failure to properly repress seed initiation.

In conclusion, progeny analysis of apomictic mutants has enabled us to directly show that the sexual pathway is the default state in apomictic Hieracium caespitosum plants and the two dominant genetic loci that control apomixis, LOA and LOP, trigger changes in this pathway. We have also shown that autonomous endosperm formation and embryo development are closely linked at the LOP locus, at least in H. caespitosum. As sequence information becomes available, these apomictic mutants will be important tools for complementation analysis that will enable the identification of genes responsible for apospory and autonomous embryo and endosperm formation.
Methods

Plant Material

Hieracium caespitosum accession C4D and gamma deletion mutants are described in Catanach et al (2006) and were kindly provided by Ross Bicknell (Crop and Food Research – New Zealand). Plants are maintained under vegetative conditions as described in Koltunow et al. (1998). Greenhouse growing conditions are as described in Chapter 2. Cross-pollination experiments were done with pollen from the apomictic accession A4 (*H. aurantiacum*) and sexual accession P4 (*H. pilosella*). All accessions used are tetraploid. Manual pollination was carried out as soon as outer florets from capitula opened and stigma are available and proceed through 2-3 days until all florets from capitula are exposed to pollen.

Seed analysis

Mature seeds were collected, separated into three classes according to morphology, and germinated as described in Koltunow et al. (1998a; 2000). Germination was scored after 15-17 days in tissue growing conditions. Seed viability was calculated as total number of germinated seedlings divided by total number of seeds.

Microscopy

Developmental staging of flowers was described in (Koltunow, Johnson et al., 1998). Clearing of ovaries was done as described in Chapter 2.
Flow cytometry

Seedlings from germination experiments were grown for at least a month before transferring into individual tissue culture pots. A small piece of leaf tissue was chopped finely with razor blade in 300μl of extraction buffer (citric acid 2.0% Tween 0.5%) and filtered in a 30 μM mesh (Partec filter). Samples were kept in ice for no longer than 3h until analysis. Before flow cytometry, 800 μl of staining solution was added (2 mg/ml DAPI in saturated dibasic sodium phosphate – final concentration of DAPI 2.5 μg/ml) and staining allowed to proceed for 3-5 min. DAPI fluorescence was measured in a BLD LSRII flow cytometer with FSC (on log scale) set to 624 volts, SSC (log scale) 273 volts and DAPI (linear scale) 250 volts and a total of 10000 events were measured. Plant parents were used as reference and progeny was interpreted as having the same reading as parent plants, therefore tetraploid or half of DAPI count, considered diploid.
CHAPTER 4

CHARACTERIZATION OF THE MULTICOPY SUPPRESSOR OF IRA1 (MSII) GENE IN APOMICTIC AND SEXUAL HIERACIUM
Abstract

MULTICOPY SUPPRESSOR OF IRA1 (MSI1) is part of the FIS complex and $msi1$ mutants display the typical $fis$ phenotype of autonomous central cell proliferation and embryo abortion, but unlike other $fis$ mutations, also displays autonomous egg cell proliferation. In apomictic Hieracium plants, egg and central cell development are fertilization-independent and seed formation is completely autonomous. In this Chapter a Hieracium MSI1 homologue, $HMSII$, is characterized and it is shown that gene expression as well as spatial distribution of $HMSII$ mRNA is conserved in ovaries of sexual and apomictic plants. Both $HMSII$ and $HFIE$ expression are deregulated in apomictic mutant 179, which has a deletion in the LOP locus which impairs autonomous seed initiation. However, neither $HMSII$ nor $HFIE$ copy number has been altered in the genome of mutant 179, suggesting that these genes are not part of LOP. Together, these data suggest that autonomous seed development is not due to lack of $HMSII$ expression in apomictic egg cells and $HMSII$ might have a similar function to $HFIE$ in controlling sexual and apomictic embryo and endosperm formation.
Introduction

The MSI1 protein was initially identified in yeast as a suppressor of *ira1*, a mutation that disrupts the RAS-cAMP pathway that controls cellular growth (Ruggieri et al., 1989). In Arabidopsis, MSI1 has functions in different stages of the plant life cycle such as in floral meristem organization, flower morphology and flowering time and ultimately fertility (Hennig et al., 2003; Bouveret et al., 2006). AtMSI1 is also part of the Chromatin Assembly Factor (CAF)-1 complex, involved in depositing nucleosomes in newly replicated DNA in actively replicating meristematic cells and other replicating tissues (Exner et al., 2006).

AtMSI also has a role in female gametophyte and seed development as part of the FIS complex (Kohler et al., 2003; Guittton et al., 2004). In addition to autonomous proliferation of the central cell, autonomous proliferation of unfertilised egg cells were also observed in *msi1* mutants (Kohler et al., 2003; Guittton and Berger, 2005). Parthenogenetic *msi1* mutant embryos arrest early, but express embryo-specific markers, confirming its embryo identity (Guittton and Berger, 2005). The MSI1 protein is a homologue of p55 in *Drosophila* and RbAp48 in humans, which are known to interact with the cell cycle regulator RETINOBLASTOMA and this interaction is also conserved in plants (Ach et al., 1997). Similar to the *fis* mutations in Arabidopsis, mutations in *rbr* cause maternal gametophytic lethality (Ebel et al., 2004). In contrast to *fis* mutants, the supernumerary nuclei in *rbr* ovules do not have endosperm or embryo identity (Ingouff et al., 2006).

Aside from RBR, MSI1 also interacts with FIE as part of a 600 kDa protein complex which also contains MEA (Kohler et al., 2003). FIE is also capable of
interacting with RBR (Mosquina et al., 2004). The ability of both MSI1 and FIE to interact with RBR implies that the function of the FIS complex is linked to cell cycle control and cell differentiation in the female gametophyte. However, the distinct phenotypes of fie and msi1 mutants indicate that MSI1 might repress fertilization independent embryo initiation through an alternative pathway that is independent of the other FIS-class genes (Guitton and Berger, 2005).

In apomictic Hieracium plants the AI cell is able to form an unreduced embryo sac containing an egg and central cell that develop without fertilization to form embryo and endosperm, respectively (Bicknell and Koltunow, 2004). In Chapter 2, it was shown that unlike Arabidopsis, HFIE is not part of a central cell repressive complex in sexual and apomictic plants, but is needed for the formation of a viable seed. In this Chapter, MSI1 function in Hieracium seed development is addressed by determining if HMSI1 expression is altered in apomictic plants, particularly in apomictic egg cells, which are capable of developing in the absence of fertilization.

Results

Isolation and sequence analysis of Hieracium MSI1

A Hieracium homologue for MSI1 was isolated from a triploid apomictic plant H. piloselloides (D3) and a sexual tetraploid plant H. pilosella (P4). Using a degenerate primer approach cDNA fragments were isolated from ovary tissue containing mature gametophytes (stage 6). Sequence analysis confirmed significant homology to MSI1 sequences from tomato (Lycopersicum esculentum - AAB70241) and Arabidopsis thaliana - NP200631). For isolation of a full-length sequence,
5’RACE and 3’RACE were carried out using a Gene Racer library from stage 6 ovary RNA of apomictic and sexual plants. A cDNA of 1600 bp was isolated and conceptual translation showed a coding sequence of 1272 bp, 5’ untranslated region (UTR) of 159 bp and 3’UTR of 169 bp. Primers were then designed to amplify the full-length coding sequences from the apomictic plant D3 and the sexual plant P4 and another apomictic plant, the tetraploid *H. caespitosum* (C4d). DNA alignment of the *Hieracium* sequences showed a 99.1% identity between the two apomictic plants, 96.6% identity between apomictic C4d and sexual P4 and 97.3% between apomictic D3 and sexual P4. At the protein level, the apomictic and sexual *Hieracium* MSI1 cDNA encoded 423 amino acids which were identical with the exception of one amino acid at position 250, which in apomictic D3 is an arginine and in apomictic C4D and sexual P4 is a glutamine (Figure 1A). The presence of a glutamine in the sequences isolated from apomictic C4D as well as sexual P4 plants strongly suggests that this amino acid substitution does not compromise the function of MSI1 in sexual and apomictic plants. Protein sequence alignment with LeMSI1, AtMSI1 and *Oryza sativum* (Os) MSI1 showed a high degree of sequence conservation with the *Hieracium* homologues. *Hieracium* sequences were even 84% identical to the distantly related monocot sequence from rice and up to 92% identical to LeMSI1 (Table 1).

**Table 1.** Amino acid sequence identity matrix of plant MSI1 homologues.

<table>
<thead>
<tr>
<th></th>
<th>C4d</th>
<th>P4</th>
<th>Os</th>
<th>At</th>
<th>Le</th>
</tr>
</thead>
<tbody>
<tr>
<td>D3</td>
<td>99.2%</td>
<td>99.7%</td>
<td>84.3%</td>
<td>88.4%</td>
<td>91.7%</td>
</tr>
<tr>
<td>C4d</td>
<td>-</td>
<td>99.5%</td>
<td>83.8%</td>
<td>88.4%</td>
<td>91.5%</td>
</tr>
<tr>
<td>P4</td>
<td>-</td>
<td>-</td>
<td>84.3%</td>
<td>88.6%</td>
<td>91.9%</td>
</tr>
<tr>
<td>Os</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>85.0%</td>
<td>85.2%</td>
</tr>
<tr>
<td>At</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>90.8%</td>
</tr>
</tbody>
</table>

D3 – apomictic *H. piloselloides*; C4d - apomictic *H. caespitosum*; P4 - sexual *H. pilosella*; Os – *Oryza sativa*; At – *Arabidopsis thaliana*; Le – *Lycopersicum esculentum*
Figure 1. (A) Protein alignment of *Hieracium* MSI1 sequences isolated from apomictic (C4d and D3) and sexual (P4) plants with *Oryza sativa* (Os - ABF97823), *Arabidopsis thaliana* (At - NP_200631.1) and *Lycopersicum esculentum* (Le - AAB70241.1). The seven WD-40 repeats are underlined in blue. Red arrow indicates the single amino acid change between *Hieracium* sequences. (B) Neighbour-Joining tree of HMSI1 sequence with the Arabidopsis MSI family members showing that HMSI1 clusters with AtMSI1 and OsMSI1. (C) Southern blot analysis with MSI1 probe. Genomic DNA from *Hieracium* apomictic (D3 – APO) and sexual (P4 - SEX) was digested with EcoRI (El), EcoRV (Ev) and HindIII (H) that do not have sites in probe.
Five proteins of the MSI family are present in the Arabidopsis genome and these have different functions during Arabidopsis development. Only MSI1 has been shown to have a function in seed development. To ascertain that the sequence isolated from Hieracium was homologous to MSI1 and not other family members, phylogenetic analysis was performed using the neighbouring joining algorithm. The NJ phylogenetic tree clearly clusters HMSI1 to AtMSI1 and OsMSI1 in contrast to the other Arabidopsis MSI sequences (Figure 1B). Southern blot analysis using a full-length MSI1 cDNA as probe on apomictic D3 and sexual P4 genomic DNA revealed that it is a low copy gene in Hieracium, with one prominent fragment being detected and up to four fragments of less intensity in the genomes of sexual and apomictic plants (Figure 1C).

Interestingly, in the process of isolating PCR fragments with degenerate primers from ovary cDNA of these plants, several clones were identified with significant homology to MSI1 but containing deletions in conserved positions (Figure 2A). Ten clones were sequenced from both apomictic and sexual plants, and out of these, four clones from the apomictic plants and five clones from the sexual plant were detected that contained deletions of 17 bp, 2 bp and ,5 bp.. The location of the deletions was found to be in the second WD-40 domain and would potentially disrupt the production of a functional protein (Figure 2B). Sequence analysis showed that non-functional clones with deletions were more closely related to each other than to functional clones from the same plant (data not shown). It is possible that non-functional alleles are present in the polyploid genomes of both sexual and apomictic Hieracium plants, similar to that described for HFIE (Tucker, 2003), and that non-functional HMSI1 clones could still be expressed. Their
**Figure 2.** (A) Schematic representation of full-length HMSII cDNA. Numbered white boxes indicate the 7 WD-40 domains. Red bar shows the position of the 24 bp deletion in the 2nd WD-40 domain and arrows point to the position of the primers used to amplify a 332 bp fragment spanning this region, shown in (B). Non-functional alleles are characterized by three tandem deletions of 17 bp, 2 bp and 5 bp that map to the second WD-40 domain, underlined in red.
presence in apomictic and sexual plants suggests that these non-functional clones are unlikely to play a significant role in apomictic development. The sequence divergence observed between potential functional and non-functional alleles allowed specific primers to be designed for functional alleles and these were used for expression analysis throughout this study.

**HMSII expression analysis**

*HMSII* expression was analysed using quantitative RT-PCR and *in situ* hybridization (Figure 3). *HMSII* transcripts were detected throughout ovary development in both apomictic D3 and cross-pollinated sexual P4 plants and transcript levels were similar during early stages of ovary development, when megaspore mother cell differentiation and spore selection occurs (Stages 2 to 4) (Figure 3A). In cross-pollinated sexual P4 plants, embryogenesis initiates and by stages 8/9, globular embryos and nuclear endosperm are formed. Fertilization-independent embryo formation in apomictic D3 plants occurs after embryo sac maturity (stage 6) and by stages 8/9, similar to sexual plants; globular embryos and nuclear endosperm are also present. This stage of embryo initiation is marked by an increase of *HMSII* expression in both plants, but is more pronounced in apomictic plants, with a 2.2-fold increase in *HMSII* levels between stages 6 and 8/9 (Figure 3A). At later stages (stages 11/12), when torpedo embryos are present, *HMSII* expression decreases to similar levels in both plants.

In Arabidopsis, mutations in *msi1* lead to parthenogenetic development of the egg cell. We hypothesized that *HMSII* expression might be spatially altered in the egg cells of apomicts relative to egg cells from sexual plants, leading to autonomous
A

HMSI1 expression in Hieracium

![Graph showing expression levels of HMSI1 in Hieracium across different developmental stages of capitula.]

B

A/S

S

APO

SEX

![Images showing developmental stages of capitula labeled A/S and S. The images are labeled with CC (capped cell) and EC (empty cell).]
Figure 3. (A) Quantitative RT-PCR analysis of HMSII expression during apomictic (APO) and sexual (SEX) ovary development. Developmental stages of capitula are as described in (Koltunow et al., 1998). Stages 2/3 – early ovule from the apomicts development, appearance of megaspore mother cell (MMC); stage 4 – sporogenesis, in ovaries, differentiation of aposporous initial cells; stage 5/6 – embryo sac maturation; stage 8/9 – early seed development, globular embryos and nuclear endosperm; stage 11/12 – late stage seed development, torpedo embryos and cellular endosperm; (B) HMSII in situ hybridization, stage 6 ovules from apomictic (APO) and sexual (SEX) plants were hybridized with an anti-sense probe (A/S) and sense (S) probe. Positive signals appear in purple colour. Bars = 20 μM. cc – central cell; ec – egg cell.
development. *In situ* hybridization with a *HMSII* probe localized specific transcripts during the mature gametophytic stage of ovary development. In both apomictic and sexual plants, *HMSII* transcripts were detected within the embryo sac (Figure 3B). A strong signal was detected in the central cell of the apomictic embryo sac with an anti-sense probe and although the signal was less intense, it was also detected in the egg cell. This pattern was similar to *HMSII* distribution in sexual embryo sacs, where hybridization signals were detected in both the central and egg cell. No signal was evident in embryo sacs hybridised with sense transcripts. Aside from the intensity of the hybridization signal in the central cell of apomictic embryo sacs, the spatial distribution of *HMSII* transcripts was similar in apomictic and sexual plants in mature gametophytes, when autonomous embryogenesis initiates in apomictic plants. Therefore, autonomous development of the egg cell is unlikely to be due to the lack of *HMSII* expression in egg cells of the apomict.

**RBR isolation and expression in Hieracium**

A similar degenerate primer strategy to that used to isolate *HMSII* sequences was used to isolate RBR-like sequences from *Hieracium*. Primers were designed to the conserved A/B pocket domain of a RBR protein alignment and PCR fragments were amplified from ovary cDNA. Full-length cDNAs were obtained after 5’ and 3’ RACE. The full-length cDNA from apomictic D3 plants was found to be 3414 bp, with an open reading frame of 3015 bp encoding a potential protein of 1005 amino acids and a 5’UTR of 190 bp and 3’UTR of 209 bp. From the sexual P4 plant, a full-length cDNA contained 3514 bp, of which 207 bp was 5’UTR, 283 bp 3’UTR and a coding sequence of 3024 bp representing 1008 amino acids (Figure 4). Although the
cDNA sequences isolated were not identical in each plant, comparative analysis of translated sequences showed that the different cDNA sequences potentially encode a single identical protein in each plant. Comparative analysis of translated sequences from apomictic D3 and sexual P4 HRBR showed high similarity between the two sequences with 97.1% sequence identity. The characteristic A/B pocket domain is present and highly conserved in both protein sequences. The putative protein sequence isolated from the apomictic plant showed a three amino acid deletion at the C-terminal end of the protein, which lies outside of the A/B pocket or C-pocket domains (Figure 4). The HRBR sequence was more similar to maize ZmRBR3, identified as a homologue of the human RBR-like pocket protein p107 and belonging to Subfamily B of monocot RBR genes (Sabelli et al., 2005), than to ZmRBR1, of the monocot Subfamily C (Lendvai et al., 2007). This is in agreement with phylogenetic analysis of plant RBR genes in monocot and dicot species, which showed that the monocot Subfamily B is more closely related to the dicot RBR Subfamily A than the monocot Subfamily C (Sabelli et al., 2005; Lendvai et al., 2007).

Table 2. Amino acid sequence identity matrix of plant RBR homologues

<table>
<thead>
<tr>
<th></th>
<th>P4</th>
<th>At</th>
<th>Os</th>
<th>ZmRBR1</th>
<th>ZmRBR3</th>
</tr>
</thead>
<tbody>
<tr>
<td>D3</td>
<td>97.1%</td>
<td>66.4%</td>
<td>58.9%</td>
<td>42.8%</td>
<td>57.8%</td>
</tr>
<tr>
<td>P4</td>
<td>66.1%</td>
<td>66.1%</td>
<td>58.9%</td>
<td>42.6%</td>
<td>57.5%</td>
</tr>
<tr>
<td>At</td>
<td>59.4%</td>
<td>59.4%</td>
<td>42.8%</td>
<td>45.1%</td>
<td>58.0%</td>
</tr>
<tr>
<td>Os</td>
<td></td>
<td></td>
<td>59.4%</td>
<td>45.1%</td>
<td>84.2%</td>
</tr>
<tr>
<td>ZmRBR1</td>
<td></td>
<td></td>
<td></td>
<td>45.1%</td>
<td>44.3%</td>
</tr>
</tbody>
</table>

D3 - apomictic *H. piloselloides*; P4 - sexual *H. pilosella*; Os – *Oryza sativa* RBR1 (AY941774); At – *Arabidopsis thaliana* (AF245395); Zm – *Zea mays* RBR1 (X98923); Zm RBR3 (DQ124423)
Quantitative RT-PCR was performed to determine *HRBR* expression in the ovaries isolated from apomictic and sexual plants (Figure 5A). *HRBR* expression was found to be similar in both plants, with high *HRBR* levels in early stages of ovary development, which then dropped progressively throughout ovary and seed development (Figure 5A). The expression of *HMSII* and *HRBR* was also analysed in specific cell types. Laser capture microdissection was carried out by Dr. Yingkao Hu in early ovules of apomictic C4d plants in order to isolate cells related to apomictic initiation. Three cell types were dissected from early ovules: integument cells, aposporous initial (AI) cells and embryo sac (ES) cells, where cell division of the AI had occurred. RNA was extracted and amplified from laser capture cells and provided by Dr. Hu for gene expression analysis. Samples were checked for DNA contamination using primers that span intron sequences (Dr. Yingkao Hu, data not shown). Conventional RT-PCR showed similar levels of *HMSII* expression in all three cell types (Figure 5B). However, *HRBR* showed differential expression. *HRBR* expression was higher in integument cells and lower in both AI cells and ES cells (Figure 5B). Additionally, a fragment migrating at a higher than expected molecular weight was observed in the ES samples (Figure 5B). This additional fragment observed in ES cells of apomictic plants was isolated and sequence analysis confirmed it matched genomic DNA from this region of *HRBR*, therefore representing an unspliced transcript (Figure 5C). Conceptual translation of the coding sequence including this 88 bp intron region indicated that this unspliced transcript would potentially cause a shift in the reading frame leading to a premature stop codon and a truncated C-terminal region of HRBR (Figure 5D). The splice variant was not detected in apomictic (C4d and D3) and sexual P4 ovary samples, indicating that it is not abundantly expressed.
**Figure 4.** Alignment of plant RBR amino acid sequences with sequences isolated from apomictic (D3) and sexual (P4) *Hieracium*. Position of the A/B and C-pocket domains are underlined red, blue and black lines, respectively. The sequence from the apomictic plant has a three amino acid deletion relative to sexual boxed in green. At – Arabidopsis thaliana (AAF79146), Os – *Oryza sativum* (NP001062372), Zm – *Zea mays* RBR1 (AAB69649) and RBR3 (AAZ99092.1).
**A**

*RBR expression in Hieracium*

![Graph showing RBR expression levels across developmental stages of capitula.](image)

**B**

![Image showing RT-PCR results for different developmental stages.](image)

**C**

<table>
<thead>
<tr>
<th>Developmental Stage</th>
<th>mRNA Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>s2</td>
<td>800</td>
</tr>
<tr>
<td>s4</td>
<td>600</td>
</tr>
<tr>
<td>s6</td>
<td>400</td>
</tr>
<tr>
<td>s8/9</td>
<td>200</td>
</tr>
<tr>
<td>s10/11</td>
<td>100</td>
</tr>
</tbody>
</table>

**D**

![Diagram illustrating the gene structure.](image)
Figure 5. (A) Quantitative RT-PCR analysis of HRBR expression during apomictic (APO) and sexual (SEX) Hieracium ovary development; (B) HMSI1 and HRBR expression in specific cell types. Ovaries from early megasporogenesis, when apomixis initiates were sectioned and used for laser capture microdissection of integument cells (I), aposporous initials (AI) and embryo sacs (ES) (Dr. Yingkao Hu). Total RNA was isolated, amplified and provided by Dr. Hu for cDNA synthesis and RT-PCR analysis. Ubiquitin was used as internal reference and loading control. HRBR showed slight down-regulation in AI cells and ES cells along with the presence of a larger fragment in the ES sample. (C) Sequence analysis of HRBR alternative transcript. The 265 bp fragment was sequenced (RB Alt) and aligned with cDNA and genomic (RB gen) sequence for HRBR; (D) schematic representation of HRBR protein showing the position of the truncation in the putative unspliced variant protein.
HMSII-HRBR interaction is conserved in *Hieracium*

The previously characterized interaction between MSI1 and RBR reported for Arabidopsis was also tested for the *Hieracium* homologues. Due to the high sequence similarity between HMSII isolated from apomictic and sexual plants, with only one amino acid change identified, HMSII from the apomictic plant D3 was used as bait in yeast two-hybrid experiments. AtMSI1 was used as a positive control. Yeast co-transformed with bait and prey vectors containing Arabidopsis and *Hieracium* RBR-AD fusion proteins were able to grow in selective medium in the absence of leucine, indicating interaction of the AtMSI1 protein with HRBR from apomictic and sexual plants (Figure 6). Similarly, growth in selective medium also indicates that HMSII interacts with both Arabidopsis and *Hieracium* RBR proteins (Figure 6). HMSII interaction with DNA a methyltransferase of the MET1 class, HMET, was also tested and both AtMSI1 and HMSII are capable of binding to HMET protein. Interaction was not observed with the UBIQUITIN CONJUGATING ENZYME (UBC), used as negative control, indicating that the interactions are specific for RBR and HMET1 sequences.

**HMSII expression is altered in a lop mutant 179**

In order to understand the possible role of HMSII and HRBR in apomictic development, expression analysis was carried out in an apomictic mutant that has lost the capacity for parthenogenesis, mutant 179. Cytological and progeny analysis of this mutant was shown in Chapter 3. As with wild type apomictic plants, mutant 179 retained the capacity for apomixis initiation and AI cells can still be observed in early
Figure 6. Yeast two-hybrid assay using Arabidopsis (At) and *Hieracium* (H) MSI1 proteins as bait and At- and HRBR from apomictic (APO) and sexual (SEX) plants and a *Hieracium* DNA methyltransferase 1 (HMET) protein as prey. Bait-prey interaction restores DNA-binding and transcriptional activation domains and induces the expression of the reporter gene for leucine biosynthesis. Growth in selection media lacking leucine indicates positive interaction. The protein UBIQUITIN CONJUGATING ENZYME was used as a negative control (neg).
stage ovules (Chapter 3) (Figure 7A). At stage 6, WT C4d apomictic plants have already initiated autonomous seed development and contain developing embryos and nuclear endosperm (Figure 7A). By contrast, stage 6 mutant 179 ovules failed to initiate autonomous seed development and contain arrested egg cells and fused polar nuclei (Figure 7A). As with unpollinated sexual P4 plants, unpollinated egg and central cells in mutant 179 ovules eventually degenerated. Further characterization of this mutant with respect to seed viability and progeny analysis following cross-pollination were described in Chapter 3. Using quantitative RT-PCR, transcript levels in apomictic C4d WT and mutant 179 stage 6 capitulum, when differences in WT and mutant 179 ovules become evident, were examined. Down-regulation of HMSII expression was observed in mutant 179 ovaries relative to ovaries of the WT apomictic plant (Figure 7B). As low levels of HFIE expression in ovules leads to embryo and endosperm abortion (Chapter 2), the expression of HFIE in mutant 179 ovaries was also examined. Consistent with a role in autonomous seed development, HFIE expression was also down-regulated in mutant 179 (Figure 7B). Decreased transcriptional expression of epigenetic regulators was specific to HMSII and HFIE, as expression of HRBR, the DNA methyltransferases HMET, HCMT and the DNA glycosylase DEMETER was not altered in mutant ovaries (Figure 7B).

Down-regulation of HMSII in mutant ovaries might indicate a role in embryo development. Deletions or rearrangements in HFIE and HMSII in mutant 179 may be the cause of HMSII or HFIE down-regulation. Therefore, HMSII and HFIE gene copy number was determined in the mutant plant relative to the parent apomictic C4d plant. Genomic DNA from mutant and WT plants were used in Southern blot analysis with a HMSII probe and showed identical band patterns for both plants, indicating that the genes is not obviously rearranged in the mutant genome (Figure 7C). Similarly, using
HFIE as a probe, we did not detect major changes in HFIE copy number between apomictic WT C4d and mutant 179 (Figure 7D). However, it cannot be ruled out that the gamma irradiation caused a point mutation in HMSII or HFIE that would not be identified by genomic Southern and lead to altered levels of gene expression.

In conclusion, expression analysis in apomictic D3 plants together with cross-pollinated sexual P4 indicates that HMSII expression is induced upon both fertilization-dependent and independent embryo formation. Furthermore, the loss of the parthenogenesis locus interrupts autonomous seed development and this is associated with reduced expression of HMSII and HFIE, suggesting that these two WD-40 proteins play a fundamental role in initiating seed development in Hieracium plants.
Figure 7 (A) Cytology of loss of parthenogenesis mutant 179. Left panel shows apomixis initiation in WT apomictic plants and the differentiation of the aposporous initial cell (shaded white); next panel shows stage 6 ovules of WT plants that typically contain early embryos (em) and syncytial endosperm nuclei (arrows). Mutant 179 still retains the ability to undergo apomixis initiation, as shown by the differentiation of AI cells (shaded white) in early mutant ovules; stage 6 mutant ovules do not initiate autonomous seed development and embryo sacs contain arrested egg cell (ec) and central cell (cc); All bars=50μM; (B) Quantitative RT-PCR analysis of epigenetic regulators in WT apomictic and mutant stage 6 capitula (C and D) Southern blot analysis of WT and mutant 179 with a HMSI1 (C) and HFIE (D) probe. Genomic DNA was digested with the restriction enzymes BamHI (B) and HindIII (H) which do not cut MSI1 probe and EcoRV (E) and HindIII (H) for the FIE probe;.
Discussion

**HMSII expression is conserved in apomictic and sexual plants**

In this work, the role of MSI1 in fertilization independent *Hieracium* seeds was investigated. The data presented here shows that *Hieracium* apomictic and sexual plants encode multiple *MSII*-like genes, some of which may represent non-functional alleles. At the amino acid level, HMSII isolated from apomictic plants is practically identical to HMSII isolated from the sexual plant suggesting that it is unlikely that autonomous egg cell proliferation in apomictic *Hieracium* is due to genetic mutations that might affect HMSII function. Similarly, autonomous seed development is also not likely to be the result of altered spatial expression of *HMSII* in the female gametophyte of apomictic *Hieracium* plants as *HMSII* transcripts were detected in apomictic and sexual female embryo sacs. While the presence of *HMSII* transcripts in female gametes strongly supports a role in the initiation of seed development, it also suggests that autonomous seed initiation is not related to the absence of *HMSII* in these cells.

**Unspliced HRBR transcript in embryo sacs could indicate an early gametophytic function**

A *HRBR* cDNA with potential to encode a functional HRBR protein was isolated from both apomictic and sexual *Hieracium* plants. The high identity of the putative proteins also indicates that genetic mutations do not affect potential HRBR function. However, expression analysis in specific cell-types during initiation in
apomictic C4d plants suggests that *HRBR* is transcriptionally regulated. An alternative transcript that includes a small intron at the C-pocket domain of the HRBR protein was identified in laser capture samples. This intron introduces an in-frame stop codon that would truncate the HRBR protein near its C-terminus, disrupting the C-pocket domain. Along with the typical A/B pocket domain, this region is highly conserved in HRBR relative to other plant RBRs (Ach et al., 1997). The C-pocket domain of human RB has been shown to interact with different proteins relative to the A/B pocket and is also required for growth suppression (Whitaker et al., 1998). The alternative splice form was only detected in embryo sac cells, after AI cells had undergone the first mitotic division to form an unreduced embryo sac and occurred together with the normally spliced form. In the Arabidopsis plant life cycle, RBR function of cell cycle arrest is also required to maintain the balance of undifferentiated stem cells versus entry into a differentiation pathway in root meristems (Wildwater et al., 2005). Depletion of *RBR* in root meristems led to the accumulation of undifferentiated stem cells, suggesting that *RBR* expression is required to promote cell differentiation (Dinneny and Benfey, 2005). The phenotype of *rbr* mutants in the female gametophyte also suggests a possible role in the control of cell fate during female gametogenesis (Ebel et al., 2004). It is likely that the production of this alternative, truncated HRBR protein alleviates the repression on the cell cycle and enables the AI to initiate mitotic divisions and enter the gametophytic program. In this case, it would be expected that such a mechanism would also be conserved in the initiation of the gametophytic program of the sexual megaspore; however, this issue remains to be addressed. Alternatively, this might be a mechanism specific to apomictic plants enabling unreduced somatic cells to acquire gametophytic identity.
Further analysis of HRBR expression in isolated sections of sexual Hieracium is needed to elucidate this.

**HMSII/HRBR interaction is conserved in apomictic and sexual Hieracium**

In human cell lines, the homologue of MSI1, RbAp46, was initially identified by its ability to interact with Rb (Qian et al., 1993). This interaction has also been conserved in plants (Ach et al., 1997). The phenotypes of msi1 and rbr mutants suggest some functional overlap in female gametophyte maturation. The interaction of HMSI1 with HRBR indicates that this part of the pathway is conserved in Hieracium and this may account for the egg cell arrest in sexual plants. The HMSI1/HRBR proteins from apomictic plants also interact, indicating that this pathway is also present in apomictic gametophytes. Interestingly, Arabidopsis RBR expression becomes restricted to the central cell after embryo sac maturity, suggesting that MSI1 function in restricting egg cell proliferation might be independent of RBR (Ingouff et al., 2006). Functional analysis of MSI1 is required to determine if it has repressive roles in sexual plants and whether this has been conserved in apomictic plants.

Additionally, MSI1 was also shown to interact with a DNA methyltransferase of the MET1 class, homologous to Arabidopsis MET1. MET1 is the main maintenance DNA methyltransferase in plants and has been shown to be directly involved in the imprinting of PcG genes (Jullien et al., 2006). In human cell lines, the expression of the MET1 homologue, Dnmt1 is increased in cells lacking Rb activity and a direct interaction between the Rb and Dnmt1 proteins has also been demonstrated (Pradhan and Kim, 2002; McCabe et al., 2005). Interestingly, this interaction requires, at least in part, the C-pocket domain of Rb and leads to decreased
Dnmt1 enzymatic activity (Pradhan and Kim, 2002). The interaction between MSI1 and MET1 suggests that DNA methylation is associated with the MSI1/RBR pathway. The effect this interaction might have on MET1 activity and therefore on the expression of gametophytic genes, such as the imprinted PcG genes remain to be determined.

**HMSI1 and HFIE might have similar functions during autonomous seed development in Hieracium**

In Arabidopsis, mutations in *msi1* show the typical maternal effect on seed development characteristic of *fis* mutants and embryos abort after the heart stage. However, *msi1* mutants also show a sporophytic recessive effect that is distinct from the maternal gametophytic effect, as embryos carrying two mutant alleles abort early with irregular cell divisions (Guitton et al., 2004). The results in Arabidopsis seem to support a role for MSI1 in embryo development that is different from the FIS-pathway. It is possible that this role has also been conserved in embryo formation in *Hieracium*. *HMSI1* expression is down-regulated in apomictic *lop* mutant 179 in which autonomous development is impaired. Interestingly, *HFIE* expression was also down-regulated to similar levels as *HMSI1* in mutant 179 ovaries. We have previously shown that depletion of *HFIE* mRNA in apomictic ovules leads to autonomous embryo and endosperm abortion (Chapter 2). Therefore, the similar expression patterns of *HMSI1* and *HFIE* in mutant ovaries, together with the role of HFIE in apomictic plants suggest that *HMSI1* may also play an important role in controlling autonomous seed formation. Contrary to AtFIE and AtMSI1, HFIE and HMSI1 do not interact in yeast-two-hybrid assays, further indicating that HMSI1 and HFIE might act
in different complexes to regulate embryo development in *Hieracium*. However, further functional analysis of HMSII in *Hieracium* plants is needed to elucidate its function in autonomous and sexual seed development.

Interestingly, Southern analysis of HMSII and HFIE in mutant 179 did not shown any significant change in gene copy number. While it cannot be ruled out that point mutations has occurred as a result of gamma radiation that are not detected by Southern analysis, this result suggests that HMSII and HFIE are not part of the LOP locus and therefore are unlikely to be the direct trigger for autonomous seed initiation. The data presented here strongly support a role for HMSII and HFIE that is downstream of such a signal, ensuring the proper formation of embryo and endosperm. This model of action implies that LOP has the capacity to mimic developmental signals fertilization in order to initiate autonomous seed development.

**Material and Methods**

**Plant Material**

All plants and growing conditions are as described in Chapter 2.

**Gene isolation**

Degenerate oligonucleotides were designed to conserved regions of the MSI1 and RBR proteins according to sequence alignments of plant proteins available in the NCBI database. For MSI1 the primers used spanned the second and fifth WD-40 domain according to the alignment of Ach et al. (1998) while the HRBR primers were
designed at the border of the A/B pocket domain. Two pairs of forward and reverse
degenerate primers were used in nested PCR reactions. The first RT-PCR reaction
was performed using stage 6 ovary cDNA from apomictic and sexual plants as
template. Amplification products from this first reaction were diluted 25x and used as
template for a second-round nested reaction. PCR fragments were cloned into pGEM
(Promega) and typically five clones were sequenced for each fragment from each
plant. These sequences were then used to design primers for Rapid Amplification of
cDNA Ends (RACE) using a Gene Racer library (Invitrogen) prepared from stage 6
ovary cDNA from apomictic and sexual plants according to manufacturer’s
instructions. RACE fragments were cloned and sequenced and aligned with internal
fragments. Specific primers were then designed to the untranslated regions to amplify
full-length cDNAs for further sequence analysis. All primer sequences are provided in
the 5’ to 3’ direction. MSI1 degenerate primers: msi1F1 –
TGGAATAATAAYACTCCTTTYYTCTAYGA; msi1F2 –
GCNYTNGARTGGCCNTCNYTAC; msi1R1 –
TADATRTTYTCNGCCATYTGGCA; msi1R2 –
TCNGGNGGNCCRTCYTCNGCRTC. RBR degenerate primers: RBRF1 –
TTYTTYAARGARYTNCCNCARTT; RBRF2 – GGNYTAGTNCTNRTHTTGGC;
RBRR1 – AGYTCRTTTNGANGGGCTCTG; RBRR2 –
CKNARNGGAGANACRTANAC. HMSI1 RACE primers: msi15R1 –
TAGGTTCTTCTCCGATGTGTG; msi15R2 – CCCCTTAACCTCAAATCAGGA;
msi13R1 – GGATGGAATCCACAAAATGAG; msi13R2 –
GGTCTGGGATCTTAGGAAAATGAG. HRBR RACE primers: HRBR5R1 –
AGGTTGTTGCCATGTAGGATT; HRBR5R2 –
GATACCGCTACTTGGGAAAATG; HRBR3R1 –
GAAGATTTTATCGCTGCTTTG; HRBR3R2
TGTGAAATTAGCTGCGGTTAGA.

**Genomic Southern Blot Analysis**

Genomic DNA was isolated from leaves using the CTAB method as previously described (Tucker et al., 2001). Restriction digests were performed overnight and samples were separated on a 0.8% agarose gel. Standard procedures were used to transfer samples to nylon membranes (Amershan). Pre-hybridization and hybridization steps were done in phosphate buffer with 7% SDS at 65°C. Probes were prepared with Rediprime random priming labelling kit (Amershan) following manufacturer’s instructions. Full-length cDNAs were used as HMSI1 and HFIE probes. Washing of the membranes was done in standard sodium citrate buffer (SSC) at 2x concentration for 10min at 65°C, followed by SSC 0.2x washes for 15min. at 65°C.

**Quantitative RT-PCR**

Ovaries from each development stage were collected in liquid nitrogen and used for total RNA isolation with the RNA Easy Isolation kit (Qiagen). One microgram of total RNA was used for cDNA synthesis using a Superscript II cDNA synthesis kit (Invitrogen) according to manufacturer’s instructions. cDNAs were diluted 10x with Milli-Q water and 2μl were used in an RT-PCR reaction containing 280nM of each primer, 1x SyberGreen mix (Thermo Scientific) in a Rotor Gene 3000 real time PCR (Corbett Research). Ct values were used to determine *HMSII* and *HRBR* mRNA
levels and these were normalized against UB expression. Reactions were carried out in triplicate with biological repetitions of each developmental series. PCR conditions were as follows: initial denaturation of 15 min at 94°C, and 40 cycles of 94°C/10s, 58°C/15s, 72°C/20s. Amplified fragments were checked on agarose gels after PCR and sequenced to confirm identity. Means and standard deviations of at least two biological repetitions are presented. Primers used for amplification were as follows:

\[\text{msi1 QF} \quad \text{GATCAGGAACAAAACACCAGAGG}; \quad \text{msiQR} \quad \text{CACCTCGAGATTCTTACGCAG}; \quad \text{HRBR QF} \quad \text{GTCCAGGTTCAACCAAGGTATC}; \quad \text{HRBR QR} \quad \text{GGTACGCATGTGTGCTCTCTC}; \quad \text{ubQF} \quad \text{ACTCCACTTGGTCTTGC}; \quad \text{ubQR} \quad \text{AGTACCGCCGTCTTCAAGC}\]

**In situ hybridisation**

A full-length HMSI1 cDNA was cloned in both orientations into pGEM T-Easy (Promega). Plasmids were linearized with SpeI and T7 RNA polymerase was used for probe synthesis with a DIG-RNA labelling kit (Boehringer Mannheim). *In situ* hybridisation was carried out as described in Guerin et al. (1997). Ovary tissue was fixed in 4% paraformaldehyde and 0.25% glutaraldehyde, dehydrated using an ethanol series from 30 to 100% in 10% increments for 30s then embedded in butyl-methyl methacrylate and cut into 4μm sections. Hybridization was carried out at 45°C for 16h, after post-hybridization treatments, NBT/BCIP colour substrate was added to the slides and monitored for colour development.
Differential Interference Contrast microscopy

Capitula at appropriate developmental stages were collected and cleared as described in Chapter 2.

Yeast-two-hybrid assay

HMSI1 coding sequences were amplified, cloned into pGEM T-Easy (Promega) and sequenced to confirm that the ORF was intact. Full-length coding sequences were then cloned in frame into pLexA vector to create fusion protein with lexA DNA binding domain (Clontech). Similarly, RETIBLASTOMA RELATED PROTEIN (RBR), full-length sequences were cloned in frame into pB42AD to create fusion proteins with the activator domain. Transformation and selection conditions were the same as described in Chapter 2.
CHAPTER 5

DIFFERENTIAL EXPRESSION OF TWO CLASSES OF DNA METHYLTRANSFERASES IN OVARIES OF APOMICTIC AND SEXUAL HIERACIUM
**Introduction**

Methylation of cytosine residues in DNA is an important modification that is associated with the formation of heterochromatin. The plant family of DNA methyltransferases has been separated into classes that differ in sequence, domain content and substrate specificity (Bestor, 2000; Finnegan and Kovac, 2000). Class I, the DNA methyltransferase MET-class, which is closely related to the mammalian Dnmt1; the Class II chromodomain methyltransferases (chromomethylases or CMT) which are specific to plants and finally, Class III of domains rearranged methyltransferases (DRM) which are related to mammalian Dnmt3 (Bestor, 2000; Finnegan and Kovac, 2000). These classes are involved in methylation of cytosine residues in symmetrical sites (CpG, CpNpG) and de novo (CpNpN) methylation, respectively (Henikoff and Comai, 1998; Genger et al., 1999; Finnegan and Kovac, 2000; Cao and Jacobsen, 2002a,b).

In plant reproductive development, DNA methylation has been shown to be involved in parent-of-origin effects in Arabidopsis seed development, including in the control of imprinted genes (Jullien et al., 2006; Adams et al., 2000). Reciprocal crosses between hypomethylated and normally methylated gametes phenocopy interploidy crosses with effects on endosperm proliferation (Adams et al., 2000; Scott et al., 1998). These effects can be at least in part explained by MET1 function in the control of the imprinted genes FWA and FIS2 in the endosperm (Julien et al., 2006).

DNA methylation is also likely to play a role in post-zygotic seed development in both fertilization products. In met1 mutants, embryo patterning is defective, which affects seed viability and these defects appear in higher frequency in met1 cmt3 double mutants, indicating that these DNA methyltransferases act synergistically to control embryo development (Xiao et al., 2006). Additionally, in
maize the endosperm also undergoes extensive DNA hypomethylation indicating that this epigenetic mark is likely to be important in the establishment of gene expression programs for proper endosperm development (Lauria et al., 2004). The effect of DNA methylation on seed development is also evident through its association with the fis-class mutations. Mutations in the FIS genes lead to autonomous endosperm development and seed abortion. Pollen from hypomethylated plants is able to rescue the mutant phenotype (Guitton et al., 2004; Luo et al., 2000). These lines of evidence strongly suggest that DNA methylation has a direct role in controlling seed development by regulating gene expression in particular events of the sexual reproductive pathway.

DNA methylation is counteracted by DNA demethylation that is promoted by the DNA glycosylase DEMETER (DME) and DME-like genes (DML) (Penterman et al., 2007). DME activates the expression of at least two imprinted genes in the central cell: MEA and FIS2 (Gehring et al., 2006; Jullien et al., 2006). DME is capable of removing methyl groups from methylated DNA and this correlates with the activation of MEDEA and FIS2 (Gehring et al., 2006; Jullien et al., 2006; Choi et al., 2002).

The establishment of DNA methylation marks on imprinted genes might also occur during gametophyte development, where it has been shown genetically that the loss of MET1 activity during post-meiotic nuclear divisions of male and female gametophyte leads to the reactivation of silenced transgenes (Saaze et al., 2003). Interestingly, once DNA methylation marks are lost, they cannot be easily reset through meiosis, gametogenesis and somatic growth even in the presence of a functional MET1 locus (Finnegan et al., 1996; Saaze et al., 2003; Takeda and Paszkowski, 2006).
Although the role of DNA methylation in controlling sexual or fertilization-dependent seed formation has received considerable attention, its role in controlling fertilization-independent or autonomous seed formation has yet to be addressed. In *Hieracium* plants, both embryo and endosperm formation are independent of fertilization. In previous Chapters, it was shown that HMET interacts with a member of the FIS complex, HMSI1. It was also shown that the function of at least one FIS-class gene, HFIE, is essential in sexual and apomictic seed formation and does not have a role in repressing sexual seed initiation as described in the sexual model plant Arabidopsis. Due to the association of DNA methylation with the fie mutant phenotype in Arabidopsis and the altered role of HFIE in *Hieracium*, the role of DNA methylation in sexual and apomictic plants was examined in this Chapter. The isolation and characterization of a homologue of the main maintenance DNA methyltransferase, MET1, from apomictic and sexual *Hieracium* plants is reported along with the expression pattern throughout apomictic and sexual ovary development. A CMT-class DNA methyltransferase and the DNA glycosylase DEMETER from *Hieracium* were also partially characterized and their expression patterns compared in apomictic and sexual plants. Collectively these data show that there are differences in HMET genomic organization that appear to alter the amino-terminus of the apomictic HMET protein, and this might have an effect on how HMET binds RBR and MEA. Expression analysis showed that HMET and the counteracting gene HDME are co-ordinately expressed in ovaries resulting in a similar expression ratio and this was consistent with global levels of DNA methylation. The partial characterization of a CMT-like *Hieracium* homologue showed that HCMT is differentially expressed in late embryogenesis in apomictic seeds relative to sexual
seeds. Together, these findings indicate that the DNA methylation machinery is differentially regulated during reproductive events of apomictic development.

Results

Isolation and sequence analysis of HMET

Degenerate primers designed to the conserved catalytic region of the MET1-class DNA methyltransferases were used to amplify sequences from Hieracium ovary cDNA. Subsequently, amplification of cDNA ends enabled the isolation of a full-length cDNA from both sexual P4 and apomictic D3 Hieracium. Conceptual translation of the cDNA fragments identified an ORF of 1547 amino acids from the apomictic plant and 1569 amino acids for the cDNA isolated from the sexual plant. Two conserved domains were identified in each protein sequence that are typical of DNA methyltransferases of the MET1-class: the catalytic domain DNA methylase and two tandem BAH (bromo adjacent homology) domains (Figure 1A). The catalytic domains are located at the C-terminal region from amino acids 1107-1536 in the apomictic sequence and amino acids 1128-1557 in the sexual. Comparative amino acid sequence analysis of full-length HMET from sexual and apomictic plants revealed strong similarity along the entire length of the protein, particularly the catalytic domain, where sequence identity was 99%. Throughout the entire length of the protein, sequence identity between HMET from sexual and apomictic Hieracium was 93%. Identity of HMET from apomictic plant with proteins from other species is shown in Table 1. The catalytic domain is the region with highest homology among plant MET proteins. The two BAH domains were also identified in Hieracium MET
sequences and their location has been conserved compared to sequences from other plant species (Figure 1). The amino-terminal region of MET1 is reportedly less conserved and this was also observed when HMET was compared to Arabidopsis DMT1 (50%), *Daucus carota* AAC39355 (59%), *Zea mays* DMT101 (47%) and *Oryza sativa* DMT702 (47%), but was highly conserved, 94% identity, between the sequences from apomictic and sexual plants.

**Table 1.** Amino acid identity of apomictic D3 HMET (*Hieracium piloselloides*) protein compared with sexual HMET (*Hieracium pilosella*) and other plant MET proteins. Where available, the nomenclature adopted by Chromatin Database (www.chromodb.org) was used. P4 – sexual *Hieracium pilosella*; At- *Arabidopsis thaliana* DMT1 (At5g49160); Dc - *Daucus carota* AAC39355; Os – *Oryza sativa* DMT702; Zm - *Zea mays* DMT101.

<table>
<thead>
<tr>
<th></th>
<th>P4 Sexual</th>
<th>At</th>
<th>Dc</th>
<th>Os</th>
<th>Zm</th>
</tr>
</thead>
<tbody>
<tr>
<td>D3 apomictic</td>
<td>93%</td>
<td>55%</td>
<td>64%</td>
<td>56%</td>
<td>58%</td>
</tr>
</tbody>
</table>
A

ATG
P4 D3

B

C BAH  Catalytic domain  (429 AA)
Figure 1. (A) Schematic representation of the structure of HMET proteins. Like other MET-1 class proteins, HMET has two BAH domains positioned just before the catalytic domain. Regions of conservation were identified in the catalytic domain and these are numbered according to Finnegan & Kovac (2000). Alignment of rapid amplification of 5’ cDNA end fragments isolated from apomictic D3 and C4d and sexual P4 Hieracium HMET showing the divergence in 5’UTR region. In-frame ATG for apomictic sequences is shown in red and upstream ATG present in sexual P4 HMET sequence is boxed green. Upstream ATG in P4 HMET sequence potentially encodes 19 amino acids. The underlined regions in schematic diagram are shown in B and C; (B) The cDNA isolated from the sexual P4 Hieracium plant indicated it has the potential to encode an additional 19 amino acids at the amino terminal that is absent from HMET isolated from the apomictic plant; (C) HMET from the apomictic D3 plant also has a two amino acids deletions prior to the first BAH domain.
Although there was extensive sequence homology between \textit{HMET} from apomictic and sexual plants, the putative HMET protein from the apomict D3 was characterized by two deletions that might be of particular interest. The full-length sexual P4 \textit{HMET} cDNA contains an upstream in-frame ATG that would extend the amino-terminus of the sexual P4 HMET protein (Figure 1A). Although it is not known whether this ATG is used as a translation initiation site, it is the only in-frame ATG in the 5’UTR of sexual P4 \textit{HMET}. Furthermore, the 5’UTR of the apomictic D3 \textit{HMET} cDNA does not contain this upstream ATG and the remainder of 5’UTR sequence is quite diverged compared to \textit{P4 HMET}. Isolation of \textit{HMET} sequence from another apomictic plant, tetraploid \textit{Hieracium caespitosum} (C4d) also confirmed the similarity with apomictic D3 sequence in the 5’UTR (Figure 1A), further suggesting that this region has been conserved between these two apomictic species. This putative upstream ATG present in the sequence of the sexual P4 \textit{Hieracium} cDNA would potentially encode for an additional 19 amino acids (Figure 1B). The translated amino acid sequence of apomictic D3 HMET also contains a two amino acid deletion upstream of the two BAH domains, at position 671 of the apomictic HMET sequence (Figure 1C). This region is characterized by three repeats of the amino acids Glu-Gly (E-G) in sexual HMET; whereas in apomictic HMET, only two repeats are present. A total of 36 amino acid substitutions were identified between the two proteins and of these, 19 are considered conservative changes. Collectively, the isolation of \textit{HMET} from sexual P4 and apomictic triploid D3 and tetraploid C4d ovaries showed that HMET from the apomictic plant may be missing 19 amino acids that could potentially have functional relevance.
**HMET organization in the genome of apomictic and sexual plants**

Given the differences in the amino terminus of the sequences of HMET derived from sexual and apomictic plants, genomic sequences upstream of the coding sequence were isolated to further characterize the *HMET* genes from sexual and apomictic plants (Figure 2). Two 1.3 Kb fragments were isolated from the genome of the sexual plant, from BbrPI and EcoRV genomic libraries prepared for PCR chromosome walking (Tucker, M. 2003). Sequence analysis showed that these fragments were identical and matched the *P4 HMET* cDNA sequence. Alignment of genomic sequence with *P4 HMET* cDNA identified a splice site, -2 nucleotides upstream of the ATG (+1), which contains the conserved right border splice recognition signal AG. The intron sequence consists of 951 bp and the left intron border was identified by comparison with *P4 HMET* cDNA sequence and it also contained the conserved GT recognition signal. A small exon of 63 bp matched the *P4 HMET* cDNA sequence and this exon contained the in-frame ATG that would potentially extend the coding sequence of *P4 HMET* protein by 19 amino acids (Figure 2A). Sequences further upstream of this fragment were not identified in this study.

Genomic *HMET* sequences were also isolated from the apomictic D3 genome. After several attempts with different primer combinations, only two fragments of approximately 1.1Kb were isolated from the EcoRV genomic library. These fragments were very similar but showed some sequence variation, such as a seven nucleotide deletion at position -49 relative to the putative translation initiation ATG (+1). Both fragments matched the cDNA coding sequence throughout the region of overlap (127 bp). The genomic sequence started to diverge from the cDNA sequence...
six nucleotides upstream of ATG (+1), indicating the position of the splice site. Interestingly, this position differed relative to P4 HMET, in which the splice site was identified -2 nucleotides upstream of the ATG (+1). Genomic sequences further upstream of the D3 HMET were not isolated to identify other exon sequences that matched the 5’UTR present in the cDNA.

Sequence analysis of the intron from sexual P4 and apomictic D3 showed 74.1% similarity. The intron sequences were very similar until position -843 bp relative to the translation start site ATG+1. From this point, extensive sequence divergence was observed, however, only limited sequence information from the apomictic plants is available. Genomic sequences further upstream were not isolated from the apomictic D3 genome therefore; it is not possible to clearly establish if an exon sequence similar to P4 HMET is also present. However, the sequence divergence observed in this region prior to this exon together with the differences in the 5’UTR between apomictic D3 and sexual P4 HMET suggests that this exon is only present in the sexual P4 genome.

To further characterize the differences in HMET genomic structure and explore the polymorphisms in HMET sequences in the polyploid genomes of the apomictic and sexual plants, a region spanning the ATG (+1) was amplified using a reverse primer at position +127 bp of the coding sequence and a forward primer at position -340 bp of the intron. The resulting fragment of 467 bp was isolated from the genomes of sexual P4 and apomictic D3 plants and a total of nine random clones were sequenced. Representative clones are shown in Figure 3. Sequence comparison of this region showed that all clones matched the cDNA throughout the 127 bp of coding sequence and in the vicinity of the ATG and the right border of the intron. The clones
Figure 2. (A) Schematic representation of genomic fragments isolated upstream of *HMET* sequence from the sexual P4 (SEX) and apomictic D3 (APO) plants. Genomic fragments isolated from sexual plant have a small exon with an in-frame ATG that matches the cDNA sequence. Genomic sequences isolated from the apomictic plant are very similar in the intron region but diverge significantly from the sexual sequence from position -711 bp relative to ATG+1 as indicated by line in the *HMET* SEX sequence. Arrows indicate the position of the primers used to amplify 467 bp *HMET* fragment from the genomes of sexual P4 and apomictic D3 *Hieracium* in Figure 3; (B) Comparison of genomic and cDNA sequences from sexual and apomictic plant indicated that different splice acceptor sites were used between the two *HMET* sequences as indicated by the arrows.
from the sexual plant were highly similar to each other except for clones 4 and 9, which contained small insertions relative to the other clones. Clone 9 has a seven nucleotide insertion at position -159 bp and several single nucleotide substitutions. Similarly, clone 4 has a seven nucleotide insertion at -253 bp. Of the nine clones sequenced from the apomictic plant, three of them (clones 6, 7 and 9) have a seven nucleotide deletion at position -51 bp, along with other nucleotide substitutions. The high similarity in the coding sequence and in the region of the splice site suggests that *HMET* genomic organization has been conserved in this region in both apomictic and sexual plants. The comparison between genomic DNA and cDNA sequences further suggests that the D3 *HMET* and P4 *HMET* genes are differentially spliced and also suggests that P4 *HMET* has an additional exon relative added to its mRNA that contains an in-frame ATG and this exon does not appear to be present in *HMET* of the apomictic D3 genome.

The sequence variation encountered in this study suggests that the genomes of the triploid apomict D3 and tetraploid sexual P4 have multiple alleles of *HMET* as this is expected for polyploid plants. However, aside from the exon sequence in the genome of P4 *HMET* described above, variations were mostly in the intron sequence and the clones from sexual and apomictic plants matched the expressed cDNA, suggesting that these multiple alleles potentially encode DNA methyltransferases with identical protein sequence.

**Isolation of a partial cDNA for HCMT**

By designing degenerate primers to the plant-specific chromodomain, a cDNA fragment was isolated from ovaries of the apomictic plant. This cDNA clone was
Figure 3. Nucleotide sequence comparison from representative clones isolated from the genomic region spanning 340 bp upstream of ATG+1 (shown in blue) and 127 bp downstream of HMET coding sequence. Two types of clones were isolated from apomictic D3, with one type (D3-6) showing a deletion at position -51 relative to the other type that does not have this deletion (D3-1). Clones isolated from the sexual P4 plant also showed differences. While P4-9 has an insertion at position -159 that is also in the D3 clones but absent in the other P4 clones, clone P4-4 shows an insertion at position -253 bp that is unique relative to other clones from both apomictic and sexual plants.
extended by 5’ and 3’RACE, however, sequence alignment with CMT from other species indicated that the isolated cDNA was unlikely to represent a full-length \textit{HCMT} coding sequence (Figure 4A). Conceptual translation of this partial cDNA indicated an ORF of 571 amino acids with sequence homology to AtCMT3 (DMT6 – \url{www.chromodb.com}) and OsDMT701 (Figure 4B). This ORF was shorter by 226 amino acids at the amino terminal region relative to AtCMT1 and 270 amino acids relative to AtCMT3. Sequence identity with AtCMT3 and OsDMT701 was 58.0% and 58.7%, respectively; while HCMT shared 54.5% identity with AtCMT1 (DMT4).

In agreement with this, a phylogenetic tree was made using the Neighbouring Joining method and showed that HCMT grouped together with AtCMT3, with AtCMT1 falling outside of this clade (Figure 4C). Sequence alignment with the MET1 class DNA methyltransferase showed that the chromodomain is embedded between catalytic domains I and IV of HMET, as previously reported for chromomethylases, and that there is significant sequence variation between the catalytic domains of MET1 and CMT-type DNA methyltransferases (Figure 4A).

\textbf{Isolation of a partial cDNA for \textit{HDME}}

In Arabidopsis, DNA methylation marks can be removed from DNA by the activity of the DNA glycosylase DEMETER. Therefore, we used a similar degenerate primer approach as to that described for \textit{HCMT} isolation to isolate \textit{DME}-like sequences from \textit{Hieracium}. A cDNA fragment containing 461 amino acids of coding sequence corresponding to the catalytic domain was obtained and used for sequence alignment (Figure 5A). This fragment contains the DNA glycosylase domain and cysteine-rich regions and is highly conserved relative to the AtDME sequence.
Figure 4. (A) Schematic representation of CMT protein. The chromodomain is embedded between catalytic domains I and IV. This region was isolated from *Hieracium* ovaries and is shown in (B); (B) Amino acid sequence analysis of partial chromomethylase sequence isolated from *Hieracium*. HCMT sequence was aligned with HMET and AtMET (DMT1 – At5g49160) sequences to show the position of the chromodomain (boxed yellow) relative to the methylase catalytic domains (boxed blue). The position of the chromodomain is conserved in CMT and lies between domains I and IV; (C) phylogenetic analysis of partial HCMT sequence and the two Arabidopsis chromomethylases, AtCMT1 (DMT4 – At1g80740) and AtCMT3 (DMT6 – At1g69770) using the Neighbouring Joining algorithm.
Sequence identity to the Arabidopsis DME protein was determined for this region and shown to be 69.5% and 69.9% to the sequences from the apomictic and sexual plant, respectively. Phylogenetic analysis groups the partial sequences from *Hieracium* to AtDME and not to the other DME-like sequences from Arabidopsis (Figure 5C). Amino acid sequence comparison between DME isolated from apomictic and sexual plants showed strong sequence identity (95.6%) throughout the isolated region. Two deletions were observed in the sequence of the apomictic protein: amino acids Thr-His-His (THH in amino acids 162-165 in P4 sequence) and Asp-Cys (DC in amino acids 228-229) (Figure 5B). Both of these lie outside of the DNA glycosylase domain as well as the regions of conserved cysteine residues. Furthermore, a similar deletion is present in the region spanning THH at position 1426 of the AtDME protein (Figure 5). Therefore, these alterations are unlikely to reflect altered activity to the putative apomictic D3 HDME protein relative to the sexual counterpart.

**HMET and HDME expression in apomictic and sexual ovary development**

MET1 and DME play antagonistic roles in the Arabidopsis female gametophyte in the control of the expression of the imprinted gene *MEDEA (MEA)*. Methylation marks imposed by MET1 are erased at the *MEA* promoter by the DNA glycosylase activity of DME, leading to specific transcriptional activation of the maternal allele of MEDEA (Xiao et al., 2003). The relationship between *HMET* and *HDME* expression was investigated to determine if their relative expression was altered in apomictic ovaries relative to their sexual counterparts. In capitula of the tetraploid sexual P4, *HMET* expression was initially up regulated in stage 2, then
Figure 5. (A) Schematic representation of DME protein from Arabidopsis. Yellow box indicates the position of a nuclear localization signal (NLS), the blue shaded box indicates the position of the DNA glycosylase catalytic domain and the conserved position of four cysteine residues is shown by vertical red arrows. Black horizontal line shows the region that was isolated from sexual P4 and apomictic D3 *Hieracium* ovaries; (B) amino acid alignment of this region with the DME protein from Arabidopsis and three DME-like proteins (DML) from the Arabidopsis genome showing the high degree of conservation of the catalytic domain, underlined in green, and the position of cysteine residues in the *Hieracium* sequences (red arrows). Accession numbers for the sequences used in the alignment are DME – AF521596, DML1 – At2g36490, DML2 – At3g10010, DML3 – At4g34060. The apomictic sequence has two deletions relative to the sequence from the sexual plant and these are boxed in blue; (C) Phylogenetic analysis by the neighbouring joining algorithm showing that the *Hieracium* sequences are more closely related to Arabidopsis DME than to other DME-like (DML) sequences.
decreased throughout capitula at stages 4 and 6 where *HMET* mRNA levels reached basal levels (Figure 6A). Throughout the developmental stages of seed formation, in capitula at stages 8/9 and 11/12, *HMET* expression remained low. In contrast to *HMET* expression, *HDME* expression levels were constant throughout all stages of capitula development analysed, except during seed maturation, at stages 11/12, when *HDME* levels dropped significantly (Figure 6B). The relative expression ratio between *HMET/HDME* was analysed for the initial events of ovule formation and female gametophyte maturation (capitula at stages 2, 4 and 6). Relative *HMET/HDME* expression reflects differences in the expression pattern between the two genes. *HMET* expression was 5.3x fold higher at stage 2, and as *HMET* expression decreased and *HDME* expression remained relatively constant, this ratio decreased (Table 2).

The expression pattern of *HMET* throughout ovary development in the apomictic triploid D3 and the apomictic tetraploid C4d mirrored that of sexual P4 ovaries, with highest levels of expression in stage 2 capitula, which then dropped gradually in stages 4 and 6 (Figure 7A). Throughout seed development, *HMET* expression was maintained at low levels, and this expression pattern was similar to *HMET* expression in sexual P4 (compare Figures 6A and 7A). However, comparisons of *HMET* mRNA abundance in apomictic and sexual ovaries showed that *HMET* expression was significantly higher in apomictic ovaries relative to ovaries of the sexual plant. Higher levels of *HMET* expression were observed throughout all developmental stages analysed in both apomictic plants.

Analysis of *HDME* expression varied in the two apomictic plants analysed. In the triploid apomict D3, there was a significant decrease in *HDME* expression at stage 4, where *HDME* levels were maintained relatively constant throughout the other developmental stages analysed. In the tetraploid apomict C4d, *HDME* expression
dropped only modestly after stage 2, and continued to drop at later stages until \( HDME \) expression reached similar levels of expression compared to the triploid apomict D3 at stages 11/12. Similar to \( HMET \) expression levels in apomictic and sexual ovaries, \( HDME \) expression was also higher in ovaries of both apomictic plants relative to the sexual P4 plant. Analysis of \( HMET/HDME \) expression ratio in the apomictic plants showed that \( HMET \) expression was 3.2x and 4.1x higher than \( HDME \) at stage 2 in the tetraploid C4d and the triploid apomict D3, respectively (Table 2). At stage 4, this ratio dropped to 1.7 in C4d but remained high in D3 (8.0, Table 2). At stage 6, \( HDME \) expression was higher, and the ratio reached <1.0 in both apomictic plants. The trend in \( HMET/HDME \) expression was similar to that observed in the sexual plant and suggests that the expression of these two genes might be developmentally coordinated.

**HCMT expression analysis**

\( HCMT \) expression was analysed in ovaries of apomictic D3 and C4d along with cross-pollinated sexual P4 plants (Figure 8). In sexual P4 ovaries, \( HCMT \) expression was strongly up-regulated in initial events of ovule formation and this corresponded to the peak of \( HCMT \) expression. \( HCMT \) mRNA levels then decreased at later stages of ovary development and was lower at stage 4, dropping to basal state levels at embryo sac maturity (stage 6; Figure 8A). During seed development, \( HCMT \) expression remained low in cross-pollinated sexual P4. \( HCMT \) expression was similar in ovaries of apomictic C4d and D3 plants. High levels of \( HCMT \) mRNA were detected at stage 2, which then decreased at stage 4 and reaching the lowest level of expression at stage 6 (Figure 8A). This expression pattern was similar to the sexual P4 at these developmental stages. During apomictic seed development, \( HCMT \) expression
Figure 6. Expression analysis of *HMET* (A) and *HDME* (B) in ovaries and seeds of the sexual P4 plant using quantitative RT-PCR. Relative expression levels were determined by normalization with the constitutively expressed ubiquitin gene. Sexual P4 flowers were cross-pollinated with a tetraploid apomictic pollen donor. Staging of capitula was according to Koltunow et al. (1998). Schematic diagram shows the developmental events of each particular stage of capitula development. Stage 2 – differentiation of the megaspore mother cell; stage 4 – meiotic division and sexual spore selection; stage 6 – embryo sac maturity; stages 8/9 – globular embryos and nuclear endosperm; stages 11/12 – torpedo stage embryo.
Figure 7. Expression analysis of *HMET* (A) and *HDME* (B) in ovaries and seed of a triploid (D3) and tetraploid (C4d) apomictic plant using quantitative RT-PCR. Relative expression levels were determined by normalization with the constitutively expressed ubiquitin gene. Staging of capitula was as according to Koltunow et al. (1998). Schematic diagrams are as in Figure 6.
was initially low in stage 8/9 of both apomictic plants. However, at late stage 11/12, 
*HCMT* expression was up regulated in apomictic capitula and this is in contrast to 
*HCMT* expression at the same developmental stage in cross-pollinated sexual seeds 
(Figure 8A). Between stage 10 and 12, when embryos mature from heart to torpedo 
stage, there is a significant increase in *HCMT* mRNA levels in the apomictic plant, 
whereas *HCMT* levels are continuously low in ovaries of the sexual plant (Figure 8A). 
Collectively, the expression data shows that *HCMT* is up regulated in stage 12 
capitula apomictic plants relative to comparative stages in cross-pollinated sexual 
plants, in which *HCMT* expression remains low.

**DNA methyltransferase expression analysis in apomictic mutants**

To verify if the differential expression patterns of *HMET* and *HCMT* are 
functionally relevant, their expression patterns were analysed in apomictic mutant 
134, which has lost capacity for apomixis initiation, producing only reduced embryo 
sacs similar to the sexual plant. In this mutant, parthenogenetic embryos are still 
formed, although the genetic component that confers this phenotype segregates 
gametophytically in the progeny, and parthenogenetic embryos are only present in 
50% of the ovules (Chapter 3). In the tetraploid apomict C4d, *HCMT* showed an 
identical expression pattern to the triploid apomictic plant D3 (Figure 8B). In mutant 
134, *HCMT* expression was highest in the initial phases of ovule development, 
continuously decreasing to basal levels in stage 6 ovaries, and this is a similar pattern 
to that encountered in both apomictic and sexual plants. However, during seed 
development, *HCMT* expression in mutant 134 was similar to that of the sexual plant, 
with low levels of *HCMT* throughout all stages of seed development analysed (Figure
Figure 8. (A) *HCMT* expression analysis in apomictic D3 and sexual P4 ovary development. Schematic diagram of the cytological events in each developmental stage is shown. Sexual P4 was cross-pollinated with tetraploid apomict pollen donor. In sexual P4 and mutant 134 the appearance of enlarged somatic cells that give rise to unreduced embryo sacs does not occur and is shown in white in contrast to the unreduced embryo sacs produced by apomictic plants shown in grey; (B) *HCMT* expression in apomictic C4d and mutant 134 ovary development. Mutant 134 was generated through gamma irradiation of the C4d WT parent plant, as described in Catanach et al. (2006) and produce reduced embryo sacs of which 50% are still capable of forming autonomous embryos.
Figure 9. *HCMT* expression analysis in C4d WT apomict and mutant 115 in ovary and seed development. Mutant 115 has been characterized as an apomictic mutant that behaves as a sexual plant (Chapter 3). Unpollinated mutant 115 does not produce embryos. Cross-pollination was carried out with a tetraploid apomictic pollen donor. The cytological events of the developmental stages of capitula are shown for WT apomicts and cross-pollinated mutant 115. As expected, *HCMT* expression is similar to the sexual P4 *Hieracium* plant.
Therefore, \( HCMT \) expression in two apomictic species, triploid D3 and tetraploid C4d, is similar and is characterized by up-regulation of \( HCMT \) expression in late stage 12 of capitula development; whereas in sexual plants and in mutant 134, \( HCMT \) levels were maintained low at this developmental stage. The relevance of \( HCMT \) differential expression was further confirmed by analysis in mutant 115, which has deletions in both LOA and LOP and therefore behaves as a sexual plant (Chapter 3). It would be expected that \( HCMT \) expression in mutant 115 would resemble that of the sexual \( Hieracium \) plant. In early stages of capitula development \( HCMT \) expression followed similar trend as in \( Hieracium \) D3 and C4d apomictic, sexual P4 and mutant 134 (Figure 9). In late stages, \( HCMT \) expression level was similar to that of the sexual plant and mutant 134 and this was confirmed for unpollinated and cross-pollinated mutant 115 (Figure 9).

\( HMET \) and \( HDME \) expression in mutant background was also analysed in the WT apomict C4d and mutant 134 ovaries. The analysis was restricted to stages 2, 4 and 6 to assess the significance of \( HMET \) and \( HDME \) in apomixis initiation. \( HMET \) expression is not altered in mutant ovaries relative to WT apomictic ovaries (Figure 10A). However, HDME expression was shown to be increased in the mutant background (Figure 10B). Higher \( HDME \) expression levels resulted in lower \( HMET/HDME \) ratios in mutant ovaries compared to apomictic plants D3 and C4d and the sexual P4 plant (Table 2). The \( HMET/HDME \) expression ratio was elevated in early stages in these plants (Table 2), and in mutant 134, this ratio dropped significantly to 0.9 and 0.6 in stage 2 and 4, respectively (Table 2). There was no correlation between the altered expression ratios with the appearance of aposporous initial cells, as the \( HMET/HDME \) expression ratio in the mutant did not resemble that
Figure 10. (A) *HMET* expression analysis in stage 2, 4 and 6 ovaries of the WT apomictic C4d plant and an apomictic mutant 134 generated by gamma-irradiation of C4d. This mutant has lost the capacity for apomixis initiation and undergoes meiosis as in the sexual plant forming reduced embryo sacs shown in white in the schematic diagram, in contrast to the reduced embryo sacs formed in the WT apomictic plant, shown in grey; (B) HDME expression analysis in C4d WT and mutant 134.
of the sexual P4 plant. This suggests that HMET and HDME might not play significant roles in apomixis initiation.

Table 2. HMET/HDME expression ratios in ovaries at stages 2, 4 and 6 of sexual P4, apomictic triploid D3 and tetraploid C4d and in the C4d-derived apomictic mutant 134. The average normalized expression values of two biological repetitions were used to determine ratios.

<table>
<thead>
<tr>
<th></th>
<th>s.2</th>
<th>s.4</th>
<th>s.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sexual P4</td>
<td>5.3</td>
<td>2.3</td>
<td>1</td>
</tr>
<tr>
<td>Apomictic D3</td>
<td>4.1</td>
<td>8.0</td>
<td>0.5</td>
</tr>
<tr>
<td>Apomictic C4d</td>
<td>3.2</td>
<td>1.7</td>
<td>0.4</td>
</tr>
<tr>
<td>Mutant 134</td>
<td>0.9</td>
<td>0.6</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Analysis of methylated cytosines in ovaries of apomictic and sexual plants

Although HMET expression is significantly higher in apomictic ovaries, this increase was accompanied by higher HDME levels resulting in similar HMET/HDME expression ratios in early stage ovaries of apomictic and sexual plants. The proportion of methylated deoxycytosine (mdC) to deoxycytosine (dC) was measured by HPLC in DNA extracted from stage 4 ovaries of the tetraploid apomict C4d and sexual P4 plants as a measure of global DNA methylation of the genome. The proportion of
mdC in the sexual plant was 7.0% (SD 0.4%) and in the apomictic plant 5.8% (SD 0.5%). Therefore, the level of global methylation in stage 4 ovaries is similar between sexual and apomictic plants and is consistent with the $HMET/HDME$ expression ratio.

Analysis of $HCMT$ expression indicated increased expression in apomictic ovaries between stages 10 and 12 and this corresponded to the heart-torpedo transition of embryo development. To assess if increased $HCMT$ expression resulted in the accumulation of mdC in a particular tissue within the developing seed a monoclonal antibody specific for methylated DNA (anti mC) was used for immunolocalization in histological sections of apomictic and sexual ovaries. Specificity of the antibody was shown by the detection of a double-stranded methylated oligonucleotide, whereas the unmethylated double-stranded oligo was not detected (Figure 11A). During stage 10, fluorescence was observed throughout the expanded embryo sac, including both heart stage embryos and cellular endosperm of apomictic and sexual origin (Figure 11B).

**Yeast-two hybrid interaction of HMET1 with HRB and AtMEA**

In Arabidopsis, MET1 function is associated with regulating the expression of imprinted FIS-class genes and pollen from a hypomethylated plant is able to rescue the $fis$ phenotypes (Julien et al., 2006; Guitton et al., 2004; Luo et al., 2000). This functional association implies that MET is associated with the FIS pathway. Furthermore, the MEA homologue in humans, ENHANCER OF ZESTE2 (EZH2) and the human Rb protein are capable of binding Dnmt1 (Viré et al., 2006; Pradhan et al., 2002). Therefore, the ability of HMET to bind FIS-class proteins was investigated. Previously (Chapter 4), it was shown that MSI1 interacts with HMET. This interaction was observed using HMET sequence isolated from the apomictic plant but
Figure 11. (A) Analysis of the specificity of the monoclonal antibody for methylated cytosine. Methylated (mdC) and unmethylated (dC) double stranded oligonucleotides were blotted on replica DEAE membranes and either stained with ethidium bromide (EtBr) to demonstrate equal amounts of each double stranded oligo (left panel) or incubated with the primary antibody (anti-mdC), which was detected by secondary antibody conjugated with alkaline phosphatase (see material and methods). Colour reaction was only observed in mdC oligos. (B) anti-mdC antibody was used in histological section of ovaries at stage 10 (right column). Methylated DNA was observed by using a secondary antibody labelled with FITC and signals are observed by green fluorescence. Sections were also counter-stained with DAPI. No major changes can be detected in mdC distribution in apomictic (APO) and sexual (SEX) embryo and endosperm. Bars=50μM.
interactions using HMET from the sexual plant were not tested. This initial observation was expanded with a focus on the RBR protein isolated from *Hieracium* and the Arabidopsis MEA protein. HMET isolated from apomictic and sexual plants were tested as bait and AtMEA, HRBR and HFIE from apomictic and sexual plants as prey. The HMET isolated from the sexual plant used in this assay contained the 19 amino acid extension at the amino-terminus that is absent in HMET of the apomictic D3 plant. The ability to grow in selection medium lacking leucine indicates that both apomictic and sexual HMET proteins are capable of interacting with RBR isolated from apomictic and sexual plants as well as AtMEA, but not HFIE (data not shown).

We conducted a semi-quantitative β-galactosidase assay to measure the ability of interacting proteins to induce the expression of the reporter gene. The intensity of expression is directly associated with the affinity of the interaction between bait and prey, with tighter interactions capable of promoting higher expression of the reporter gene. β-galactosidase activity was always higher when using HMET from the apomictic plant relative to the sexual counterpart (Figure 12), suggesting that HMET from apomictic and sexual plants have different affinities for the proteins tested. However, it cannot be ruled out the affinity differences observed are due to translation efficiency of the HMET proteins in yeast. Further analysis is needed to rule out this possibility.
Figure 12. β-galactosidase assay of positive interactions. Colonies were grown in liquid selection media without uracil, histidine, thymidine and leucine in the presence of the chromogenic substrate ONPG. HMET isolated from apomictic D3 and sexual P4 were used as bait and tested for interaction with Retinoblastoma-related protein (RB) from apomictic and sexual *Hieracium* and Arabidopsis RB and MEDEA (MEA). The results shown are an average of two independent experiments in triplicate. Higher β-galactosidase activity was consistently observed when using HMET from apomictic D3 with all proteins tested relative to sexual counterpart.
Discussion

**Apomictic D3 HMET gene structure is distinct from that of sexual P4 HMET**

The isolation of regions immediately upstream of the putative HMET translation start site showed marked differences in HMET gene structure between the apomictic D3 and sexual P4 plants. The apomictic D3 HMET gene does not appear to have a small potentially coding exon, thus resulting in a protein 19 amino acids shorter than sexual P4 HMET protein. Furthermore, the genome of the apomictic D3 contains HMET alleles with small deletions that are distinct from the alleles in the genome of the sexual P4 plant. It is possible that these alterations might have a role in transcriptional regulation of HMET expression. Transcriptional control of MET1 expression is little understood in plants. In animals, Dnmt1 expression has been shown to be under the control of the cell cycle control protein Rb which represses Dnmt1 expression by binding to E2F sites in the Dnmt1 promoter, modulating Dnmt1 expression according to cell cycle progression (McCabe et al., 2005). In Arabidopsis, it has been recently shown that MET1 expression is up regulated in MSI1 co-suppressed lines indicating that MSI1 represses MET1 expression and this is most likely achieved through the direct binding of MSI1 to the MET1 promoter (Jullien, P. & Berger, F., personal communication). This suggests that as in animals, the RBR/MSI1 pathway is likely to control MET1 expression in plants. In an attempt to understand the altered HMET expression levels in apomictic Hieracium ovaries genomic sequences upstream of the putative translation initiation site were isolated. In Arabidopsis, over 75% of the E2F binding sites are located up to 400 bp upstream of the putative translation initiation site (ATG+1) (Ramirez-Parra et al., 2003). A search
of cis-regulating binding sites in this region using PLACE software (Higo et al., 1999) failed to identify putative E2F binding sites that could be targeted by RBR. However, the apomictic *Hieracium* D3 genome contains alleles with conserved deletions at -51 bp relative to ATG (+1). It is tempting to speculate that these deletions could affect MSI1/RBR-mediated repression in these alleles, thus contributing to increased *HMET* expression in apomictic D3 plants.

It was demonstrated using yeast two-hybrid assays that HMET from apomictic and sexual plants can also bind to *Hieracium* RBR protein. Interestingly, the difference in amino acid sequence, including a 19 amino acid extension present in HMET from sexual P4 *Hieracium*, could affect this interaction as shown by reporter assays as an indirect measurement of binding affinity. Higher reporter activity was observed when HMET from apomictic D3 *Hieracium* was the RBR partner and this could suggest that binding to RBR is stronger relative to HMET from the sexual plant. The functional significance of this interaction is as yet unclear. In human cells, the Rb protein can also control Dnmt1 activity post-transcriptionally, as the physical interaction between pRb and Dnmt1 proteins leads to decreased Dnmt1 catalytic activity (Pradhan et al., 2002). RBR function in plant development is related to the control of the balance of cell differentiation and cell division (Ebel et al., 2004; Wildwater et al., 2005). It is possible that the altered relationship between HMET and HRBR is related to differences in cell fate decisions that occur in apomictic plants, particularly in early stages of ovule development, when a somatic cell changes fate and enters a gametophytic program to produce the unreduced apomictic embryo sac. Alternatively, these differences in HMET/HRBR binding might only reflect a compensating mechanism to buffer DNA methyltransferase activity in response to higher mRNA levels.
**HMET and HDME expression are coordinated in Hieracium ovary development**

Plant reproductive development is distinct from animals in that sexual spores undergo mitotic divisions. Genetic analysis has shown that MET1 activity is essential to maintain epigenetic methylation marks during this phase of the plant life cycle (Saze et al., 2003; Takeda and Pazskowski, 2006). DNA methylation marks can be actively removed from DNA by the DNA glycosylase DEMETER (Xiao et al., 2003) and the interplay of MET1 and DME activity controls the activation of the maternally expressed gene MEDEA (Xiao et al., 2003; Gehring et al., 2006). The dynamic interaction between MET and DME in sexual and asexual reproduction in Hieracium was examined by analysing relative HMET/HDME expression levels in early ovary development. HMET expression pattern is similar in two apomictic plants, triploid D3 and tetraploid C4d and sexual tetraploid P4. However, HMET expression was shown to be higher in both apomictic species analysed relative to the sexual plant. Similarly, HDME expression levels were also higher in apomictic plants relative to the sexual plant. As the result of this, the expression ratios between HMET/HDME were similar in apomictic and sexual plants at these early developmental stages; indicating that HMET and HDME expression are coordinated. Increased HDME levels may compensate for higher HMET expression levels in apomictic ovaries. Due to the antagonistic function of HMET and HDME, it would be expected that a similar expression ratio would result in similar global DNA methylation levels. This is supported by HPLC analysis of global DNA methylation in apomictic and sexual ovaries at megasporegenesis, which did not show major differences in the proportion of mdC. Interestingly, the HMET/HDME ratio drops significantly at ES maturity. This could suggest that removal of methylated cytosines by increased DME activity is an
important aspect in the formation of a mature gametophyte. However, this does not seem to be related to the apomictic process, as this occurred in sexual and apomictic ovaries. In Arabidopsis, DME expression is restricted to the central cell of the female gametophyte and is highest in mature female gametophytes (Choi et al., 2002). It is possible that this increase in HDME expression relative to HMET expression leads to the activation of embryo sac specific genes essential for embryo sac maturation and seed initiation. Localization of HMET and HMDE expression at this stage is needed to further address this issue.

These data presented here do not support a role of DNA methylation in apomixis initiation and this is further corroborated by analysis of HMET expression in loa mutant 134 which does not make AI cells. However, it cannot be ruled out that changes in specific loci might occur between sexual and apomictic plants that could be significant to the apomictic process. This could be addressed by altering the DNA methylation status in apomictic and sexual Hieracium ovaries.

**HCMT expression is associated with late stage seeds derived from unreduced gametophytes**

The CMT3-like sequence isolated from Hieracium showed differential regulation during apomictic and sexually-derived embryogenesis, in particular at late stages of embryo development, in the transition from heart to torpedo stage, where it is up regulated in apomictic seeds. The functional relevance of this differential expression is not yet clear. Expression analysis in loa mutant 134 showed that HCMT expression pattern was similar to that of the sexual plant. This might suggest that HCMT expression is up regulated in stage 12 seeds that are derived from unreduced
embryo sacs, relative to seeds derived from reduced embryo sacs, in which \textit{HCMT} is maintained at low levels at stage 12. However, mutant 134 ovaries segregate for the autonomous seed component of apomixis (Chapter 3), so that approximately 40\% of the ovules contain parthenogenetic embryos. Therefore, we cannot rule out that \textit{HCMT} expression is low in mutant 134 at this stage due to the high proportion of ovules that do not contain embryos. Immunostaining to localize methylated cytosine did not detect a specific region in apomictic seeds where methylated DNA was accumulating. However, we cannot rule out that changes at specific loci occur that are not detected by this method. Further analysis of \textit{HCMT} expression in pollinated mutant embryo sacs and localization could elucidate if \textit{HCMT} expression in fact marks embryos that are derived from unreduced embryo sacs. In Arabidopsis, CMT3 activity is needed for cytosine methylation at the CNG context and genetic screens for suppressors of reporter and endogenous gene silencing has identified \textit{CMT3} as the main CNG maintenance chromomethylase (Lindroth et al., 2001; Bartee et al., 2006). Recently, genome-wide high-resolution mapping of DNA methylation showed that CNG methylation is more likely to be associated with functionally relevant genes rather than retrotransposons (Zhang et al., 2006). Genetic analysis has also shown that CNG targeting mechanism depends on the siRNA pathway (Chan et al., 2006). It is tempting to speculate that \textit{HCMT} expression in embryos derived from unreduced embryo sacs could be related to establishing apomictic reproductive identity in the embryo, by regulating specific gene expression programs.
Material and Methods

Plant Material

All plant materials and growing conditions as per Chapters 2.

HMET and HCMT isolation

Degenerate oligonucleotides were designed to conserved regions of the catalytic domain based on alignment of MET1 sequences from other plant species. Two pairs of forward and reverse degenerate primers were used in nested PCR reactions. The first RT-PCR reaction was performed using stage 6 ovary cDNA from apomictic and sexual plants as template. Amplification products from this first reaction were diluted 25x and used as template for a second round nested reaction. PCR fragments were cloned into pGEM (Promega) and typically five clones were sequenced for each fragment from each plant. These sequences were then used to design primers for Rapid Amplification of cDNA Ends (RACE) using Gene Racer library (Invitrogen) prepared from stage 6 ovary cDNA from apomictic and sexual plants according to manufacturer’s instructions. RACE fragments were cloned and sequenced and aligned with internal fragments. Specific primers were then designed to the untranslated regions to amplify full-length cDNAs for further sequence analysis.
Quantitative RT-PCR

Tissue preparation, RNA extraction, cDNA synthesis and RT-PCR conditions were carried out as per Chapter 4.

Primers used for amplification were as follows (5’-3’):

- HDMEQF  ACCATGTCGAACTGCAATGA
- HDMEQR  TGATCAAATCCCTCACACA
- HCMTQF  CTTGGTCCAATCAAAGAAAGGT
- HCMTQR  AACAAGTCCTCTCCATAACCAA
- HMET13QF AAGATCGAATCCTCACCGTTC
- HMET13QR GCCTGACTTTCTGTTTGCT

Yeast-two-hybrid assay

HMET coding sequences were amplified, cloned into pGEM T-Easy (Promega) and sequenced to confirm ORF was intact. Full-length cds were then cloned in frame into pLexA vector to create fusion protein with lexA DNA binding domain (Clontech).

Similarly, RETIBLASTOMA RELATED PROTEIN (RBR), full-length sequences were cloned in frame into pB42AD to create fusion proteins with the activator domain. Transformation, growing conditions and selection were performed as described in Chapter 2.
**Dot blot analysis of antibody specificity**

Specificity of the monoclonal antibody against methylated cytosines (Aviva Systems Biology, USA) was tested as described in Oakeley et al. (1997). Complementary oligonucleotides 5’AGCTCAGACTTTGAGGCCTTAA / 5’TTAAGGCCTCAAAGTCTGAGCT with methylated cytosines were annealed by heating an equimolar solution in 100mM NaCl and allowing the mixture to cool at room temperature. The double-strand oligos were precipitated in ethanol. Methylated and unmethylated annealed oligos were spotted on replica DEAE membranes at 5, 2.5 and 1μg. One membrane was stained with ethidium bromide at 500μg/ml for 1 minute. The other was incubated with 1:1000 dilution of the antibody in DEAE wash buffer (50mM NaCl, 10mM Tris, ph7.5, 1mM EDTA, 1% Triton X-100) for 16h at 4°C. The membrane was then washed with DEAE buffer from 3x5min and incubated with the secondary anti-mouse alkaline phosphatase conjugate at 1:2000 dilution for 1h at RT. NBT/BCIP solution (Sigma) was used as substrate for colour development. As a control, the same primer pair with unmethylated cytosines was used.

**Immunostaining**

A monoclonal antibody specific to methylated DNA was used for immunolocalization. BMM embedded sections (3μm) were incubated in acetone for 20min to remove BMM and washed 2 x 2min with PBS (1L containing 8g NaCl, 0.2g KCl, 1.15g Na2HPO4, 0.2g KH2PO4, 0.2g NaN3, pH 7.3-7.6). Sections were then incubated in 1% blocking reagent (Roche) prepared with PBS for 45min. at room temperature. The mouse monoclonal antibody against methylated cytosine (Aviva
Systems Biology, USA) was used as primary antibody at a 1:100 dilution in PBS for 16h at RT. Sections were washed 3x5min in PBS before a 1h incubation with anti-mouse IgG-FITC conjugate diluted 1:100 with 1% blocking solution in PBS. Sections were then washed 3 x 5min in PBS, air dried and mounted with Vectashield containing DAPI (Vector Laboratories, UK). Sections were observed in a Zeiss Axiophot microscope using appropriate filter combinations. For FITC the filter set 487909 (Blue 450-490, Zeiss) was used and for DAPI the 487901 (UV-H 365, Zeiss) was used. Images were recorded with a Spot digital camera.

**HPLC analysis of global DNA methylation**

The protocol used was based on Johnston et al. (2005). Genomic DNA was isolated from stage 4 ovaries of apomictic and sexual *Hieracium* plants and digested with RNase A (Sigma) for 16h at 37°C. Samples were precipitated with isopropanol and 25μg used for digestion with nuclease P1 (Promega) for 16h at 37°C and alkaline phosphatase (Promega) for 2h at 37°C. Samples were separated in a Synergi C18 250 x 4 mm column (Phenomenex) with buffer A 100% (0.5% methanol; 10mM KH$_2$PO$_4$ pH 3.7) for 5 min and a gradient of 3-20min. with buffer B (10% methanol; 10mM KH$_2$PO$_4$, pH 3.7) for 20min. A standard curve for dC and mdC was used to determine the concentration in each sample. Global DNA methylation was calculated as mdC% = mdC/mdC+dC. The assay was done in triplicate and the results reflect the average and SD of two biological repetitions.
CHAPTER 6

DISCUSSION
Discussion

The FIS genes and the apomixis hypothesis

The observation that mutations in the FIS-class genes leads to the initiation of autonomous endosperm development in Arabidopsis has lead to the hypothesis that these genes are altered or the function of the complex compromised in apomictic plants (Koltunow and Grossniklaus 2003; Bicknell and Koltunow 2004). The specific aim of this thesis was to test this hypothesis in the autonomous apomictic plant Hieracium by isolating orthologues of the Arabidopsis FIS-class genes to examine their function in apomixis.

The results presented here suggest that a FIS complex identical to the repressive FIE/MEA/MSI1/RBR complex is not formed in apomictic and sexual Hieracium plants. Two main lines of evidence support this: (1) autonomous central cell proliferation was not observed in sexual Hieracium lines in which the expression of HFIE was down-regulated in ovules; and (2) protein interaction studies show that HFIE does not interact with the MSI1, HRBR and AtMEA proteins as described in Arabidopsis. Molecular modelling of AtFIE and HFIE proteins showed specific differences in the loop structures of these proteins that protrude from the beta-propeller and could alter surface-charge which might explain the differences in interaction capabilities. Ultimately, domain swapping experiments could help determine the specific changes in sequence and structure of the HFIE protein that are responsible for abolishing these interactions.
HFIE has a post-fertilization function in seed formation

Data presented here directly show that HFIE is essential for the formation of embryo and endosperm in apomictic and sexual Hieracium plants. In this case, the repressive functions that are observed in Arabidopsis are absent in Hieracium. It is possible that this repressive role of the FIS complex might be specific to the Brassicaceae lineage. In support of this view, the SET domain protein MEA is not found outside of the Brassicaceae (Spillane et al., 2007). Furthermore, previous work in Hieracium to isolate a FIS2 homologue was unsuccessful (Chaudhury et al., 2001), suggesting that this gene may not be present in Hieracium. Therefore, HFIE function in embryo and endosperm formation might be considered similar to the post-fertilization function in Arabidopsis seed development. This suggests that the primary role of FIE might be in controlling seed development.

It should be emphasized that while a HFIE does not participate in a repressive complex it is likely that a repressive complex of some form might function in Hieracium to restrict egg and central cell proliferation because in sexual Hieracium the female gametophyte arrests at maturity and seed development requires signals from fertilization. The identification of genes encoded at the LOP locus in Hieracium provide further insights into the mechanisms of female gametophyte repression and seed activation that are present in Hieracium.

The finding that HFIE and the FIS-complex is altered in Hieracium relative to the Arabidopsis counterparts also raises questions as to the role of this complex in other plants. To date, an autonomous phenotype has yet to be shown in other sexual model plants such as maize and rice. It remains unclear how these molecular mechanisms that control parent-of-origin effects, such as genomic imprinting, have
evolved in sexual *Hieracium* plants. Genomic imprinting appears to have evolved to control nutrient allocation to the developing embryo and the clear examples of imprinted genes are involved in controlling endosperm proliferation. Different species appear to support different maternal: paternal contribution for endosperm formation. While in some, such as maize, this relative contribution is strict 2m:1p and deviations result in seed abortion, others, such as Arabidopsis are more flexible. Ultimately, autonomous apomictic plants have completely overcome any imprinting barrier and are solely under maternal control.

**HFIE and HMSI1 are unlikely to be LOP determinants**

The presence of *HFIE* and *HMSI1* in the genome of *lop* mutant 179 suggests that these genes are not specifically involved in promoting autonomous seed formation and have a broader role in supporting embryo and endosperm growth, independent of their origin. Both of these genes were shown to have decreased expression in mutant 179 ovaries, suggesting that they are needed for the initiation of an autonomous seed. We cannot rule out the possibility that point mutations have affected the function of these genes in the gamma deletion mutants. The isolation of *HMSI1* and *HFIE* from mutant 179 background and comparison to the WT gene sequences in the apomictic plant would address this issue.

Additionally, a comparative approach of gene expression in ovaries of mutant 179 plants and their WT apomictic parent might identify genes that are de-regulated and likely candidates to perform such repressive roles. This approach could be further enhanced using laser microdissection of egg and central cell in mutant 179 and WT for comparisons in specific cell types. The use of these mutants, although generated
by gamma radiation, has the benefit of being a nearly isogenic background in relation to other comparative studies that use different ploidy plants or plants of hybrid origin. Complementation studies of candidate genes could be readily scored by the formation of viable autonomous seeds, in contrast to completely sterile seeds that are formed in unpollinated mutant 179.

**Does autonomous endosperm depend on autonomous embryo initiation in *Hieracium***?

Cytological and progeny analysis of apomixis mutants shows that embryo and endosperm formation are closely linked in developmental timing and either one gene allows the initiation of both or the genes stimulating seed initiation in apomictic *H. caespitosum* are closely linked. In Arabidopsis it has been shown that a positive signal from an egg cell fertilized with pollen carrying a mutation in the *CDKA;1* gene triggers the proliferation of the central cell that is able to form functional endosperm with no paternal contribution in the absence of the repressive function of the FIS complex (Nowack et al., 2006; 2007). The data provided here support a model in which such a mechanism could also be present in apomictic *Hieracium* (Figure 1). The positive signal would be provided by LOP and the autonomous formation of the embryo. Alternatively, LOP might act directly in both the egg and central cell to stimulate autonomous embryo and endosperm formation, respectively. This model could be tested in pseudogamous apomictic plants. These plants require fertilization for the formation of endosperm although parthenogenetic development of the egg cell still occurs. Examining and testing FIS-related interactions in these plants would shed light into the relevance of these interactions for autonomous endosperm development.
Figure 1. Model for the formation of autonomous endosperm in Arabidopsis (A) and *Hieracium* (B). Model described in text. Fert. – Fertilization; EC – egg cell; CC – central cell; LOP – Loss of Parthenogenesis locus
HMSI1 function in *Hieracium*

Due to the autonomous egg cell phenotype of the *msi1* mutants in Arabidopsis, *HMSI1* mRNA distribution in mature gametophytes of sexual and apomictic *Hieracium* plants was analysed. *HMSI1* mRNA was present in both female gametes independent of the mode of reproduction, suggesting that the ability to form autonomous embryos is not due to the lack of *HMSI1* expression in apomictic egg cells. HMSI1 binds to both HMET and HRBR. It is unclear what role HMSI1 might have in seed initiation. *HMSI1* expression is de-regulated in an apomixis mutant 179, which has lost the capacity to initiate autonomous development. *HFIE* expression was also down-regulated to similar levels in this mutant, which is consistent with the phenotypes of the AtMEA::hpHFIE apomictic lines. This suggests that *HMSI1* could also have a role in autonomous embryo development similar to *HFIE*. However, no interaction between HFIE and HMSI1 was observed in this study, further indicating that embryo development might require both HMSI1 and HFIE that act in distinct complexes (Figure 2). Southern analysis showed that neither HMSI1 nor HFIE genomic organization were altered in mutant 179 and this suggests that the deletion that impaired autonomous seed initiation did not affect *HMSI1* or *HFIE* copy number. This data indicates that, although HMSI1 and HFIE appear to play essential roles in seed formation, this role is not specific to apomictic development. The data support a model in which autonomous seed development mimic signals normally provided by fertilization that mobilize HFIE-dependent and HMSI1-dependent function that enable the formation of a viable embryo (Figure 2).

In the course of this study, it was not possible to identify a protein isolated from *Hieracium* that is capable of binding to HFIE. Therefore it remains unclear what
Figure 2. Model for the roles of HFIE and HMSI1 in *Hieracium* seed formation. The locus that controls autonomous seed formation (LOP) in apomictic *Hieracium* mimics the signals provided by fertilization in sexual *Hieracium* plants so that both pathways use the same machinery to initiate seed development. HFIE function is likely through the interaction with a putative SET-domain protein, while HMSI1 would act independent of HFIE in a complex that possibly involves HRBR and HMET.
type of PcG complex involving HFIE is formed in *Hieracium* embryo sacs and seed. The interaction observed between HFIE and AtCLF suggests that other *Hieracium* SET-domain protein homologues are likely candidates. The isolation of these homologues from *Hieracium* might enable the identification of HFIE PcG complex members.

Interestingly, we also detected an interaction between HMSI1 and a DNA methyltransferase of the MET1 class. The analysis of *met1* mutants in Arabidopsis has shown that DNA methylation is important in establishing early patterns of cell division during embryogenesis and correct expression of embryo identity genes (Xiao et al., 2006). These phenotypes are similar to the recessive sporophytic effect of *msi1* mutants. The interaction between HMSI1 and HMET1 reported here could indicate that MSI1 and MET1 associate to establish or maintain proper gene expression programs that are essential for embryogenesis. Further analysis of *HMET* mRNA distribution is needed to confirm if it overlaps with *HMSI1*.

**Role of HCMT expression in late stage seed development derived from unreduced embryo sacs**

In the course of this work we identified differences in the expression of another DNA methyltransferase belonging to the CMT-class. The isolation of a partial *CMT3* homologue from *Hieracium* and expression analysis during ovary and seed development revealed a peak in expression at late stages of seed development that is specific to apomictic plants. This was further confirmed by analysing *HCMT* expression in apomictic mutants 134 and cross-pollinated 115, both of which produce embryos from meiotically-derived embryo sacs. This result suggests that late stage
$HCMT$ expression is a marker for seeds that are derived from unreduced embryo sacs. However, we cannot rule out the possibility that the lower $HCMT$ expression in mutant 134 and 115 backgrounds is due to the number of ovules that do not produce an embryo. The functional relevance of this epigenetic mark in apomictic embryo development remains unclear. Analysis of seed abortion in Arabidopsis $cmt3$ mutants revealed only a modest effect on seed viability, although the analysis of $met1$ $cmt3$ double mutants showed a stronger effect on the ability to produce viable seeds compared to the single mutants (Xiao et al., 2006). This suggests that CMT3 and MET1 most likely perform redundant functions in embryogenesis. However, it is currently unknown if viability is affected at late stage of seed development in the double mutant.

It is possible that $HCMT$ plays a role in late stage embryo development that is specific to apomictic plants. The transition stage from heart stage embryo stage in Arabidopsis corresponds to the differentiation of shoot apical meristem and the definition of specific regions that define stem cell competency. It is tempting to speculate that alterations in DNA methylation in specific cell types could be related to this process. In support of this, the Argonaute protein ZWILLE has an important role in establishing and maintaining stem cell competency in the meristem. Argonaute proteins are involved in silencing mechanisms that depend on small interfering RNA (siRNA). siRNA can promote transcriptional silencing by directing DNA methylation marks to homologous regions. This process has been shown to involve CNG methylation and therefore depend on CMT3 activity. In this speculative scenario, the establishment of apomictic or sexual reproductive fate might occur in these meristematic cells and be maintained throughout the sporophytic life cycle.
Alternatively, differences in DNA methylation could be related to endosperm development. This tissue is particularly sensitive to imbalances in DNA methylation, as shown by reciprocal crosses with hypomethylated gametes (Adams et al., 2000). Furthermore, in maize and sorghum extensive hypomethylation was shown in the endosperm compared to embryo and leaf tissue (Zhang et al., 2007; Lauria et al., 2004). Although we did not observe differences in the distribution of methylated cytosines in specific regions of heart stage embryos or endosperm in *Hieracium* apomictic and sexual plants we cannot rule out that these differences might occur in specific loci of specific cells. Localization the expression of *HCMT* in developing seeds might further elucidate where *HCMT* function is important in developing apomictic seeds.

**Function of LOA**

The genetic analysis of apomictic mutants has enabled us to further understand the functions of the LOA locus, the locus that enables plants to produce unreduced embryo sacs. Apomictic plants that have lost this ability still undergo meiosis normally and produce reduced embryo sacs that can still retain the capacity for parthenogenetic development. This implies that the LOA locus most likely carries with it a function to suppress the sexual developmental pathway, either as a separate gene function or as part of a gene function that enables somatic cells to enter a gametophytic program. Cytological characterization of the progeny derived from a *loa* mutant also revealed that embryo formation was not normal. Both fertilization independent and fertilization-dependent embryo development was severely affected. Because this mutant was derived from a WT apomictic background, we conclude that
meiotic recombination exposed progeny to embryo lethal mutations that have accumulated in the genome of the apomictic plant. This implies that apomixis might have evolved as a mechanism to prevent sterility and lethality caused by deleterious mutations. Under this scenario, whatever the mechanism that enables the differentiation of a somatic cell, must also act, or be closely linked to a mechanism to prevent meiosis. In fact, the occurrence of diplospory as a mechanism to prevent meiosis is in agreement with this hypothesis. In diplosporous plants, the production of an unreduced gametophyte occurs through the reversion of megaspore mother cell meiosis, reconstitution of a somatic nuclei, and progression of the gametophytic program without meiotic recombination. Therefore, it appears that in diplosporous plants, the molecular mechanism that allows gametophyte formation by a somatic cell is also linked to avoidance of meiosis. Due to the mode of action of LOA, suppression of meiosis, it is also conceivable that introgression of this trait to other plants, even closely related species, could cause mis-regulation of initial events of megaspore mother cell differentiation and ovule abortion. This could explain the transmission distortion of LOA in a segregating population (Catanach et al., 2006). In fact, transmission distortion appears to be associated with the inheritance of the apomictic trait in several species (Grimanelli et al., 1998; Roche et al., 2001). It is likely that this represents a common mechanism in the evolution of apomixes in different phylogenetic groups.

Conclusion

This thesis has for the first time analysed the function of the FIS-class genes in apomictic reproduction and characterized their role in autonomous seed initiation. The
FIS gene *FERTILIZATION INDEPENDENT ENDOSPERM (FIE)* has been recruited differently in *Hieracium* relative to the model plant Arabidopsis. HFIE is not part of a repressive complex in the female gametophyte of sexual or apomictic *Hieracium* plants, as it has been shown for Arabidopsis FIE. HFIE plays a fundamental role in apomictic and sexual seed development, and this post-zygotic role is also retained in Arabidopsis FIE. This suggests that this could be the primary function of FIE, with the pre-fertilization repressive role being specifically recruited in Arabidopsis and the Brassicaceae. We also propose a model in which the lack of a HFIE repressive role of HFIE might have facilitated the appearance of autonomous endosperm development. The tight linkage between autonomous endosperm and embryo shown in this work might be related to our findings of HFIE function. This opens new questions as to whether autonomous endosperm linkage to autonomous embryo relies on the absence of a function FIS complex with repressive roles (Figure 1) and can be explored by expanding FIE protein interaction studies in apomictic plants that require fertilization of the central cell for endosperm formation, pseudogamous apomicts. These results encourage further study of the function and diversity of PcG complexes in seed development and their role in facilitating autonomous endosperm.

The transfer and use of apomixis technology in crops will depend on further gene identification efforts and future functional analysis of these genes. This work has shown that the *Hieracium* model system is a powerful tool for the developmental study of these genes. Together with the use of the apomictic mutants that were characterized in this work, future studies can focus on the use of these mutants to assist the identification of LOA and LOP and complement the individual components of genes related to apomixis initiation or autonomous seed initiation.
Appendix 1. Supplementary data for Chapter 2.

Supplementary Figure to Chapter 2. Revised amino acid alignment of FIE proteins, modified from Tucker, M. (2003). The sequence of HFIE from apomictic D3 and apomictic C4d and sexual Hieracium plants were aligned and shown to have only one amino acid change (vertical arrowhead). Comparison with FIE from Arabidopsis (AtFIE) showed some differences and these are important to the structure of intervening loops in the beta-propeller structure. The differences highlighted in the model are pointed out with one (*) and two (**)
**Supplementary Table Chapter 2.** Primers used for yeast two hybrid constructs. Restriction used for cloning are in italic. The in-frame ATG is shown in bold, except for HRB, where fusion was with the second amino acid of the protein

<table>
<thead>
<tr>
<th>Primer (RE site)</th>
<th>Vector</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>AtFIE lexA F (BamHI) AtFIE lexA R (XhoI)</td>
<td>pLexA</td>
<td>tacggatccGAATTCGAAGATAACCTTAGGG cagctcagCTACTTGGAATACGCGGCTCCAG</td>
</tr>
<tr>
<td>HFIE lexA F (BamHI) HFIE lexA R (XhoI)</td>
<td>pLexA</td>
<td>gaagggattcaATTCGAAGATAACCTTAGGGG tacctcagTGAGGTGCAACTGATCTCCCA</td>
</tr>
<tr>
<td>AtRB AD F (BglII) AtRB AD R (XhoI)</td>
<td>pB42AD</td>
<td>catagatctCTATGGGAAGAAGTTCCAGCCCTCCAG gaactcagCTATGAATCTGGTGCTCCG</td>
</tr>
<tr>
<td>HRBR AD F (NcoI) HRBR AD R (XhoI)</td>
<td>pB42AD</td>
<td>gcacggatGACCTTCAGGGCGAATACCC gcgcctcagTTCAACTCTTTGACCCTTGATGGA</td>
</tr>
<tr>
<td>AtMSI1 AD F (EcoRI) AtMSI1 AD R (NcoI)</td>
<td>pB42AD</td>
<td>cttggaattcATGGGAAGAGCAGAGAGGA gaaccatggCTAAAGAGCTTTTGATGGGTT</td>
</tr>
<tr>
<td>HMSI1 AD F (EcoRI) HMSI1 AD R (BglII)</td>
<td>pB42AD</td>
<td>ccaggaattcATGGGAAGAGCAGAGAGGA ccaggaattcCTACAGCACCTCGAGAGTCTCTC</td>
</tr>
<tr>
<td>AtMEA AD F (EcoRI) AtMEA AD R (XhoI)</td>
<td>pB42AD</td>
<td>ccaggaattcATGGGAAGAGAACCATGAGG tacctcagCTAACGAGCTGGACGGCTCC</td>
</tr>
<tr>
<td>AtCLF AD F (EcoRI) AtCLF AD R (BglII)</td>
<td>pB42AD</td>
<td>ccaggaattcATGGCGTCAGGAAGCTTGTGCTCC</td>
</tr>
<tr>
<td>AtCLF AD R (BglII)</td>
<td>pB42AD</td>
<td>ccaggaattcCTAAAGGAGCTTTTGATGGGTCTCTC</td>
</tr>
</tbody>
</table>
Appendix 2.

Primers List – Chapter 5.

**HMET RACE primers**

HMET 3R1 - AGCCCTTGCTTCACCTTCTTG  
HMET 3R2 - GGAAACGAGCATTCACTCTGG  
HMET 5R4 - GACTTCAGCATTTCAACAGGT  
HMET 5R3 - CTTTGATGGCTTTCCGAGTCTA  
HMET 5R2 - TACAATAGCGGCGGTTTGATGAG  
HMET 5R1 – AAGACCTGACACAAAAAGCACAC

**HMET genomic isolation**

HMET gF1- AGCTTAACACTTCTGTTCAC  
HMET gR3 - CGCCAGGTCAGCAAAAAATGCC  
HMET gR4 - GCCTTTTGTACAGGTAAAATTC  
HMET 127 rev - GGTTTTTCGTTGTGTTTTGTCT

**HCMT RACE primers**

CMT 3R1 - ATGATCATCGCCCACTTTAAACT  
CMT 3R2 - TGGTCCTGTATTATGCAATGAC  
CMT 3R3 - AAAACCATTTGGACGACTATGG  
CMT 5R1 - AAGAAGCCTTTAGCGAATCTGA  
CMT 5R2 - CGATGTCCATGTAGACAAACGAG  
CMT 5R3 – TGTTTCTGAACCGGTTGAATC

**HDME RACE primers**

HDME 3R1 – GATCCTTGCTCATACCTCCTTG  
HDME 3R2 – AGGACAATGTAGCTCCAAGAA  
HDME 5R1 – CAGGGGTACCCATCCAAG  
HDME 5R2 – TGGTGAAGTGTAGTGACGTACGCA  
HDME 5R3 - GCCGTACACACTCCACACTTTTCA  
HDME 5R4 - TGGTCTTTTCACAAGCCGTATT
References


Kim, H. H., Y; Park, J; Park, M; Park, M; Kim, MD; Kim H; Lee, MH; Moon, J; Lee, I & Kim, J (2004). "A genetic link between cold responses and flowering time through FVE in Arabidopsis thaliana." Nature Genetics 36: 167-171.


Lagace, M., Chantha, S., Major, G. and Matton, D. P. (2003). "Fertilization induces strong accumulation of a histone deacetylase (HD2) and of the other chromatin-remodelling proteins in restricted areas of the ovules." Plant Molecular Biology 53: 759-769.


Peaston, A., Evsikov, A., Graber, J., de Vries, W., Holbrook, A., Solter, D. and
Knowles, B. (2004). "Retrotransposons regulate host genes in mouse oocytes

Penterman, J., Zilberman, D., Huh, JH., Ballinger, T., Henikoff, S. and Fischer, RL.
DNA demethylation in the Arabidopsis genome. Proc Natl Acad Sci U S A
104: 6752-6757

of the C-terminal WD-40 repeat of the human Groucho/TLP1 transcriptional


maintenance human DNA (cytosine-5) methyltransferase and modulates its
activity. EMBO Journal 21, 779-788.

retinoblastoma-binding protein related to a negative regulator of Ras in yeast


"DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses

Rodrigues, JCM and Koltunow, AM. (2004). Epigenetic aspects of sexual and asexual

"Demethylation-induced developmental pleitropy in Arabidopsis." Science

Research 32, W321-W326.


