LOCATION AND EXPRESSION OF GENES RELATED TO THE CYTOPLASMIC MALE STERILITY SYSTEM OF *BRASSICA NAPUS*

BY

RACHEL GWYNETH GEDDY

A thesis submitted to the Faculty of Graduate Studies and Research of McGill University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Biology Department
McGill University,
Montreal, Quebec, Canada
January, 2006

© Rachel Gwyneth Geddy, 2006

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-25154-6
Our file Notre référence
ISBN: 978-0-494-25154-6

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Dedicated to my father, David Geddy, the best, kindest and gentlest man who ever lived.

"See first, think later, then test. But always see first.
Otherwise you will only see what you were expecting."

From **So Long, and Thanks for All the Fish** by Douglas Adams

ABSTRACT

Cytoplasmic male sterility (CMS) is a maternally inherited defect in the production of pollen, the male gamete of the flower. This sterility can be suppressed by nuclear Restorer of Fertility (Rf) genes that normally downregulate the expression of the CMS-associated novel mitochondrial gene. In *Brassica napus*, *nap* CMS and *pol* CMS are associated with related chimeric mitochondrial genes *orf222* and *orf224*, respectively. CMS in both *nap* and *pol* is associated with a polar loss of locule development, loss of synchronous locule development and clumping of sporogenous tissue away from the tapetal cell layer, as well as secondary effects on petal and bud formation. In *nap* CMS, early accumulation of *orf222* transcripts in the locule regions of developing anthers is associated with sterility, while the absence of *orf222* transcripts from the locules is associated with fertility restoration. Accumulation of novel antisense transcripts of *atp6* in a cell specific manner which matches that of sense transcripts of *orf222* and *atp6* in *nap* CMS anthers may be indicative of a post-transcriptional regulatory mechanism associated with CMS in flower buds.

Restoration of fertility in *Brassica napus nap* and *pol* CMS is associated with nuclearly encoded genes *Rfn* and *Rfp*, respectively. These restorers are very closely linked to one another, and may be allelic. Further efforts to isolate *Rfp* have narrowed the genomic region to approximately 105 kb of a syntenic region in *Arabidopsis thaliana*. Cosmid clones isolated from a library of *Brassica rapa* genomic DNA introgressed with *Rfp* have been successfully sorted into contigs

through the application of the amplified fragment length polymorphism technique. The region to which *Rfp* is mapped is syntenic to a region of *Arabidopsis* DNA that is a duplication of a second location at the 23 megabase region of chromosome 1 of that genome. This region contains pentatricopeptide (PPR) motif-encoding genes that are highly related to other restorers of fertility of other species. By inference, *Rfp* from *Brassica napus* may encode PPR motifs. The PPR genes related to these previously characterized restorers of fertility are often found alongside the restorer genes existing as mini-clusters of several PPR-encoding genes. This is likely caused by selective pressure acting on PPR-encoding genes that resulted in diversification and multiplication of these genes. In addition, the PPR genes of this duplicated region are not syntenically located, whereas the non-PPR-encoding genes maintain their syntenic locations. The same is true for orthologous comparisons between *Arabidopsis* and other plant species. PPR genes are therefore malleable and capable of alteration in response to changing environmental pressures, such as the evolution of sterility inducing genes.

RÉSUMÉ

La stérilité mâle cytoplasmique (CMS) est un défaut de production de pollen spécifié par des gènes uniques mitochondrial. Cette stérilité, qui est d'hérédité maternelle, peut être dominée par un gène de restauration de fertilité nucléaire (Rf) qui peut réduire l'expression du gène mitochondrial. Deux types de CMS dans *Brassica napus, nap* et *pol*, sont associés avec des gènes chimères semblables, *orf222* et *orf224*, respectivement. La CMS de *nap* et *pol* font des changements à la structure du bouton floral, incluant l'absence de développement des locules d'une manière polaire, la perte de synchronisme du développement des locules, l'agrégation des cellules sporogénaires et des effets secondaires sur la formation de structure des pétales et des boutons floraux. Dans la CMS *nap* le produit de la transcription de l'*orf222* cumule tôt dans les régions où les locules se développent. Quand la CMS *nap* est dominée par la restauration de fertilité, les produits de la transcription de l'*orf222* sont absente des locules. L'accumulation de l'ARN antisens unique de *atp6* d'une manière spécifique cellulaire qui mimète l'ARN sens de *orf222* et *atp6* dans des anthères CMS *nap* peut indiquer un mécanisme de régulation post-transcriptionel associé avec la CMS des boutons floraux.

La restauration de fertilité dans le CMS *nap* et *pol* de *Brassica napus* est associée avec deux gènes nucléaires, *Rfn* et *Rfp*, respectivement. Ces gènes Rf sont liés très étroitement l'un à l'autre, et peuvent être alléliques. Des efforts continuent pour isoler *Rfp*; maintenant la région où ils se trouvent est à presque 105 kb d'une région synténique de *Arabidopsis thaliana*. Des clônes ont été isolés d'une

bibliothèque de cosmides contenant d'ADN de *Brassica rapa* avec *Rfp*. Ces clônes ont été organisés dans des contigs par l'utilisation de la technique polymorphisme de longueur des fragments amplifiés (AFLP). La region où se trouve *Rfp* est syntenique à une région d'ADN de *Arabidopsis* qui elle-même sert de duplication d'une autre région du chromosome 1 de ce génome. Cette région d'*Arabidopsis* contient des gènes qui codent des motifs pentatricopeptides (PPR) et qui sont très similaires aux autre gènes Rf des autres espèces. Par inférence, *Rfp* de *Brassica napus* peut aussi coder pour des motifs PPR. Les gènes PPR de restauration de fertilité sont fréquemment trouvés dans un groupe de gènes PPR similaires. La source de ces groupes de gènes PPR est probablement la pression sélective qui résultent dans la diversification et multiplication de ces gènes. On trouve aussi que les gènes PPR de cette région double ne sont pas synténique, mais que les autres gènes non-PPR maintiennent leur ordre synténique. Cela est aussi vrai dans les comparaisons entre orthologues de *Arabidopsis* et des autres plantes. Les gènes PPR sont donc malléables et peuvent changer en réponse aux changements des pressures sélectives environnementales, par exemple, l'évolution des gènes causent la stérilité.

ACKNOWLEDGEMENTS

Students who continue to the end of a Ph.D. have never done so alone; they have done it with the support and encouragement of those who shared the burden with them and carried them through this most challenging of life experiences. First, to Dr. Gregory Brown, my supervisor, I thank you for the projects you shared with me and the faith which you placed in me and my ability to complete them. Also, for your generous financial support, I thank you. I also thank you for your guidance, advice and diplomacy. Your support has molded me into the independent, critical person I am today. Thanks also to NSERC for funding provided for my work through an NSERC Discovery Grant made to Dr. Brown.

To my wonderful parents you both helped me and supported me in more ways than anyone can count. My father, to whose memory this thesis is dedicated, passed away during the course of this degree, but his hard work and persistence in the face of adversity showed me how strong I could be. He gave me everything he could and was available whenever I needed him. I miss him sorely. My mother is an inspiration to everyone around her. She is my friend, my mentor and my solid rock of reason when all else is in chaos. With their support, this thesis came to fruition.

To my husband Dan, you are my life. You kept me whole in many ways. Pulling me together when I fell apart, cooking wonderful meals for me, showing me there is more to life than the lab, sharing in my thoughts about work and life. Thank you for your unwavering support and help (especially with editing!). Our travels together have just begun. Here's to the future!

My friends have all provided a safe haven in which to rest. Céline Mouchel helped me to think outside the laboratory test kit to expand my scientific and personal experiences. She helped me with talks to perfect them for presentations,

and salved my wounded ego with beer when required. To Nabil Elrouby, my thanks for all the scientific and personal input over the years. Imoinda Romain, Eaner Xu, Erica Dudgeon, Sylvia Gajewski, Ivan Pasic, Anton Marinov, Laura Istanboulian, Andrea Nicholls and the Hockey Gang, you all gave my mind a good rest each time I went home. To my extended family, especially Uncle Don and Aunt Connie and Jerry and Stefanie Gajewski, I thank you for your love and support.

To my supervisory committee members throughout the years, Drs. Candace Waddell, Paul Lasko, Christian Hardke and Tamara Western, my thanks for your input and advice. To the administrative staff, especially Susan Bocti, our wonderful graduate student secretary without whom we would all miss our deadlines, I thank you. To Becky Dolan, my thanks for your friendship and support. To my unexpected friend Colleen Parish, your last minute kindness meant the world to me. To Céline Frechina, Nadim Maghzal and Ryan Nantel-Smith, my thanks for your help. To Gregoire Morisse, thanks for being there. My thanks also go to my lab mates over the years, including Laetitia Mahé, whose work was published with mine. I also thank the staff at the Genome Centre of Quebec and DNALandmarks for their sequencing work and help on my project. Thanks also to Jenny Piché for her help in editing my Résumé.

Words have the power to spur you on to greater achievements. A lab supervisor of mine once told me "Don't be afraid of hard work". Through this thesis I have looked hard work straight in the eye and challenged it. My thanks go to everyone who inspired me to greater things.

CONTRIBUTIONS OF THE AUTHORS

Chapter 1 – This chapter was written by Rachel Geddy and edited by Dr. Gregory G. Brown.

Chapter 2 - This work has been published in the Plant Journal according to the citation stated below. The manuscript presented is co-authored by Laetitia Mahé, whose work is presented in the results section entitled "Altered regulation of the APETALA3 promoter in CMS versus fertile *B. napus*", and by Dr. Gregory G. Brown in his capacity as research supervisor.

Geddy, R., Mahé, L. and Brown, G.G. (2004) Cell-specific regulation of a *Brassica napus* CMS-associated gene by a nuclear restorer with related effects on a floral homeotic gene promoter. Plant J. **41**, 333-345.

Chapter 3 - This chapter was written by Rachel Geddy and edited by Dr. Gregory G. Brown.

Chapter 4 - Initial experimental work, presented in the results section under the heading "Background work", was performed by several scientists, in particular Natasa Formanov and Rachel Stollar. Experiments performed during the course of this thesis are included in the results section under the heading "Current Work", and was assisted by Céline Frechina, Nadim Maghzal and Ryan Nantel-Smith. This chapter was written by Rachel Geddy and edited by Dr. Gregory G. Brown.

Chapter 5 - This work is being prepared as manuscript that will be submitted for publication. The manuscript presented is co-authored by Dr. Gregory G. Brown in his capacity as research supervisor.

TABLE OF CONTENTS

ABSTRACT.....I

RESUME.....III

ACKNOWLEDGEMENTS.....V

CONTRIBUTIONS OF THE AUTHORS.....VII

TABLE OF CONTENTS.....VIII

LIST OF FIGURES.....XI

LIST OF TABLES.....XIII

LIST OF ABBREVIATIONS.....XIV

CHAPTER 1: LITERATURE REVIEW.....1

INTRODUCTION.....2

I. THE STAMEN.....3

II. GENES AFFECTING STAMEN DEVELOPMENT.....6

III. CYTOPLASMIC MALE STERILITY.....11

IV. RESTORERS OF MALE FERTILITY.....18

V. *BRASSICA NAPUS*.....23

VI. EXPERIMENTAL GOALS, APPROACH AND RATIONALE.....27

REFERENCES.....28

CHAPTER 2: CELL-SPECIFIC REGULATION OF A *BRASSICA NAPUS* CMS-ASSOCIATED GENE BY A NUCLEAR RESTORER WITH RELATED EFFECTS ON A FLORAL HOMEOTIC GENE PROMOTER.....40

INTRODUCTION.....41

MATERIALS AND METHODS.....45

RESULTS.....52

NORMAL ANTHER DEVELOPMENT IN *B. NAPUS*.....52

DEVELOPMENT OF CMS ANTHERS.....56

MITOCHONDRIAL TRANSCRIPT EXPRESSION IN CMS AND RESTORED BUDS.....60

ALTERED REGULATION OF THE APETALA3 PROMOTER IN CMS VERSUS FERTILE *B. NAPUS*.....67

DISCUSSION.....72

NOVEL MORPHOLOGICAL ASPECTS OF CMS IN *BRASSICA*.....72

SPATIAL AND TEMPORAL REGULATION OF MITOCHONDRIAL TRANSCRIPT ACCUMULATION.....72

ORF222 TRANSCRIPT ACCUMULATION AND THE CMS PHENOTYPE.....75

REFERENCES.....81

CHAPTER 3: CELL-SPECIFIC ANTISENSE TRANSCRIPTS OF A NUCLEAR AND A MITOCHONDRIAL GENE ENCODING SUBUNITS OF THE ATP SYNTHASE COMPLEX IN *BRASSICA NAPUS*.....87

INTRODUCTION.....88

MATERIALS AND METHODS.....91

RESULTS.....95

SENSE AND ANTISENSE TRANSCRIPTS OF A MITOCHONDRIALLY-ENCODED SUBUNIT OF THE ATP SYNTHASE IS EXPRESSED IN A CELL-SPECIFIC MANNER DURING FLORAL DEVELOPMENT.....95

ANTISENSE ATP6 ARE QUALITATIVELY DIFFERENT FROM THE CORRESPONDING SENSE TRANSCRIPTS.....98

DISCUSSION.....103

REFERENCES.....106

CHAPTER 4: TOWARDS IDENTIFICATION OF THE RESTORER OF FERTILITY FOR POLIMA CMS.....110

INTRODUCTION.....111

MATERIALS AND METHODS.....114

RESULTS.....119

BACKGROUND WORK.....119

CURRENT WORK..... 120

DISCUSSION..... 139

REFERENCES..... 141

CHAPTER 5: PENTATRICOPEPTIDE REPEAT ENCODING GENE LOCATION IN PLANTS IS NOT CONSERVED IN ORTHOLOGOUS AND PAROLOGOUS COMPARISONS..... 144

INTRODUCTION..... 145

MATERIALS AND METHODS..... 149

RESULTS..... 151

LOCATIONS OF PPR GENES ARE HIGHLY VARIABLE BETWEEN CO-LINEAR REGIONS OF ARABIDOPSIS AND BRASSICA OR RAPHANUS GENOMES..... 151

VARIATION IN PPR GENE LOCATION BETWEEN TWO PARALOGOUS GENE REGIONS ON CHROMOSOME I OF ARABIDOPSIS..... 161

PPR GENES IN THE RADISH RFO REGION HAVE BEEN SUBJECT TO DIVERSIFYING SELECTION..... 175

DISCUSSION..... 181

REFERENCES..... 186

GENERAL CONCLUSIONS AND SUMMARY..... 190

ORIGINAL SCHOLARSHIP AND DISTINCT CONTRIBUTIONS TO KNOWLEDGE..192

APPENDIX 1: COMPLIANCE CERTIFICATES

LIST OF FIGURES

Figure 1.1 Anatomy of the *Brassica napus* flower..... 4-5

Figure 1.2 Cartoon representation of CMS-associated mitochondrial gene regions from various plant species..... 12-13

Figure 1.3 Relationship between different *Brassica* species and their haploid chromosome number (N)..... 25-26

Figure 2.1 Development of fertile (a-f) and CMS (g-l) anthers in *Brassica napus* 53-54

Figure 2.2 Polarity of locule formation, locule fusion and petal abnormalities in CMS floral buds..... 57-58

Figure 2.3 *In situ* hybridization of floral bud tissue with *orf222* sequences..... 61-62

Figure 2.4 *In situ* hybridization of floral bud tissue with antisense cytochrome *b*, *atp6* and ATP synthase gamma subunit probes..... 65-66

Figure 2.5 AP3 promoter activity in CMS and male-fertile flowers as measured by expression of an AP3::GUS construct..... 69-70

Figure 3.1 *In situ* hybridization of *Brassica napus nap* CMS floral bud tissue with *atp6* and cytochrome *b* sequences in the microspore mother cell stage..... 96-97

Figure 3.2 Gel blot analysis of mitochondrial RNA with sense and antisense transcripts of the *atp6* gene..... 99-100

Figure 4.1 Cartoon depicting the placement of probes with reference to the syntenic region of *Arabidopsis*..... 121-122

Figure 4.2 Polymorphism P2 of the *Rfp* region used to identify linked cosmids from a *Brassica rapa* library introgressed with *Rfp*..... 123-124

Figure 4.3 Single Nucleotide Polymorphism of the *Rfp* region used to identify linked cosmids from a *Brassica rapa* library introgressed with *Rfp*..... 126-127

Figure 4.4 Restriction Fragment Length Polymorphism of the *Rfp* region used to identify linked cosmids froma *Brassica rapa* library introgressed with *Rfp*..... 129-130

Figure 4.5 Amplified length polymorphism of selected cosmids..... 134-135

Figure 5.1 Predicted genes in *Brassica* and the syntenic segment of the *Arabidopsis* genome..... 152-153

Figure 5.2 Predicted genes in two regions of *Brassica napus* and the syntenic segment of the *Arabidopsis* genome..... 155-156

Figure 5.3 Predicted genes in radish and the syntenic segment of the *Arabidopsis* genome..... 159-160

Figure 5.4 Comparison of two *Arabidopsis* duplicated regions on chromosome 1 163-164

Figure 5.5 Phylogenetic relationship between PPR proteins of the 4.3 and 23 megabase region of *Arabidopsis* chromosome 1 and PPR genes found elsewhere in the *Arabidopsis* genome..... 171-172

Figure 5.6 Alignment of protein coding sequences of PPR encoding restorer of fertility genes with PPR genes of the 4.3 and 23 megabase regions..... 176-177

LIST OF TABLES

Table 4.1 Clones recovered with their respective probes and anlysed by the Amplified Fragment Length Polymorphism technique..... 131-132

Table 4.2 Putative contigs found among cosmids analysed by the Amplified Fragment Length Polymorphism technique 136-137

Table 5.1 Conservation of genes within two regions of *Arabidopsis* chromosome 1..... 165-166

Table 5.2 Relatedness of genes of the 4.3 and 23 Mb syntenic regions of *Arabidopsis*..... 167-168

Table 5.3 Structural annotation and targeting of PPR-encoding proteins from various plants..... 173-174

Table 5.4 Synonymous and nonsynonymous nucleotide substitution in pairwise comparisons of sequences from radish and *Arabidopsis*..... 179-180

LIST OF ABBREVIATIONS

+	Similarity
AFLP	Amplified fragment length polymorphism
AG	*Agamous*
aox	Alternative oxidase
AP	*Apetala*
At	*Arabidopsis thaliana*
ATP	Adenosine triphosphate
ATPase	Adenosine triphosphosphate synthetase
B.	*Brassica*
bp	Basepairs
BC	Backcross
C	Celcius
Cam	Campestris
CMS	Cytoplasmic male sterility
cob	Cytochrome *b*
COP	Convergent overlapping gene pairs
cox	Cytochrome oxidase
cv.	Cultivar
DNA	Deoxyribonucleic Acid
dsRNA	Double stranded RNA
Fil	Filamentous flower
g	Gram
g	Gravity
GUS	Beta-glucuronidase
H	Hour
I	Identity
Ka	Rate of non-synonymous nucleotide substitution
Kb	Kilobasepairs
Ks	Rate of synonymous nucleotide substitution
LRR	Leucine rich repeat
LFY	*Leafy*
MADS	MCM1, Agamous, Deficiens and Serum Response Factor
Mb	Megabase
Mg	Milligram
µg	Microgram
µL	Microlitre
µm	Micrometer

min.	Minute
miRNA	micro RNA
mL	Millilitre
mM	Millimolar
mm	Millimeter
mRNA	Messenger RNA
mt	Mitochondrial
nad/NADH	Nicotinamide adenine dinucleotide
NAD(P)H	Nicotinamide adenine dinucleotide phosphate
nap	Napus
NAT	Natural antisense transcript
NBS	Nucleotide binding site
NCBI	National Center for Biotechnology Information
NCS	Non-Chromosomal Stripe
Ogu	Ogura
ORF	Open reading frame
PCR	Polymerase Chain Reaction
Pi	*Pistillata*
pol	Polima
PPR	pentatricopeptide repeat
R gene	Disease resistance gene
RdRP	RNA dependent RNA polymerase
Rf	Restorer of fertility
RFLP	Restriction fragment length polymorphism
Rfn	Restorer of napus fertility
Rfp	Restorer of polima fertility
rps	Ribosomal protein subunit
RNA	Ribonucleic Acid
RNAi	RNA interference
ROS	Reactive oxygen species
RT-PCR	Reverse Transcriptase Polymerase Chain Reaction
S	Sulfur
Sep	*Sepellata*
TPR	Tetratricopeptide repeat
UFO	Unusual floral organs
UTP	Uridine triphosphate
w/v	Weight per volume
X-Gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside

CHAPTER 1

LITERATURE REVIEW

Introduction

Cytoplasmic male sterility (CMS) is a maternally inherited failure to produce functional pollen and has been correlated with the expression of novel, often chimeric mitochondrial genes (Hanson and Bentolila, 2004). Nuclear-encoded restorer of fertility, or Rf, genes can suppress the male sterile phenotype by downregulating the expression of the novel, CMS-inducing mitochondrial gene. The oilseed rape species *Brassica napus* contains two endogenous, CMS conferring cytoplasms, *nap* and *pol*, which are restored by their respective restorer genes *Rfn* and *Rfp*; these restorers act to reduce levels of novel ORF-containing mitochondrial transcripts (Brown, 1999).

Several Rf genes have recently been characterized and most have been found to encode a protein containing pentatricopeptide (PPR) repeats (Hanson and Bentolila, 2004). Although significant progress has been made in identifying the genes that specify CMS and fertility restoration, little is understood about how expression of these genes influences the physiological, developmental or biochemical processes such as how pollen formation is disrupted. Pollen development can be disrupted at any of several different stages, depending on the plant species and CMS cytoplasm. In all cases, CMS is thought to be related to mitochondrial dysfunction resulting from the expression of abnormal mitochondrial genes. In many CMS systems, vegetative plant growth and most aspects of flower development, including ovule development and female fertility, are unaffected or only slightly perturbed (Budar and Pelletier, 2001). It is unclear how mitochondrial dysfunction can have such diverse effects on pollen and stamen development without otherwise affecting plant development and function. It is possible that male sterility related phenotypes result from the

interplay of tissue and cell-specific regulatory processes involving stamen development, mitochondrial gene expression, and mitochondrial function.

I. The Stamen

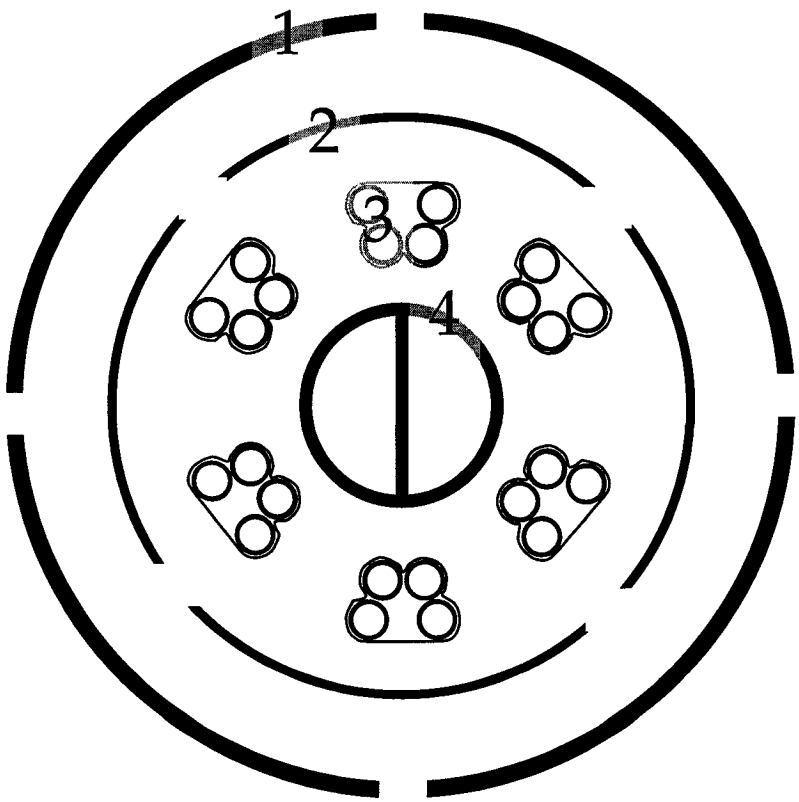
Many angiosperms, or flowering plants, reproduce sexually, with male and female reproductive organs found in the flower. The male reproductive organ is known as the stamen and is made up of a long filament, through which water and nutrients are transported, and the anther, where the male gametes develop. The female reproductive organ is termed the carpel, and is where ovules develop and in which fertilization and seed development take place. In approximately 96% of angiosperms, the flower contains both male and female organs (Briggs and Walters, 1997). In fertile *Brassica napus*, the filament positions the mature anther above the stigmatic surface of the carpel where pollen, released through dehiscence of the anther, can lodge, germinate, and fertilize the ovules (Briggs and Walters, 1997).

Flowers commonly have a structure made up of concentric rings of organs termed whorls (Figure 1.1a). In *Brassica* and allied genera the outermost whorl, whorl 1, is comprised of four sepals which are green leaf-like structures covering the developing flower bud. In *Brassica*, the second whorl consists of four petals; the petals and sepals together comprise the perianth. The next whorl (whorl 3) consists of two lateral (short) and four medial (long) stamens. The innermost whorl (whorl 4) comprises the female reproductive organ, which, in *Brassica*, is composed of two fused carpels.

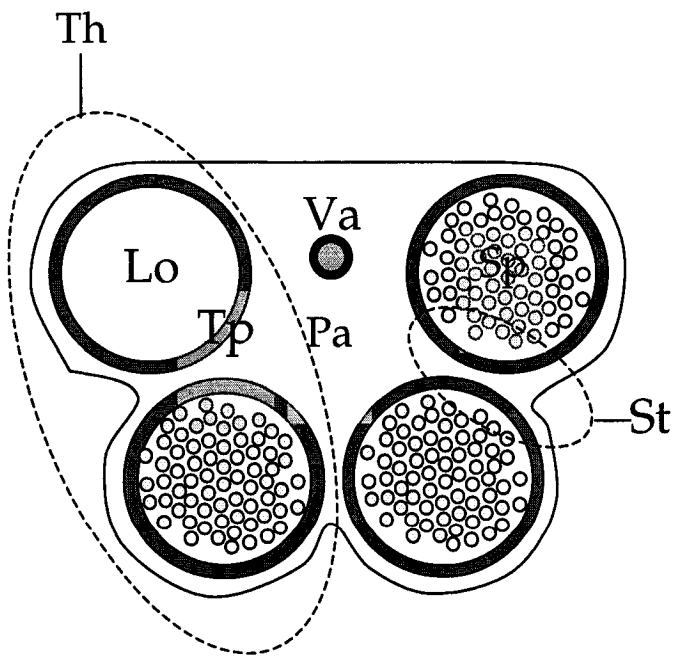
Figure 1.1 Anatomy of the *Brassica napus* flower

a) Cartoon representing a cross section through a *Brassica napus* flower. Whorl 1 comprises four sepals, whorl 2 comprises four petals, whorl 3 contains six anthers and whorl 4 is formed of two fused carpels.
b) Cartoon representing anther anatomy. Parenchyma (Pa) is the connective tissue which holds the four locules (Lo) together. The locules contain the growing sporogenous cells (Sp) and are lined by the nurse cell layer called the tapetum (Tp). The vasculature (Va) runs up the centre of the parenchyma and is the central extension of the filament. The two symetrical pairs of locules are called theca (Th). When pollen has developed to maturity, the stomium (St) breaks down through apoptosis releasing the pollen from the anther.

(a)



(b)



Individual flowers develop from the floral meristem. Floral organs are initiated in sequence, starting with the sepals followed by the stamens, petals and carpels (Polowick and Sawhney, 1986). This order is the same in *Arabidopsis* (Smyth *et al.*, 1990). Petal growth is slow after initiation, but accelerates as the anther starts to develop (Smyth *et al.*, 1990). The anther itself is composed of four lobes called locules, two on the abaxial (sepal) side and two on the adaxial (carpel) side (Figure 1.1b). These are held together by connective tissue and there is a central extension of the filament which houses the vasculature. Between each symmetrical pair of lobes (called the theca) there is the stomium, which breaks down through apoptosis to release pollen at maturity. The tapetum is the cell layer lining the locules which acts as a nutritive layer for the developing sporogenous cells (reviewed by Goldberg *et al.*, 1993). A detailed description of sporogenous cell development is provided in Chapter 2.

II. Genes affecting stamen development

Floral development follows a well studied series of events. Many of the genes controlling floral morphogenesis belong to a widespread family of transcription factors containing MADS-box sequences (**M**cm1, **A**gamous, **D**eficiens and **S**erum response factor were the first four members discovered). MADS-box transcription factors play an important role in floral patterning. Floral induction, the change from vegetative to reproductive development, is achieved through the action of meristem identity genes, notably *LEAFY* (*LFY*) and *APETALA1* (*AP1*) in *Arabidopsis*. *LEAFY*, a non MADS-box transcriptional activator of MADS-box containing *AP1*, is expressed uniformly in primordial floral tissue (Schultz and Haughn, 1991). Early constitutive expression of *LEAFY* results in early floral induction and ectopic expression of *AP1* (Parcy *et al.*, 1998). However, the

absence of *LFY* expression only results in the delayed expression of *AP1*, indicating that other activators also participate in the induction of *AP1* trancription (Liljegren *et al.*, 1999). *APETALA1* is initially activated throughout the floral primordium, but is later restricted to the developing sepal and petals. It serves the dual purpose of specifying floral identity early in development and acting as a class A homeotic gene later in floral development (Theissen *et al.*, 2000; Zhao *et al.*, 2001).

Once initial floral induction has taken place, individual organs are specified through several classes of homeotic genes, each with their own regulators (for review see Ng and Yanovsky, 2001 and Lohmann and Weigel, 2002). Class A homeotic genes, such as *AP1*, act in whorl 1 of the flower to specify sepal development, and work together with B-class genes in whorl 2 to define petal development. B- and C-class homeotic genes work in concert to specify stamen development in whorl 3, while C-class gene action leads to the production of carpels in whorl 4. In this model, A- and C-class genes repress each other, strictly maintaining their own zones of activity. The homeotic genes of floral development are, for the most part, transcriptionally regulated since their promotors are active where their function is required. One exception to this rule is the non-MADS gene *APETALA2*. This A-class gene is expressed uniformly in all whorls (Weigel and Meyerowitz 1994; Theissen *et al.*, 2000; Zhao *et al.*, 2001).

Stamen development involves several well studied B- and C-class genes, and their regulators. One B-class MADS-box gene, *APETALA3* (*AP3*), is activated by *AP1* (Lamb *et al.*, 2002). After this initial activation, *AP3* is autoregulatory, as is *PISTILLATA* (*PI*), another B-class MADS-box gene. *AP3* and *PI* form a heterodimer which binds to elements termed CarG boxes in the promotor of *AP3*

and are responsible for the autoregulatory function of *AP3*. These elements are also found in the PI promoter, although their role in *PI* regulation is uncertain (Riechmann *et al.*, 1996; Hill *et al.*, 1998; Tilly *et al.*, 1998). The *AP3* and *PI* gene products are post-transcriptionally activated by the MADS-box *SEPELLATA* genes; *SEPELLATA* genes have also been shown to play a role in floral determination. Evidence suggests that the quaternary structures composed of several MADS domain proteins may act in different combinations to specify the various floral organs. In the case of stamen development, *PI*, *SEP3*, *AGAMOUS* (*AG*) and *AP3* may form a unique complex not found in other whorls (Honma and Goto, 2001). It is also thought that the B-class regulatory gene *UFO* (*UNUSUAL FLORAL ORGANS*) acts to activate the activity of *AP3* in whorls 2 and 3 via the ubiquitination and subsequent degradation of the negative regulators of *AP3* transcription (Samach *et al.*, 1999).

AGAMOUS, a C-class MADS-box gene is repressed, perhaps directly, by *AP2* in whorls 1 and 2, but activated in whorl 4 by *LEAFY* and *WUSCHEL* (*WUS*), and in whorl 3 by *UFO* and *LEAFY* (Lenhard *et al.*, 2001; Lohmann *et al.*, 2001; Bomblies *et al.*, 1999). By contrast, *AG* negatively regulates A-class genes, preventing their expression in the last two whorls (Sieburth *et al.*, 1998). In addition, *AG* acts later in sporogenesis, activating *sepellata* via *sporocyteless*, a regulator of male gamete formation (Ito *et al.*, 2004), making it an important player in early stamen differentiation. Recent genome-wide analysis of *Arabidopsis thaliana* also shows that *AGAMOUS* may cause the shift from floral meristem induction to floral organ differentiation via activation of gibberrelin synthesis (Gomez-Mena *et al.*, 2005)

Many floral determination genes and their regulators have orthologs in Antirrhinum, leading to the suggestion that the ABC model of floral organ determination may apply to "most, if not all, flowering plants" (Ng and Yanofsky, 2001). In fact, a *Brassica* homolog of *AP3* was recently discovered (Pylatuik *et al.*, 2003), adding to the evidence of the conservation of floral developmental patterning between species. Recent studies have also shown that C-class gene function can be traced back to before the divergence of angiosperms and gymnosperms and has been conserved across 300 million years of evolution (Zhang *et al.*, 2004).

Stamen development is governed by an extremely complex pattern of gene expression. In tobacco it has been estimated that as many as 10000 genes are specifically expressed during stamen development (Kamalay and Goldberg, 1980; 1984). Defects in such anther specific genes can lead to physical defects resulting in the sterility of the anther, although this does not preclude the sterility-inducing potential of genes which are expressed more generally in the plant. Floral homeotic genes described above can, when misexpressed, lead to male sterility. The studies of Sanders *et al.* (1999) provide insight into the other varieties of male sterile phenotypes that can be observed in *Arabidopsis thaliana.* These authors screened a collection of T-DNA tagged mutants and a near-saturation EMS mutagenized population for mutations specifically leading to male sterility. These plants were then classified into several different phenotypic categories. Mutants in the first category had undeveloped anthers. This defect appears early in anther development and inhibits the formation of the anther proper and, thus, sporogenous tissue. Genes causing this defect are likely involved in early patterning of the anther and laying down the groundwork for subsequent cellular differentiation. For example, the gene *FILAMENTOUS FLOWER* (*FIL*)

controls inflorescence and flower architecture and, when mutated, results in a complete absence of an anther atop the filament (Komaki *et al.*, 1988).

A second category of mutant is characterized by the pollenless phenotype, in which anthers form that lack sporogenous tissue but contains otherwise morphologically normal locules. This mutation could be caused by aberrations in meiosis or defects in processes of the surrounding cell layers such as the tapetum. It could also be a result of the unexpected destruction of sporogenous tissue. For example, in *pollenless-2* mutants, both the tapetum and abnormal meiotic cell products degenerate, leaving empty locules. This differs from *pollenless3-2* in which only the defective meiotic cell products degrade (Sanders *et al.*, 1999).

The third category of male sterile mutations involves defects in the release of mature pollen from the locules of the anther. Such dehiscence defective phenotypes are likely caused by misexpression of pathways of apoptosis, or programmed cell death, in the cells of the stomium. If these cells remain intact, the locules cannot open and any pollen, even fully functional pollen, would remain trapped and lead to sterility. For example, *non-dehiscence-1* anthers fail to degenerate the stomium, but instead degenerate the connective tissue and endothecial tissue between the two locules, resulting in one large locule per theca (Sanders *et al.*, 1999).

The last category of male sterility in the plants studied is that of defective pollen. The pollen, in this case, aborts and cannot fertilize an ovule, even if all other parts of the stamen development seem normal. Defects in male reproductive development can be attributed to all manner of defects; metabolism, improper meiosis, premature cell death etc. For example, mutant *tdm1* meiotic cells

undergo a third round of meiosis with chromosome duplication, resulting in defective pollen (Ross *et al.*, 1997), and *mei1* mutants create fragmented chromosomes and micronuclei during pollen development (He *et al.*, 1996). Clearly, defects at any one of many stages of stamen development can lead to male sterility.

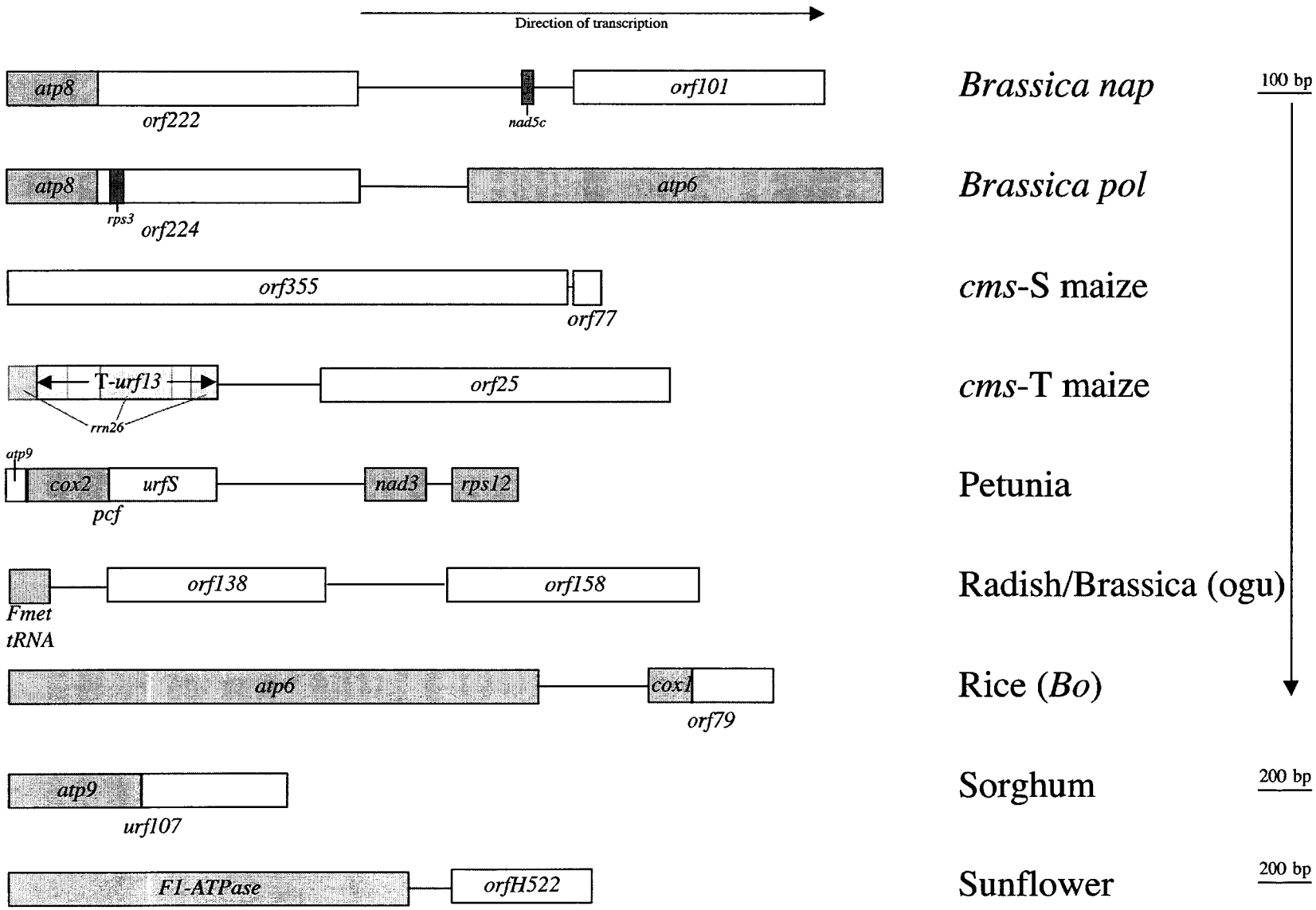
III. Cytoplasmic Male Sterility

Cytoplasmic male sterility (CMS) is a condition which results in a plants' inability to produce pollen. CMS is the most common cause of male sterility in gynodioecious plants (Lewis, 1941), and has been identified in a wide cross-section of angiosperms. The sterile phenotype is inherited through the female, and is caused by unusual genes or open reading frames (ORFs) encoded in the mitochondrial genome.

CMS-associated mitochondrial genes share several common characteristics (Figure 1.2). The structure of many of these genes is chimeric; they are composed of portions of known mitochondrial genes fused to unique, unknown sequences. Maize (*Zea mays*) has several different CMS cytoplasms, the best studied of which is the Texas, or T cytoplasm. The *cms*-T associated open reading frame T-*urf13* is homologous to portions of the mitochondrial large ribosomal RNA gene *rrn26* and is expressed using the 5′ upstream region of the gene encoding subunit 6 of the ATP synthase, *atp6* (Wise *et al.*, 1999). Similarly, the rice CMS *Bo* cytoplasm contains a portion of the gene for cytochrome oxidase subunit 1 (*cox1*) fused to unique sequences (Kadowaki *et al.*, 1990). CMS-associated genes in both sorghum and petunia both contain portions of a gene encoding subunit 9 of the ATP synthase (*atp9*; Tang *et al.*, 1996). In petunia the CMS-associated gene *pcf* also

Figure 1.2 Cartoon representation of CMS-associated mitochondrial gene regions from various plant species.

Unknown open reading frames (orfs) are often co-transcribed with other genes or orfs (shown here as boxes with intergenic regions shown as lines; shading is for clarity only). One exception shown here is from sorghum. Many of the transcriptional units are made up of portions of other genes. Two exceptions are radish *ogu* and *cms-S* from maize. All genes represented here are transcribed from left to right. The scale of the diagram is shown at the right; *Brassica nap* through rice Bo are 100bp/cm while sunflower and sorghum are 200bp/cm.


Direction of transcription
atp8
orf222
orf101
nad5c
Brassica nap
100 bp
atp8
rps3
orf224
atp6
Brassica pol
orf355
orf77
cms-S maize
T-urf13
rrn26
orf25
cms-T maize
atp9
cox2
urfS
pcf
nad3
rps12
Petunia
orf138
orf158
Fmet
tRNA
Radish/Brassica (ogu)
atp6
cox1
orf79
Rice (Bo)
atp9
urf107
Sorghum
200 bp
F1-ATPase
orfH522
Sunflower
200 bp

includes a portion of the gene for cytochrome oxidase subunit 2 (*cox2)* (Hanson *et al.*, 1999). It is speculated that these novel open reading frames may disrupt the function of the normal mitochondrial genes whose portions they have appropriated, leading to impaired oxidative phosphorylation. This, in turn, would lead to an overall deficit of energy available, and, ultimately, to male sterility (Levings, 1993).

CMS-associated genes are frequently co-transcribed with standard mitochondrial genes. The "Ogura" CMS of radish is associated with the novel mitochondrial gene *orf138* which is co-transcribed with the gene for f-Met tRNA (*trnfM)* and another open reading frame, of unknown function *orf158* (Bellaoui *et al.*, 1997). In rice, the ORF associated with the "Bo" CMS, *orf79,* is co-transcribed with *atp6* (Kadowaki *et al.*, 1990), whereas sunflower's *orfH522* is co-transcribed with the mitochondrial gene encoding the alpha subunit of the ATP synthase, *atpA* (Smart *et al.*, 1994).

In spite of these commonalities, CMS shows variation in the way it is manifested. CMS in carrot (Linke *et al.*, 1999), tobacco (Kofer *et al.*, 1991; Zubko *et al.*, 1996) and wheat (Ogihara *et al.*, 1997) leads to homeotic transformation of stamens into carpelloid structures; the elimination of the anther is the cause of the sterility and is likely caused by peturbation of floral homeotic gene function (see above section II). Homeotic transformations found in these systems may be the result of direct or indirect effects on B- and C-class homeotic gene pathways via a disruption in the interaction of the nuclear and extranuclear gene products (Linke *et al.*, 1999). By contrast, the sterility of maize CMS-S is caused by pollen abortion, and that of petunia, sunflower and maize *cms*-T is associated with premature degradation of the tapetum. In tobacco, impairment of the tapetum causes pollen

abortion (Hernould *et al.*, 1998), whereas in maize CMS-S pollen collapses at the late starch filling stage (Wen and Chase, 1999). In sunflower, abortion of young meiotic cells follows premature degradation of the tapetum and is observed to cause sterility (Smart *et al.*, 1994).

This study is focused on the analysis of CMS in the oilseed rape species *Brassica napus*. Cytoplasmic male sterility (CMS) in this plant can be conditioned by three cytoplasms called *nap*, *pol*, and *ogu* (Brown, 1999; Bellaoui *et al.*, 1997).

One type of cytoplasmic male sterility occurs in *Brassica napus* plants containing the *nap* sterility-inducing cytoplasm. Analysis of the mitochondrial genome of *nap* cytoplasm plants has revealed the presence of a rearrangement or recombination event which is associated with the sterility phenotype (L'Homme *et al.*, 1997). Analysis of this gene region shows a chimeric open reading frame inserted upstream of, and co-transcribed with, an exon of the *nad5* gene (*nad5c*), which encodes a subunit of the NADH dehydrogenase of the electron transport chain (L'Homme *et al.*, 1997). This open reading frame, called *orf*222, shows an expression pattern which differs between fertile, sterile, and restored plants. Transcripts are strongly expressed in CMS plants, but significantly reduced in abundance in fertility restored plants and absent in fertile plants (L'Homme *et al.*, 1997). The 5' portion of *orf*222 shows a high degree of sequence similarity with the *atp8* gene (also referred to as *orfB*), which encodes a subunit of the mitochondrial ATP synthase. The remaining portion of *orf*222 shows limited similarity with a set of related orfs of unknown function in the *Arabidopsis thaliana* mitochondrial genome (Brown, 1999).

The chimeric open reading frame associated with *pol* CMS is called *orf224* and it has high sequence homology to *orf222*. *orf224* is upstream of, and co-transcribed with *atp6*, which encodes a subunit of the FoF1-ATPase of the mitochondrion. Extensive work has been done to elucidate the size and origins of transcripts of this gene region (Menassa *et al.*, 1999).

The sequences of *orf222* and *orf224* share a high degree of similarity at both the nucleotide and protein levels, with 85% and 79% homology, respectively. In addition, the hydropathy profile of these two proteins is remarkably similar, suggesting that the structure of the protein produced may be unaffected by the sequence changes (L'Homme *et al.*, 1997). The question of how these CMS-associated genes act is still the subject of debate, and further analysis is needed to shed light on their function.

Another type of cytoplasmic male sterility has been found in tobacco (Chetrit *et al.*, 1992), but more closely resembles the non chromosomal stripe (NCS) mutations of maize (Lauer *et al.*, 1990; Newton *et al.*, 1990) than it does other CMS mutations. CMSI and CMSII tobacco plants contain large deletions in their mitochondrial genomes which result in a near-total loss of Complex I function of the respiratory chain. The deletions comprise the entire *nad7* gene and the upstream region of *nad1*. Also, there is a reduction in *NAD9* subunits (Pla *et al.*, 1995; Gutierres *et al.*, 1997). Loss of Complex I function is compensated by the activity of alternative NAD(P)H dehydrogenases. In addition, there is an increase in alternative oxidase activity (*aox*) which is attributed to the increased reactive oxygen species (ROS) resulting from the lack of Complex I activity (Sabar *et al.*, 2000).

Alternative oxidase activity is also enhanced in *NCS2* mutation of maize; the *NCS2* mutation results in the fusion of *nad7* and *nad4*, and reduces complex I function, affecting AOX and NAD(P)H dehydrogenase in a similar fashion as CMSI and CMSII (Marienfeld and Newton, 1994). Alternative oxidase activity is stimulated by the application of either rotenone, an inhibitor of complex I activity, or cyanide, an inhibitor of complex IV activity (Karpova *et al.*, 2002). The *NCS5* and *NCS6* mutations are deletions in *cox2*, and affect the function of complex IV of the electron transport chain (Lauer *et al.*, 1990; Newton *et al.*, 1990). NCS mutations are lethal, but can be propagated in the heteroplasmic condition, where both mutant and normal mitochondria are present. This results in stripes of healthy green cells and yellow clonal homoplasmic mutant cells (Newton *et al.*, 1990; Lauer *et al.*, 1990; Marienfeld and Newton, 1994).

Tobacco CMSI and CMSII and maize NCS mutations differ from other CMS lesions primarily in that the effect of the mutation causes abnormal development in both vegetative and reproductive organs, rather than being restricted to male reproductive organs alone. These mitochondrial mutations are deletions of genes responsible for proper functioning of the electron transport chain and efficient oxidative phosphorylation. This differs from most types of CMS in which novel open reading frames are responsible for the male sterile phenotype. Although it is common for CMS-associated ORFs to contain sequences homologous to genes of the electron transport chain, complete functional copies of these essential genes exist elsewhere in the mitochondrial genome. Since NCS, CMSI and CMSII mutations affect processes essential to every cell in the plant, the effect is seen throughout the organism. However, the phenotypes that result from expression of CMS-associated ORFs are restricted to male reproductive tissues. It is unclear

why the latter CMS mitochondrial mutations do not affect female reproduction or other aspects of plant development.

IV. Restorers of Male Fertility

Genes termed Restorers of fertility (Rf) reverse the sterility phenotype of CMS plants and allow fertile pollen production to proceed. These genes are not, however, found in the mitochondrion; they are located in the nucleus. The nature of these nuclearly-encoded restorers and the mechanism by which they act on genes coded in the mitochondrial genome has been the subject of much recent research.

The CMS/Rf systems may have evolved as a means to promote outcrossing in plants. Many plants have developed a hermaphroditic reproductive system, and coexist in gynodioecious populations alongside female plants. The hermaphroditic state is advantageous to stationary organisms which cannot move around to find mates, and allows plants to self-pollinate to reproduce. However, when a plant is male sterile, as can occur with the loss or inactivation of a restorer gene, the plant must use pollen from another plant to reproduce. This strengthens the population by creating conditions where heterosis or hybrid vigor takes place (Birchler *et al.*, 2003). In this condition, the progeny of a cross is much stronger than the progeny of a self-pollination, and offers a distinct advantage to gynodioecious populations. In agricultural applications, strengthening a population of plants in this manner can be extremely useful, removing the necessity to hand emasculate plants to force out-crossing.

In *Brassica napus*, most cultivars contain the dominant restorer gene for *nap* CMS, *Rfn*. The cultivar Bronowski, however, does not contain *Rfn*, and is, therefore, made male sterile by the *nap* cytoplasm (L'Homme *et al.*, 1997; Li *et al.*, 1998). Normal, male fertile lines of the Bronowski cultivar contain the *cam* (for *campestris*) cytoplasm. This cytoplasm does not appear capable of causing male sterility, and Bronowski with the *cam* cytoplasm is male fertile. The Polima or *pol* cytoplasm is also capable of conferring CMS on *Brassica napus*. The nuclear restorer of the *pol* CMS is designated *Rfp*. *Rfp* is not found in most *B. napus* genotypes and hence, unlike the *nap* cytoplasm, the *pol* cytoplasm is capable of conferring CMS on most *B. napus* varieties (Brown, 1999).

Rf2, one of several genes involved in the restoration of maize *cms*-T, was the first restorer of fertility cloned . This restorer of fertility, *Rf2*, is one of several restoring loci found for the T-cytoplasm CMS (Cui et al., 1996). It encodes an aldehyde dehydrogenase, and is thought to act in one of two ways. In the metabolic hypothesis, *Rf2* is thought to scavenge acetaldehyde and, thereby, eliminate any toxic effect that acetaldehyde may have when present with the CMS- associated URF13 protein. Alternatively, the interaction hypothesis has *Rf2* interacting directly or indirectly with *URF13*, thus diminishing *URF13*'s negative effect on fertility (Cui *et al.*, 1996; Liu *et al.*, 2001). Another locus *Rf1* (Schnable and Wise, 1994) exhibits a direct effect on *URF13* accumulation and is associated with the presence of a 1.6 kb mitochondrial transcript and a 70-80% reduction in URF13 protein. The effect of *Rf1* is somewhat redundant since a similar restoration of fertility can be observed from the combined effect of two loci *Rf8* and *Rf**, which restore fertility in conjunction with *Rf2* (Dill *et al.*, 1997) and whose post-transcriptional regulation of *URF13* involves transcript processing and RNA editing of *urf13* (Dill *et al.*, 1997).

At least two other CMS systems exist in maize, such as CMS-C which is restored by *Rf4* (Sisco, 1991) and CMS-S, which can restored by several nuclear genes, including the recessive gene *rfl1* (Wen et al., 2003). *rfl1*'s action leads to decreased CMS-associated transcript abundance, and it's absence, or presence in the heterozygous condition results in a decrease in *ATPA* protein, an important component of the respiratory chain.

More recently, restorer genes that directly and specifically supress the expression of mitochondrial CMS-associated ORFs have been cloned and characterized. All of these have been found to encode proteins containing pentatricopeptide repeats (PPRs, see below). The first of the PPR restorer genes, *Rf*, was cloned from Petunia (Bentolila *et al.*, 2002). Subsequent to this discovery, *Rf1* from rice (Kazama and Toriyama, 2003; Akagi *et al.*, 2004; Komori *et al.*, 2004), *Rfo* from Ogura radish (Brown *et al.*, 2003, Desloire *et al.*, 2003) and *Rfk* from Kosena radish (Koizuka *et al.*, 2003) were also found to encode PPR genes. In rice, the restorer gene acts to process CMS associated transcripts, thus reducing the abundance of complete, deleterious *atp6* transcripts (Kazama and Toriyama, 2003). The petunia restorer locus contains two separate PPR genes, only one of which restores fertility. The mode of action of the Petunia Rf gene is not known with certainty; the restorer gene reduces the level of the *pcf*-encoded protein (*PCF*) and affects the sizes of *pcf*-encoding transcripts (Bentolila *et al.*, 2002). In radish, *Rfo* does not affect transcript levels or size, but instead seems to affect the steady-state levels of the CMS-associated gene product *ORF138* either at the translational or post-translational level (Brown *et al.*, 2003).

Pentatricopeptide repeats were first defined by Small and Peeters (2000) who distinguished them from tetratricopeptide repeats (TPRs) based on several unique characteristics. The tetratricopeptide repeat is 34 amino acids in length, whereas the pentatricopeptide repeat is 35 amino acids in length, and there are, on average, twice as many repeats in PPR containing genes. Both TPRs and PPRs form alpha helices; TPRs are thought to bind proteins based on the size and biochemical makeup of the protein's central spiral groove. However, PPR genes differ in that their amino acid sidechains which project into the central groove are almost exclusively hydrophilic, resulting in a positively charged region at the bottom of the groove. This groove has sufficient width to hold a single RNA chain, and it's positive charge would bind well to RNA's phosphate backbone (Small and Peeters, 2000).

A recent study has described the PPR gene content of the *Arabidopsis* genome (Lurin *et al.*, 2004). This work has extended the observations of the few PPR genes thus described in *Arabidopsis* and has resulted in fuller documentation of the characteristics of PPR genes. The PPR genes in *Arabidopsis* fall into four major subfamilies, and, in particular, there is one cluster of closely related PPR genes on chromosome 1 which form the densest group of PPRs in the genome. This group consists of 19 genes and several pseudogenes, and shows sequence similarity to the restorers of fertility of petunia and ogura radish. This is also the same chromosome arm on which the restorers of fertility of *nap* and *pol* CMS are localized in the related genome of *Brassica napus*, based on synteny with *Arabidopsis*. The PPR genes found in this region of the *Arabidopsis* genome were shown by sequence analysis to be mitochondrially targeted (Lurin *et al.*, 2004).

The family of PPR genes in plants is greatly expanded compared to other eukaryotes. Since the complexity of plant genomes are not substantially greater than that of animals, the expansion of PPR genes in plants suggests that other proteins in other eukaryotes may fill the roles of plant PPR proteins (Lurin *et al.*, 2004). Disease resistance genes in plants are another example of a large family of related genes, each coding for a specific function. These disease resistance proteins are nucleotide binding and contain both terminal inverted repeats and leucine rich repeats. Their repeat structure differs in small ways from one disease resistance protein to another and each one is highly specific in its function (Myers *et al.*, 1999; Richter and Ronald, 2000). PPR proteins may also be highly sequence specific when binding RNA. Simply by altering the biochemical makeup of the superhelical groove, small changes in the PPR gene sequence could alter the RNA binding capacity of the resultant protein. This would account for the fact that PPR proteins cannot substitute for each other even if their target RNAs are highly homologous. For example, in *Brassica* the *nap* and *pol*, restorers *Rfn* and *Rfp* are closely linked and may be allelic, and their target RNAs (*orf222* and *orf224*) are 85% homologous, yet *Rfn* cannot restore *pol* CMS and *Rfp* cannot restore *nap* CMS (Li *et al.*, 1998).

It is also worth noting that specific cytosine to uracil RNA editing is commonly found in plant mitochondrial and plastid transcripts (Simpson, 1999). Since PPRs are predicted to bind RNA and to be involved in transcript processing, PPRs may have a significant role to play in this area of post-transcriptional processing. It has also been proposed that PPRs may be able to recruit catalytic domains which could catalyze downstream transcript processing events (Lurin *et al.*, 2004).

V. *Brassica napus*

The cruciferae, or mustard, family of plants comprises over 3000 known species. This family includes *Brassica rapa* (formerly *Brassica campestris*) and *Brassica napus* species commonly known as rapeseed (Rollins, 1993). It is thought that *Brassica napus* arose about 10000 years ago as a result of a cross made between *Brassica oleracea* and *Brassica rapa* (Figure 1.3; Warwick *et al.*, 2000).

Original cultivars of rapeseed contained erucic acid in their oil which was known to have harmful effects on lab animals. There was also a significant amount of glucosinolates in the meal, which impart a strong odor and sharp taste, which is distasteful to livestock. The breakdown products of glucosinolates also reduce the nutritional effectiveness of rapeseed meal as a livestock feed. In 1967, the *Brassica napus* variety Bronowski was found to be low in glucosinolates. The Argentine *Brassica napus* variety Tower was developed to be "double low", meaning low in glucosinolates and low in erucic acid. Varieties such as Tower and Westar, which contain less that two percent erucic acid in the oil and less than thirty micromoles of glucosinolates per gram of air-dried oil-free meal, were called Canola (**Can**adian **o**il **l**ow **a**cid) in 1979, a designation which is indicative of these high quality "double low" varieties (www.canola-council.org/rapeseed.aspx).

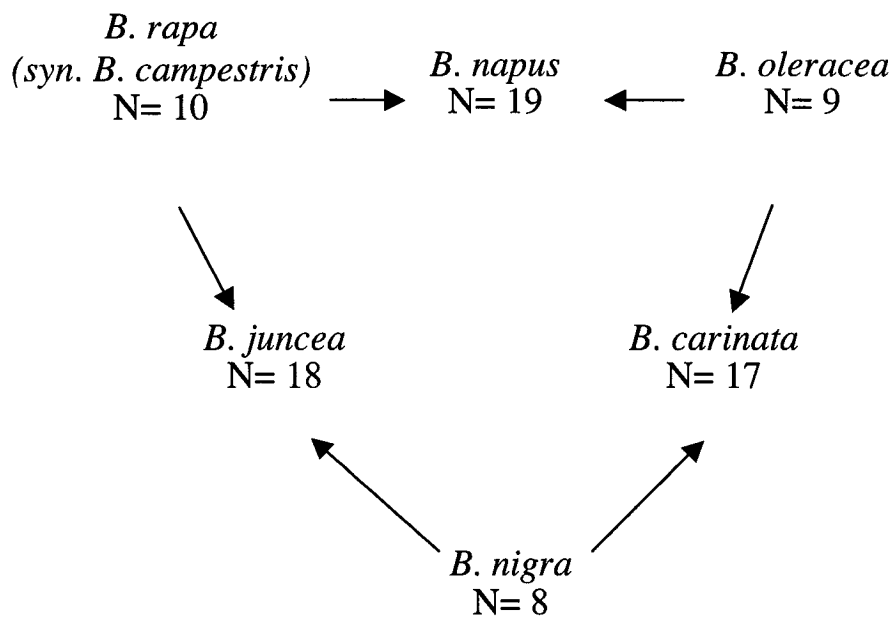
Today, canola is an important part of the Canadian economy. In 2005, 13.05 million acres of canola were grown in Canada, approximately half of which is grown in Saskachewan alone (www.canola-council.org/acreageyields.html). In 2004, Canola was the largest crop by value in Canada (www.canola-coucil.org/

facts_export.html). Thus, studying canola is of great importance to the Canadian economy and worldwide nutrition. Please note the position of the page break.

Brassica napus, an amphidiploid allotetraploid, underwent chromosome doubling and has a haploid number of chromosomes equal to 19. This correlates to the addition of the haploid number of chromosomes from its parental species *B. oleracea* (n=9) and *B. rapa* (n=10) (U, 1935; Warwick *et al.*, 2000; Rana *et al.*, 2004). The genome sizes of these species are large; *B. rapa* and *B. oleracea* are approximately 500 and 600 megabasepairs, respectively, while *B. napus* is estimated to be 1200 megabasepairs (Arumuganathan and Earle, 1991). These crucifers are also related to the model organism *Arabidopsis thaliana*; *Arabidopsis* has only 5 haploid chromosomes and a compact genome of 145 megabasepairs (Arabidopsis Genome Initiative, 2000). The Brassicaceae and *Arabidopsis* diverged about 14-24 million years ago (Yang *et al.*, 1999; Koch *et al.*, 2000) while *B. rapa* and *B. oleracea* diverged from each other about 4 million years ago (Rana *et al.*, 2000).

Figure 1.3 Relationship between different *Brassica* species and their haploid chromosome number (N).

Arrows indicate plant species' genetic contribution to the creation of new species through interspecific crosses. For example, *Brassica napus* (N=19) is a species derived from the cross between *Brassica rapa* (N=10) and *Brassica oleracea* (N=9). (Adapted from U, 1935.)


B. rapa
(syn. B. campestris)
N= 10
B. napus
N= 19
B. oleracea
N= 9
B. juncea
N= 18
B. carinata
N= 17
B. nigra
N= 8

VI. Experimental goals, appproach, and rationale

Cytoplasmic male sterility has been widely studied at the molecular level in many species. We sought to expand the knowledge base for *Brassica napus* CMS through a full characterization of the physical manifestation of this phenomenon. Flower buds were sectioned on a microtome and examined by light microscopy at various stages of development, and a comparison was drawn between developing fertile and sterile flower buds. To determine if any correlation could be found between CMS-associated transcript accumulation and the physical defect of CMS flowers *in situ* hybridization on CMS and fertility-restored floral tissue was performed using probes corresponding to several mitochondrial genes including *orf*222 .

To further elucidate the picture of CMS in *Brassica napus*, map-based cloning of the restorer of fertility, *Rfp*, was undertaken. This was a continuation of work previously begun in the lab. Screening of a cosmid library of *Brassica rapa* introgressed with *Rfp* was intended to simplify the process by exploiting it's smaller genome size (Figure 1.3). Some sequenced *Brassica* DNA from the region of *Rfp* allowed us to further analyse the nature of PPR genes by comparison to sequences from *Arabidopsis thaliana* and other species. A bioinformatic approach including BLAST sequence alignments and phylogenetic tree creation was used to show evidence of changes of location and relatedness of PPR genes as well as diversifying selective pressure acting on them.

References

***Arabidopsis* Genome Initiative (AGI)** (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408**, 796-815.

Akagi, H., Nakamura, A., Yokozekik-Misono, Y., Inagaki, A., Takahashi, H., Mori, K. and Fujimura, T. (2004) Positional cloning of the rice *Rf-1* gene, a restorer of BT-type cytoplasmic male sterility that encodes a mitochondria-targeting PPR protein. Theor. Appl. Genet. **108**, 1449–1457.

Arumuganathan, K. and Earle, E.D. (1991) Nuclear DNA Content of Some Important Plant Species. Plant Mol. Biol. Rep. **9**, 211-215.

Bellaoui, M., Pelletier, G. and Budar, F. (1997) The steady-state level of mRNA from the Ogura cytoplasmic male sterility locus in *Brassica* cybrids is determined post transcriptionally by its 3' region. EMBO J. **16**, 5057-5068.

Bentolila, S., Alfonso, A. and Hanson, M.R. (2002) A pentatricopeptide repeat-containing gene restores fertility to cytoplasmic male-sterile plants. PNAS **99**, 10887-10892.

Birchler, J.A., Auger, D.L. and Riddle, N.C. (2003) In search of the molecular basis of heterosis. Plant Cell **15**, 2236-2239.

Bomblies K., Dagenais N. and Weigel D. (1999) Redundant enhancers mediate transcriptional repression of AGAMOUS by APETALA2. Dev. Biol. **216**, 260-264.

Briggs, D. and Walters, S. M. (1997) *Plant Variation and Evolution*, 3rd ed. New York: Cambridge University Press.

Brown, G.G. (1999) Unique aspects of cytoplasmic male sterility and fertility restoration in *Brassica napus*. J. Heredity **90**, 351-356.

Brown, G.G., Formanova, N., Jin, H., Wargachuk, R., Dendy, C., Patil, P., Laforest, M., Zhang, J., Cheung, W.Y. and Landry, B.S. (2003) The radish *Rfo* restorer gene of Ogura cytoplasmic male sterility encodes a protein with multiple pentatricopeptide repeats. Plant J. **35**, 262-272.

Budar, F and Pelletier, G. (2001) Male sterility in plants: occurrence, determinism, significance and use. Canadian Royal Academy of Sciences **324**, 543-550.

Chetrit P., Rios R., De Paepe R., Vitart V., Gutierres S. and Vedel F. (1992) Cytoplasmic male sterility is associated with large deletions in the mitochondrial DNA of two *Nicotiana sylvestris* protoclones. Curr. Gen. **21**, 131-137.

Cui, X., Wise, R.P. and Schnable, P.S. (1996) The *rf2* nuclear restorer gene of male-sterile T-cytoplasm maize. Science **272**, 1334-1336.

Desloire, S., Gherbi, H., Laloui, W., Marhadour, S., Clouet, V., Cattolico, L., Falentin, C., Giancola, S., Renard, M., Budar, F., Small, I., Caboche , M., Delourme, R. and Bendahmane, A. (2003) Identification of the fertility restoration locus, *Rfo*, in radish, as a member of the pentatrico-peptiderepeat protein family. EMBO Rep. **4**, 588-594.

Dill, C.L., Wise, R.P. and Schnable, P.S. (1997) *Rf8* and *Rf** mediate unique T-*urf13*-transcript accumulation, revealing a conserved motif associated with RNA processing and restoration of pollen fertility in T-cytoplasm maize. Genetics **147**, 1367-1379.

Goldberg, R.B., Beals, T.P., and Sanders, P.M. (1993) Anther development: basic principles and practical applications. Plant Cell **5**, 1217-1229.

Gomez-Mena C., de Folter S., Costa M.M., Angenent G.C. and Sablowski R. (2005) Transcriptional program controlled by the floral homeotic gene AGAMOUS during early organogenesis. Development **132**, 429-438.

Gutierres, S., Sabar, M., Lelandais, C., Chetrit, Pl., Diolez, P., Degand, H., Boutry, M., Vedel, F., DeKouchkovsky, Y. and DePaepe, R. (1997) Lack of mitochondrial and nuclear-encoded subunits of complex I and alteration of the respiratory chain in *Nicotiana sylvestris* mitochondrial deletion mutants. PNAS **94**, 3436-3441.

Hanson, M. and Bentolila, S. (2004) Interactions of mitochondrial and nuclear genes that affect male gametophyte development. Plant Cell **16 (Suppl.)**, S154-S169.

Hanson, M.R., Wilson, R.K., Bentolila, S., Kohler, R.H. and Chen, H.C. (1999) Mitochondrial gene organization and expression in petunia male fertile and sterile plants. J. Heredity **90**, 362-368.

He, C., Tirlapur, U., Cresti, M., Peja, M., Crone, D.E. and Mascarenhas, J.P. (1996) An *Arabidopsis* mutant showing aberrations in male meiosis. Sex. Plant Rep. **9**, 54–57.

Hernould, M., Suharsono, S., Zabaleta, E., Carde, J.P., Litvak, S., Araya, A. and Mouras, A. (1998) Impairment of tapetum and mitochondria in engineered male-sterile tobacco plants. Plant Mol. Biol. **36**, 499-508.

Hill, T.A., Day, C.D., Zondlo, S.C., Thackeray, A.G. and Irish, V.F. (1998) Discrete spatial and temporal cis-acting elements regulate transcription of the *Arabidopsis* floral homeotic gene APETALA3. Development **125**, 1711-1721.

Honma, T. and Goto, K. (2001) Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. Nature **409**, 525-529.

Ito, T., Wellmer, F., Yu, H., Das, P., Ito, N. Alves-Ferreira, M., Riechmann, J.L. and Meyerowitz, E.M. (2004) The homeotic protein AGAMOUS controls microsporogenesis by regulation of SPOROCYTELESS. Nature **15**, 356-360.

Kadowaki, K., Suzuki, T. and Kazama, S. (1990) A chimeric gene containing the 5' portion of *atp6* is associated with cytoplasmic male-sterility of rice. Mol. Gen. Genet. **224**, 10-16.

Kamalay, J.C. and Goldberg, R.B. (1980) Regulation of structural gene expression in tobacco. Cell **19**, 935-946.

Kamalay, J.C. and Goldberg, R.B. (1984) Organ-specific nuclear RNAs in tobacco. PNAS **81**, 2801-2805.

Karpova, O.V., Kuzmin, E.V., Elthon, T.E. and Newton K.J. (2002) Differential expression of alternative oxidase genes in maize mitochondrial mutants. Plant Cell **14**, 3271-3284.

Kazama, T. and Toriyama, K. (2003) A pentatricopeptide repeat-containing gene that promotes the processing of aberrant *atp6* RNA of cytoplasmic male-sterile rice. FEBS Letters **544**, 99-102.

Koch, M., Haubold, B. and Mitchell-Olds, T. (2000) Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase in *Arabidopsis*, *Arabis*, and related genera. Mol. Biol. Evol. **17**, 1482–1498.

Kofer, W., Glimelius, K. and Bonnett, H.T. (1991) Modifications of mitochondrial DNA cause changes in floral development in homeotic-like mutants of tobacco. Plant Cell, **3**, 759-769.

Koizuka, N., Imai, R., Fujimoto, H., Hayakawa, T., Kimura, Y., Kohno-Murase, J., Sakai, T., Kawasaki, S. and Imamura, J. (2003) Genetic characterization of a pentatricopeptide repeat protein gene, *orf687*, that restores fertility in the cytoplasmic male-sterile Kosena radish. Plant J. **34**, 407-415.

Komaki, M.K., Okada, K., Nishino, E. and Shimura, Y. (1988) Isolation and characterization of novel mutants of *Arabidopsis thaliana* defective in flower development. Development **104**, 195-203.

Komori, T., Ohta, S., Murai, N., Takakura, Y., Kuraya, Y., Suzuki, S., Hiei, Y., Imaseki, H. and Nitta, N. (2004) Map-based cloning of a fertility restorer gene, *Rf-1* in rice (*Oryza sattiva L.*). Plant J. **37**, 315-325.

Lamb, R.S., Hill, T.A., Tan, Q.K.-G. and Irirsh, V.F. (2002) Regulation of APETALA3 floral homeotic gene expression by meristem identity genes. Development **129**, 2079-2086.

Lauer, M., Knudsen, C., Newton, K.J., Gabay-Laughnan, S. and Laughnan, J.R. (1990) A partially deleted mitochondrial cytochrome oxidase gene in the NCS6 abnormal growth mutant of maize. New Biology **2**, 129-186.

Lenhard, M., Bohnert, A., Jurgens, G. and Laux, T. (2001) Termination of stem cell maintenance in *Arabidopsis* floral meristems by interactions between WUSCHEL and AGAMOUS. Cell **105**, 805-814.

Levings, C.S. 3rd. (1993) Thoughts on Cytoplasmic Male Sterility in *cms*-T Maize. Plant Cell **5**, 1285-1290.

Lewis, D. (1941) Male sterility in natural populations of hermaphrodite plants: the equilibrium between females and hermaphrodites to be expected with different types of inheritance. New Phytologist **40**, 56–63.

L'Homme, Y., Stahl, R.J., Li, X.-Q., Hameed, A. and Brown, G.G. (1997). *Brassica nap* cytoplasmic male sterility is associated with expression of a mtDNA region containing a chimeric gene similar to the *pol* CMS-associated *orf224* gene. Curr. Genet. **31**, 325-335.

Li, X.-Q., Jean, M., Landry, B.S., and Brown, G.G. (1998). Restorer genes for different forms of *Brassica* cytoplasmic male sterility map to a single nuclear locus that modifies transcripts of several mitochondrial genes. PNAS **95**, 10032-10037.

Liljegren, S.J., Gustafson-Brown, C., Pinyopich, A., Ditta, G.S. and Yanofsky, M.F. (1999) Interactions among APETALA1, LEAFY, and TERMINAL FLOWER1 specify meristem fate. Plant Cell **11**, 1007-1018.

Linke, B., Nothnagel, T. and Borner, T. (1999) Morphological characterization of modified flower morphology of three novel alloplasmic male sterile carrot sources. Plant Breeding **118**, 543–548.

Liu, F., Cui, X., Horner, H., Weiner, H. and Schnable, P.S. (2001) Mitochondrial aldehyde dehydrogenase activity is required for male fertility in maize. Plant Cell **13**, 1063-1073.

Lohmann, J.U., Hong, R.L., Hobe, M., Busch, M.A., Parcy, F., Simon, R. and Weigel, D. (2001) A molecular link between stem cell regulation and floral patterning in *Arabidopsis*. Cell **105**, 793-803.

Lohmann, J.U. and Weigel, D. Building Beauty: The Genetic Control of Floral Patterning. Dev. Cell **2**, 135-142.

Lurin, C., Andre, C., Aubourg, S., Bellaoui, M., Betton, F., Bruyere, C., Caboche, M., Debast, C., Gualberto, J., Hoffmann, B., Lecharny, A., LeRet, M., Martin-Magniette, M.-L., Mireau, H., Peeters., N., Renou, J.-P., Szurek, B., Taconnat, L. and Small, I. (2004) Genome-wide analysis of *Arabidopsis* pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. Plant Cell **16**, 2089-2103.

Marienfeld, J.R. and Newton, K.J. (1994) The maize NCS2 abnormal growth mutant has a chimeric *nad4-nad7* mitochondrial gene and is associated with reduced complex I function. Genetics **138**, 855-863.

Menassa, R., L'Homme, Y. and Brown, G.G. (1999) Post-transcriptional and developmental regulation of a CMS-associated mitochondrial gene region by a nuclear restorer gene. Plant J. **17**, 491-499.

Myers, B., Dickerman, A., Michelmore, R., Sibaramakrishnan, S., Sobral, B. and Young, N. (1999) Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. Plant J. **20**, 317-332.

Newton, K.J., Knudsen, C., Gabay-Laughnan, S. and Laughnan, J.R. (1990) An abnormal growth mutant in maize has a defective mitochondrial cytochrome oxidase gene. Plant Cell **2**, 107-113.

Ng, M. and Yanovsky, M.F. (2001) Function and evolution of the plant MADS-box gene family. Nat. Rev. Genet. **195**, 186-195.

Ogihara, Y., Futami, K., Tsuji, K. and Murai, K. (1997) Alloplasmic wheats with *Aegilops crassa* cytoplasm which express photoperiod-sensitive homeotic transformations of anthers, show alterations in mitochondrial DNA structure and transcription. Mol. Gen. Genet. **255**, 45–53.

Parcy, F., Nilsson, O., Busch, M.A., Lee, I. and Weigel, D. (1998) A genetic framework for floral patterning. Nature **395**, 561-566.

Pla, M., Mathieu, C., De Paepe, R., Chetrit, P. and Vedel, F. (1995) Deletion of the last two exons of the mitochondrial *nad7* gene results in lack of the NAD7 polypeptide in a *Nicotiana sylvestris* CMS mutant. Mol. Gen. Genet. **248**, 79-88.

Polowick, P.L. and Sawhney, V.K. (1986) A scanning electron microscopic study on the initiation and development of floral organs of *Brassica napus* (Cv. Westar). Am. J. Botany **73**, 254-263.

Pylatuik, J.D., Lindsay, D.L., Davis, A.R. and Bonham-Smith, P.C. (2003) Isolation and characterization of a *Brassica napus* cDNA corresponding to a B-class floral development gene. J. of Exp. Bot. **54**, 2385-2387.

Rana, D., van den Boogaart, T., O'Neill, C.M., Hynes, L., Bent, E., Macpherson, L., Park, J.Y., Lim, Y.P. and Bancroft, I. (2004) Conservation of the microstructure of genome segments in *Brassica napus* and its diploid relatives. Plant J. **40**, 725-733.

Riechmann, J. L., Wang, M. and Meyerowitz, E. M (1996). DNA-binding properties of *Arabidopsis* MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA and AGAMOUS. Nucl. Acids Res **24**, 3134-3141.

Richter, T. and Ronald, P. (2000) The evolution of disease resistance genes. Plant Mol. Biol. **42**, 195-204.

Rollins, R.C. (1993) The Cruciferae of continental North America: systematics of the mustard family from the Arctic to Panama. Stanford University Press, Stanford.

Ross, K.J., Fransz, P., Armstrong, S.J., Vizir, I., Mulligan, B., Franklin, F.C. and Jones, G.H. (1997) Cytological characterization of four meiotic mutants of *Arabidopsis* isolated from T-DNA-transformed lines. Chromosome Research **5**, 551-559.

Sabar, M., De Paepe, R. and de Kouchkovsky, Y. (2000) Complex I impairment, respiratory compensations, and photosynthetic decrease in nuclear and mitochondrial male sterile mutants of *Nicotiana sylvestris*. Plant Phys. **124**, 1239-1250.

Samach, A., Klenz, J.E., Kohalmi, S.E., Risseeuw, E., Haughn, G.W. and Crosby, W.L. (1999) The UNUSUAL FLORAL ORGANS gene of *Arabidopsis thaliana* is an F-box protein required for normal patterning and growth in the floral meristem. Plant J. **20**, 433-445.

Sanders, P.M., Bui, A.Q., Weterings, K., McIntire, K.N., Hsu, Y.-C., Lee, P.Y., Truong, M.T., Beals, T.P., and Goldberg, R.B. (1999) Anther developmental defects in *Arabidopsis thaliana* male-sterile mutants. Sex. Plant Rep. **11**, 297–322.

Schnable, P. and Wise, R. (1994) Recovery of heritable, transposon-induced mutant alleles of the *rf2* nuclear restorer of T-cytoplasm maize. Genetics **136**, 1171-1185.

Schultz, E.A. and Haughn, G.W. (1991) *LEAFY*, a Homeotic Gene That Regulates Inflorescence Development in *Arabidopsis*. Plant Cell **3**, 771-781.

Sieburth, L.E., Drews, G.N. and Meyerowitz, E.M. (1998) Non-autonomy of *AGAMOUS* function in flower development: use of a Cre/loxP method for mosaic analysis in *Arabidopsis*. Development **125**, 4303-4312.

Simpson L. RNA editing-An Evolutionary Perspective. In: Atkins J , Gesteland R, editors. *The RNA World. Cold Spring Harbor,* New York: Cold Spring Harbor Laboratory Press, 1999 pp. 585-608.

Sisco, P.H. (1991) Duplications complicate genetic mapping of *Rf4,* a restorer gene for *cms*-C cytoplasmic male sterility in corn. Crop Science **31,** 1263-1266.

Small, I.D. and Peeters, N. (2000) The PPR motif is a TPR-related motif prevalent in plant organellar proteins. Trends Biochem. Sci. **25,** 46-47.

Smart, C.J., Moneger, F., and Leaver, C.J. (1994) Cell-specific regulation of gene expression in mitochondria during anther development in Sunflower. Plant Cell **6,** 811-825.

Smyth, D.R., Bowman, J.L., and Meyerowitz, E.M. (1990) Early flower development in *Arabidopsis*. Plant Cell **2,** 755-767.

Tang, H.V., Pring, D.R., Shaw, L.C., Salazar, R.A., Muza, F.R. Yan, B. and Schertz, K.F. (1996) Transcript processing internal to a mitochondrial open reading frame is correlated with fertility restoration in male-sterile sorghum. Plant J. **10,** 123–133.

Theissen, G., Becker, A., Di Rosa, A., Kanno, A., Kim, J.T., Munster, T., Winter, K.U. and Saedler, H. (2000) A short history of MADS-box genes in plants. Plant Mol. Biol. **42,** 115-149.

Tilly, J.J., Allen, D.W. and Jack, T. (1998) The CArG boxes in the promoter of the *Arabidopsis* floral organ identity gene *APETALA3* mediate diverse regulatory effects. Development **125,** 1647-1657.

U. N. (1935) Genome analysis in Brassica with special reference to the experimental formation of Brassica napus and peculiar mode of fertilization. Japanese Journal of Botany 7: 389-452.

Warwick, S.I., Francis, A. and LaFleche, J. (2000) Guide to the Wild Germplasm of *Brassica* and Allied Crops, 2nd Edition. Agriculture and Agri-Food Canada, Ottawa.

Weigel, D. and Meyerowitz, E.M. (1994) The ABCs of floral homeotic genes. Cell **78**, 203-209.

Wen, L.-Y., Ruesch, K., Ortega, V., Kamps, T., Gabay-Laughnan, S. and Chase, C. (2003) A nuclear *restorer-of-fertility*, mutation disrupts accumulation of mitochondrial atp synthase subunit α in developing pollen of *S* male-sterile maize. Genetics **165**, 771-779.

Wen, L.Y. and Chase, C. (1999) Mitochondrial gene expression in developing male gametophytes of male-fertile and S male-sterile maize. Sex. Plant Rep. **11**, 323–330.

Wise, R.P., Gobelman-Werner, K., Pei, D., Dill, C.L. and Schnable P.S. (1999) Mitochondrial transcript processing and restoration of male fertility in T-cytoplasm maize. J. Heredity **90**, 380–385.

Yang, Y.W., Lai, K.N., Tai, P.Y. and Li, W.H. (1999) Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between *Brassica* and other angiosperm lineages. J. Mol. Evol. **48**, 597–604.

Zhang, P., Tan, H.T.W., Pwee, K-H. and Kumar, P.P. (2004) conservation of class C function of floral organ development during 300 million years of evolution from gymnosperms to angiosperms. Plant J. **37**, 566-577.

Zhao, D., Yu, Q., Chen, M. and Ma, H. (2001) The *Ask1* gene regulates B function gene expression in cooperation with UFO and LEAFY in *Arabidopsis*. Development **128**, 2735-2746.

Zubko, M., Zubko, E., Patskovsky, Y., Khvedynich, O., Fisahn, J., Gleba, Y. and Schieder, O. (1996) Novel 'homeotic' CMS patterns generated in *Nicotiana* via cybridization with *Hyoscyamus* and *Scopolia*. J. Exp. Bot. **47**, 1101-1110.

CHAPTER 2

CELL-SPECIFIC REGULATION OF A *BRASSICA NAPUS* CMS-ASSOCIATED GENE BY A NUCLEAR RESTORER WITH RELATED EFFECTS ON A FLORAL HOMEOTIC GENE PROMOTER

This work has been published in the Plant Journal according to the citation stated below. The manuscript presented is co-authored by Laetitia Mahé, whose work is presented in the results section entitled "Altered regulation of the APETALA3 promoter in CMS versus fertile *B. napus*", and by Dr. Gregory G. Brown in his capacity as research supervisor.

Geddy, R., Mahé, L. and Brown, G.G. (2004) Cell-specific regulation of a *Brassica napus* CMS-associated gene by a nuclear restorer with related effects on a floral homeotic gene promoter. Plant J. **41**, 333-345.

Introduction

Cytoplasmic male sterility (CMS) is a maternally inherited trait found in many different plant species. In those cases where CMS has been examined in detail at the molecular level, it has been found to correlate with the expression of novel, often chimeric mitochondrial genes (Hanson and Bentolila, 2004). Nuclear restorer of fertility, or Rf, genes can suppress the male sterile phenotype, and are also widespread in occurrence. Rf genes usually affect only a particular type of CMS and act, in most cases, by specifically downregulating the expression of the corresponding novel, CMS-inducing mitochondrial gene. Several Rf genes have recently been characterized and most have been found to encode a protein containing pentatricopeptide (PPR) repeats (Hanson and Bentolila, 2004). PPR repeat-containing proteins represent a large gene family in plants, and other PPR-containing proteins, both in plants and fungi, post-transcriptionally regulate the expression of organelle genes (Hanson and Bentolila, 2004; Meierhoff *et al.*, 2003; Small and Peeters, 2000).

Although significant progress has been made in identifying the genes that specify CMS and fertility restoration, little is understood about how expression of these genes influences the physiological, developmental or biochemical processes such that pollen formation is disrupted. Pollen development can be disrupted at any of several different stages, depending on the plant species and CMS cytoplasm. In maize *cms*-S, for example, anther development is normal and pollen abortion occurs during gametophyte development (Gabay-Laughnan *et al.*, 1995) while in petunia, sunflower, and maize *cms*-T, pollen development proceeds normally

only through the tetrad stage (Hanson and Conde, 1985). In these latter examples, external anther development appears normal, but there is a premature breakdown of the tapetum, the nurse cells that line the pollen sacs. In the alloplasmic CMS of Nicotiana (Kofer *et al.*, 1991; Zubko *et al.*, 2001), as well as in some other species (Linke *et al.*, 2003; Murai *et al.*, 2002; Robison and Wolyn, 2002), diverse aberrations in stamen morphology are observed, ranging from the absence of stamens to stamens that have structural features of carpels or petals. In all cases, CMS is thought to be related to mitochondrial dysfunction resulting from the expression of abnormal mitochondrial genes.

In the above-mentioned CMS examples, vegetative plant growth and most aspects of flower development, including ovule development and female fertility, are unaffected or only slightly perturbed. It is unclear how mitochondrial dysfunction can have such diverse effects on pollen and stamen development without otherwise affecting plant development and function. It is possible that male sterility-related phenotypes result from the interplay of tissue and cell-specific regulatory processes involving stamen development, mitochondrial gene expression, and mitochondrial function *per se*. Surprisingly, however, only a few studies documenting developmental regulation of mitochondrial gene activity in CMS (Conley and Hanson, 1994; Smart *et al.*, 1994) have been conducted.

The oilseed rape species *Brassica napus* contains two endogenous, CMS conferring cytoplasms, *nap* and *pol*. A third cytoplasm, *cam*, is not known to confer male sterility. *pol* CMS is associated with expression of the novel open reading frame (ORF), *orf224*, situated upstream of and co-transcribed with a normal

mitochondrial gene, *atp6* (Brown, 1999; L'Homme and Brown, 1993; Singh and Brown, 1991) while *nap* CMS is associated with a different but related ORF, *orf222*, that is co-transcribed with *nad5*c, the central exon of a trans-spliced gene, and a short ORF of unknown function (Brown, 1999; L'Homme *et al.*, 1997). Unlike most other CMS-associated ORFs, *orf224* and *orf222* are highly similar to one another, with approximately 85% nucleotide and 79% amino acid similarity (L'Homme *et al.*, 1997); for each ORF, approximately 60 amino acids at the extreme N-terminus of the encoded protein are derived from a normal mitochondrial gene, *atp8* (or *orfB*). In other species known to contain more than one form of CMS, restorer genes for the different forms map to different nuclear loci; however, the *nap* and *pol* CMS restorers, *Rfn* and *Rfp*, respectively, represent different alleles or haplotypes of a single nuclear locus (Li *et al.*, 1998).

For both the *nap* and *pol* CMS systems, the nuclear restorer gene acts to reduce levels of novel ORF-containing transcripts, and this reduction in transcript levels is accompanied by the appearance of additional, shorter transcripts that do not contain the complete ORF coding sequence. *Rfp* acts by conditioning the processing of *orf224/atp6* transcripts; these processing events appear to be more extensive in stamens than in other floral organs (Menassa *et al.*, 1999). *Rfn* action is similarly developmentally regulated (H. Elina and G.G. Brown, unpublished observations) and likely acts by mediating RNA-processing events. In this study, we describe in detail the development of anthers and surrounding floral structures for these forms of CMS in comparison with that of normal *B. napus* flowers. We relate the developmental differences occurring in the CMS versus restored flowers with differences in mitochondrial gene expression. Further, we

show that the stamen/petal-specific promoter of the *APETALA3* (AP3) gene is regulated differently in sterile and fertile flowers, suggesting a possible link between the molecular and developmental features of *Brassica* CMS.

Materials and Methods

Plant material

Brassica napus cytoplasms are designated according to the convention of Kemble and Barsby (1988) and are indicated by italicized designations. The different *B. napus* strains used in this study have been described previously (L'Homme *et al.*, 1997). Plants were grown in the McGill Phytotron growth chambers under normal growth conditions as described (Singh and Brown, 1991).

Pollen viability

Pollen viability was assessed as described by Alexander (1969). Pollen was incubated on a coverslipped slide for 24 h in a solution made with 10 mL 95% ethanol, 50 mL distilled water, 25 mL glycerol, 5 g phenol, 1 mL glacial acetic acid, 10 mg Malachite green (added 1 mL of a 1% solution in 95% ethanol), 5 g chloral hydrate, 50 mg acid fuchsin (added 5 mL of a 1% solution in water), and 5 mg orange G (0.5 mL of a 1% solution in water). Aborted pollen stained green and non-aborted pollen stained red.

Pollen germination

Pollen germination was tested as described by Carpenter *et al.* (1992). Pollen was incubated for 8 h on a plate containing 15% w/v sucrose, 0.4 mM $Ca(NO_3)_2$,

0.4 mM H_3BO_4, and 1% w/v agarose. Pollen tube growth was then viewed under a light microscope.

Nucleic acid extraction

Total RNA was extracted from seedlings and floral bud tissue using the RNA isolation reagent RNAwizTM by Ambion (Austin, TX, USA) according to the supplier's instructions. Total cellular DNA was extracted from seedlings using the DNeasy® Plant Maxi Kit (Qiagen Inc., Mississauga, ON, Canada) according to the supplier's instructions. Mitochondrial RNA was extracted using a method similar to Singh and Brown (1991) with the following changes. Plant tissue was homogenized in 5 volumes of high ionic strengh buffer as described by Singh and Brown (1991), with the omission of beta-mercaptoethanol, and with the addition of a second grinding of the homogenate caught in the double layer miracloth filter in an additional 5 volumes of buffer. Mitochondria were lysed at room temperature for 10 min with gentle shaking. Preciptation of RNA was carried out at 4°C overnight and centrifuged at 13 000 *g* for 15 min before being resuspended in water and stored at 70°C. Mitochondrial DNA was extracted essentially as described by Kemble (1987) with the addition of a second grinding of the homogenate caught in the cheesecloth/miracloth filter in an additional 5 volumes of buffer.

Hybridization probes

Primers were designed to amplify *B. napus* mitochondrial genes by the polymerase chain reaction, using purified *B. napus* mtDNA as a template. In the case of the gamma subunit of the ATP synthase, primers were designed according to the sequence of the homologous gene from *Arabidopsis thaliana.* These primers were then used to amplify the corresponding sequence from total cellular *B. napus* DNA. The sequences of the various primers employed for amplification are as follows:

orf222: 5'TAAAGCTTGGTGG-AAAAGATCGTACAAGTA3',
5'TAGGTACCGAAAGACCCGGAAGTG-GTGATC3';
atp6: 5'ATATGCGGCCGCTTGTCCCATTGATTCCTA3',
5'ATCGTCTAGAAGAAAATCTCATTCATACATAGCA3';
ATP synthase gamma subunit: 5'GTTCGTCGAGCTCAAGATGCAG3',
5'AAGATT-CCTGCAGTGCTCTTAG3';
cob: 5'GTTACTGAAGGGTTGTCGA-CTG3',
5'GGGCGAATTAAGCGTCGACG3'.

These primers amplified the following segments of the corresponding genes, with nucleotide 1 being the first nucleotide of the respective coding sequence: *orf222*: nucleotides 173 through 609; *atp6*: nucleotides 50 through 650; ATP synthase gamma subunit: nucleotides 302 through 1006; *cob*: nucleotides 73 through 281.

PCR products were cloned into the vector pCR 2.1 (Invitrogen, Carlsbad, CA, USA). These inserts were then cut at restriction enzymes designed to fall within the primer sequences, and the fragments obtained were gel-purified and recloned into pBluescript SK+ (Stratagene, La Jolla, CA, USA). These clones were sequenced to confirm their identity, and Northern and Southern analyses were performed to confirm the probes hybridized in expected patterns before use in *in situ* hybridization. RNA probes were made via *in vitro* transcription from the T3 or T7 promotors located within the vector using an *in vitro* transcription kit provided by Promega (Madison, WI, USA).

Northern and Southern blot analyses

Northern and Southern blots and hybridizations were performed according to the supplier's instructions using Hybond-N+ membrane by Amersham Biosciences (Piscataway, NJ, USA).

In situ *hybridization*

In situ hybridization was according to Drews *et al.* (http://www.its.caltech.edu/~plantlab/protocols/insitu.html), with minor alterations. Sections were placed onto Superfrost Plus slides from Fisher Scientific (Pittsburg, PA, USA) without any other preparation. Labeled probes were prepared by *in vitro* transcription of a gel-purified polymerase chain reaction-amplified templates, using 35S-labeled UTP. Following hybridization and washing, the slides were immersed in NTB-2 photographic emulsion (Kodak, Rochester, NY, USA),

exposed for 16 days and developed. Slides were then stained with toluidine blue, viewed through a Zeiss Axioplan microscope fitted with a Spot digital camera (Diagnostic Instruments, Sterling Heights, MI, USA) and photographed under both bright and darkfield illumination.

The pAP3::GUS construct and plant transformation

A portion of the *AP3* promoter region (Jack *et al.*, 1992) extending from the start of the coding sequence to a position 458 bp upstream was introduced as an *Xba*I/*Bam*HI fragment into the polylinker of the Brassica GUS expression vector pRD420 (Datla *et al.*, 1992). This construct corresponds roughly to the 496 GUS expression construct of Tilly *et al.* (1998) which is expressed in Arabidopsis with the same temporal and spatial pattern as the full-length 3.7 kb *AP3* upstream sequence.

The constructs were introduced into male fertile and *pol* CMS *B. napus* cv. 'Westar' cotyledons according to Moloney *et al.* (1989). Two different T0 transgenic events from both fertile and the sterile transformants were analyzed for GUS expression. Succeeding T1 and T2 generations from each primary transformant were developed by self-fertilizing T0 plants (the leakiness of the CMS phenotype allowed a small number of selfed progeny to be recovered from sterile T0 and T1 plants). In addition, fertility-restored plants were generated by crossing sterile T1 lines with the fertility restorer line Westar-Rf (Jean *et al.*, 1997). Flowers on two separate inflorescences were examined for each GUS positive individual. In total, 12 fertile and 12 sterile GUS-positive plants from the T1 and

T2 and fertility-restored hybrid generations were studied. GUS expression patterns among all the examined fertile plants were indistinguishable as were GUS expression patterns among all the examined sterile plants.

Detection of in situ *GUS activity*

GUS activity was detected in whole inflorescences by immediately fixing them in ice cold 90% acetone for approximately 10 min following removal from the inflorescence. These were rinsed in 50 mM Na_3PO_4, 0.5 mM of potassium ferricyanide ($K_3Fe(CN)_6$) and 0.5 mM potassium ferrocyanide ($K_4Fe(CN)_6$) and vacuum infiltrated with 1.5 mM of X-GLUC added in the same solution. Vacuum infiltration was performed two times within 15 min with continuous tapping to remove air bubbles around tissues. Tissue was gently agitated at 37°C for 1 h, then at room temperature overnight. Chlorophyll was removed by passage through an ethanol series (30, 50, 70, 85, 95, 100, 100%). Inflorescences immersed in ethanol were photographed using a dissecting microscope fitted with a digital camera.

To characterize GUS activity in tissue sections of inflorescences, after overnight treatment in X-GLUC, were submerged in a medium consisting of 1/3 cryostat embedding medium (Histo PrepTM from Fisher Scientific) and 2/3 20% sucrose. The inflorescences were placed on the surface of the medium and allowed to sink to prevent formation of air bubbles. For buds, physical constriction and vacuum infiltration was necessary to remove air bubbles trapped inside. Tissue was mounted in partially frozen embedding medium, and maintained at 80°C until

sectioning. Sectioning was performed with a cryostat maintained at 20°C. Sections (12 mm) were placed on slides and allowed to thaw dry in air. Sections were photographed using a Zeiss Axioplan microscope fitted with a Spot digital camera (Diagnostic Instruments).

Results

Normal anther development in B. napus

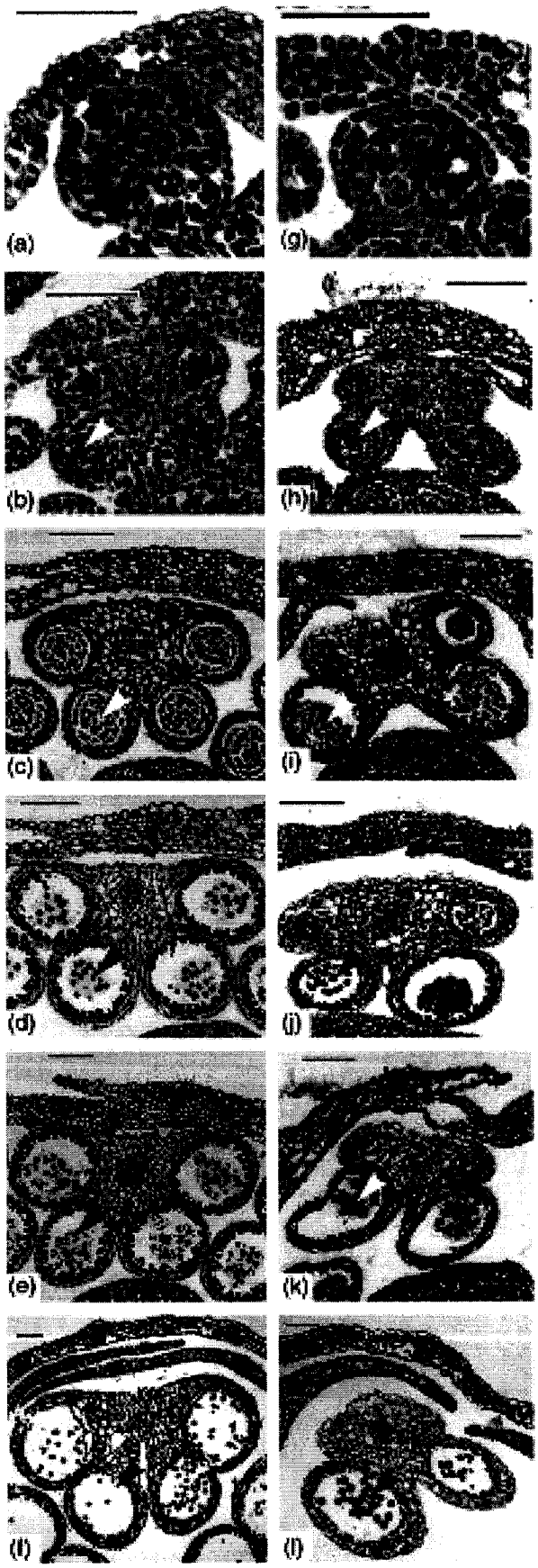
Although the morphology of normal anther and flower development has been characterized in cruciferous species (Sanders *et al.*, 1999; Smyth *et al.*, 1990; Theis and Röbbelen, 1990), no detailed morphological comparison of the development of fertile versus *pol* or *nap* CMS anthers is available. To address this, serial sections of fertile and CMS floral buds, representing different stages of anther development, were compared. Seven stages were used as benchmarks for further examination, as they were deemed representative of significant changes in floral development. The stages of fertile anther development are largely those described by Theis and Röbbelen (1990). For convenience they are described briefly again here.

Stage 1, primordial anther (Figure 2.1a): the emerging anther tissue mass is oval-shaped in cross section; no evidence of cellular differentiation within the anther is evident.

Stage 2, sporogenous cell (Figure 2.1b): characterized by the development of sporogenous cells and the emergence of a characteristic asymmetric anther shape with two adaxial (proximal to the carpel) and two abaxial sporogenous cell masses.

Figure 2.1 Development of fertile (a-f) and CMS (g-l) anthers in *Brassica napus*.

(a, g) Primordial anthers: the developing anther is oblong and shows no obvious signs of cellular differentiation; no differences between fertile and CMS anthers are evident at this stage.
(b, h) Sporogenous cell (SC) stage: the anther has assumed a four-lobed appearance; sporogenous cells (arrows) are evident within the developing locules; asymmetry and asynchrony of locule formation in CMS anthers is evident.
(c, i) Microspore mother cell (MMC) stage: irregularly shaped MMCs and a tapetum are evident in the locules (arrow); non-sporogenic tissues organize into the connective and vasculature; MMCs are clumped in some locules of CMS anthers (arrow).
(d, j) Meiocyte stage: meiocytes (arrows) have developed a callose wall and entered meiosis; the tapetum becomes vacuolated; variation in the appearance of gametophytic tissues among different locules is evident in CMS anthers.
(e, k) Tetrads of microspores enclosed in callose are visible in CMS and fertile anthers (arrows); clumps of dense tissue appear in some locules of CMS anthers.
(f, l) Pre-dehiscent pollen grain stage: anther increases in size; pollen (arrows) develops exine wall; tapetum degeneration. Size bars = 100μm.


(a)
(g)
(b)
(h)
(c)
(i)
(d)
(j)
(e)
(k)
(f)
(l)

Stage 3, microspore mother cell (Figure 2.1c): the sporogenous cell mass has differentiated into irregularly shaped microspore mother cells; the tapetal cell layer of the locule and the vasculature of the anther are visibly differentiated from the surrounding connective tissue.

Stage 4, meiocyte (Figure 2.1d): microspore mother cells have separated from one another and from the tapetum, and have differentiated into rounded meiocytes which become enclosed in callose and enter meiosis. During this period the locules continue to enlarge while the meiocytes within do not. The tapetum becomes vacuolated.

Stage 5, tetrad (Figure 2.1e): anther growth slows as the meiotic divisions generate tetrads of microspores.

Stage 6, microspore (not shown): tetrads dissociate into individual microspores; tapetum degeneration evident.

Stage 7, pollen (Figure 2.1f): anther size increases as the micropores differentiate into mature pollen grains; the anther wall thins near the stomium, and eventually degrades through apoptosis to release the pollen. Petals, the last floral organ to mature, are now evident in anther cross sections.

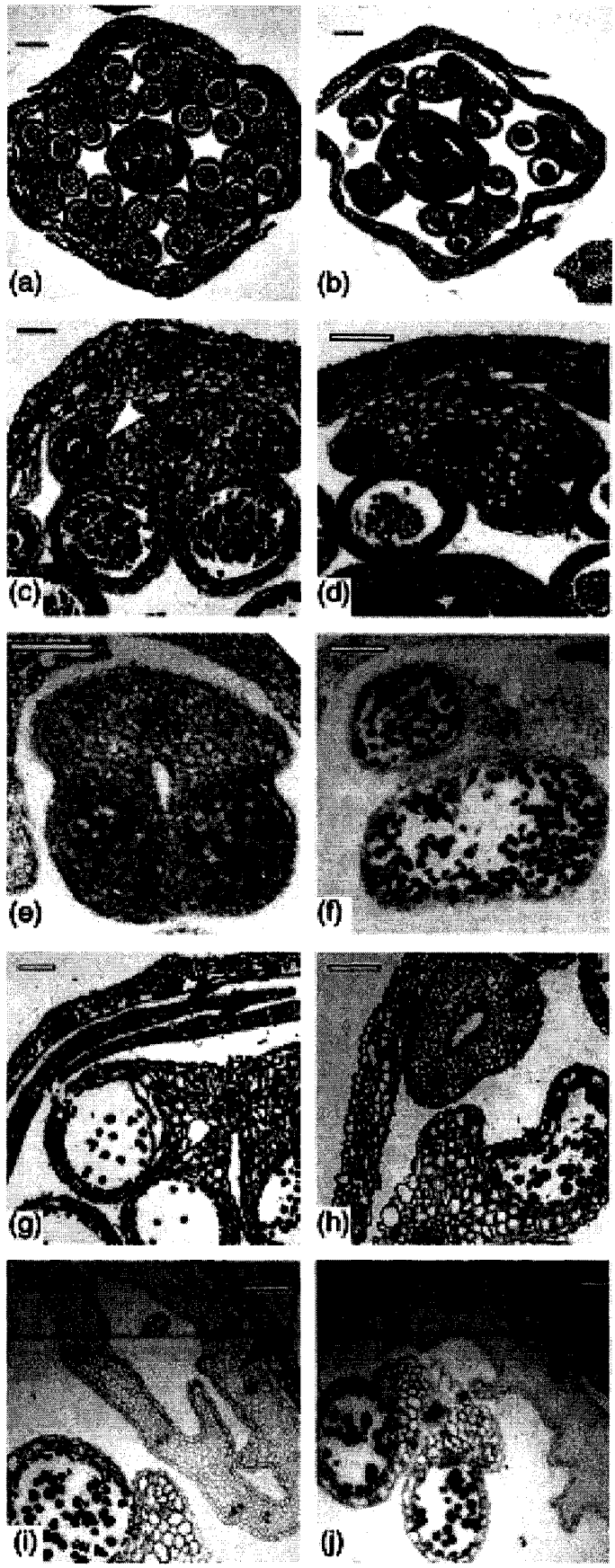
Development of CMS anthers

We observed no morphological difference between fertile and sterile buds at the earliest, primordial anther stage (Figure 2.1a,g). Differences begin to appear during the differentiation of primordial anther tissue into sporogenous cells (stage 2, Figure 2.1b,h) and vascular and parenchymal tissue (stage 3, Figure 2.1c,i). In contrast to these early stages in fertile buds, during which four laterally symmetrical locules develop, with two identical adaxial locules and two identical abaxial locules, in sterile buds the loss of between one and three locules per anther is evident (Figures 2.1h-l and 2.2b-f). The number of locules in sterile flowers varies from anther to anther (Figure 2.2b). In addition, there is a polarity to the formation of locules within an anther in sterile buds: adaxial locules are more likely to develop, while abaxial locules are more likely to be sterile.

Microspore mother cells of developing CMS locules are often clumped and detached from the tapetum (Figure 2.1i). Interestingly, this does not inhibit their differentiation: meiosis in these locules takes place normally, and tetrads differentiate into pollen as in fertile lines. Fertile anthers develop synchronously, with all locules both within and between anthers passing through different stages of development together (Figure 2.2a). In contrast, *nap* and *pol* CMS anthers can display a loss of synchronous locule development and up to three different developmental stages have been observed per anther in these lines.

Figure 2.2 Polarity of locule formation, locule fusion and petal abnormalities in CMS floral buds.

(a) Fertile floral bud at the microspore mother cell (MMC) stage.
(b) CMS floral bud at the MMC stage.
(c) CMS anther with two synchronously developing adaxial locules (MMC stage); note the small locule in the left abaxial quadrant of the anther (arrow).
(d) CMS anther with only a single adaxial locule.
(e) Early CMS anther with partial fusion of adaxial locules below an anther 'notch.'
(f) Pre-dehiscent pollen grain stage anther lacking a septum between the adaxial locules.
(g) Late-stage fertile bud displaying normal petal (arrow) development.
(h, i) Tubular petals (arrows) in CMS anthers.
(j) Thickened petal (arrow) of a CMS anther.
Size bars = 100μm.


(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)

Additional defects commonly observed in CMS anthers are the fusion of two adaxial locules (Figure 2.2e,f) and the premature separation of the tapetum from the locule wall. Instead of separating from the connective tissue during microspore stage, in sterile anthers the tapetum often separates at the microspore mother cell phase (Figure 2.2c). Despite the numerous morphological changes observed in CMS anthers, the pollen grains that form within developed locules are viable and germinate at the same rates as fertile pollen (data not shown).

No significant differences between the development of *nap* and *pol* CMS anthers were observed (data not shown). This indicates that mechanisms underlying these two CMS types are similar. It is consistent with the similarity in the mitochondrial genes that are associated with the two types of CMS and with the finding that restorer genes for the two types represent different haplotypes of the same nuclear locus. It should be noted that, despite the overall similarity in their male sterile phenotypes, *nap* cytoplasm is normally less effective than *pol* in conferring male sterility and that the *pol* CMS variety examined here, Westar (*pol*), is not as completely sterilized by *pol* cytoplasm as most other strains (Fan *et al.*, 1986). We also observed no differences between the development of fertile *cam* cytoplasm anthers and *nap* or *pol* cytoplasm fertility-restored anthers (data not shown).

Both *nap* and *pol* CMS affect the development of floral organs other than the stamen. In both cases, the bud is misshapen as the sepals and petals grow around malformed anthers. In addition, there is a frequent change of petals from flat

structures, which contour the bud, to tubular structures or to abnormally thickened and mis-shaped organs (Figure 2.2h-j). Open flowers have shrunken petals, in part as a result of this malformation, and also shortened stamen filaments, a common abnormality in other CMS plants.

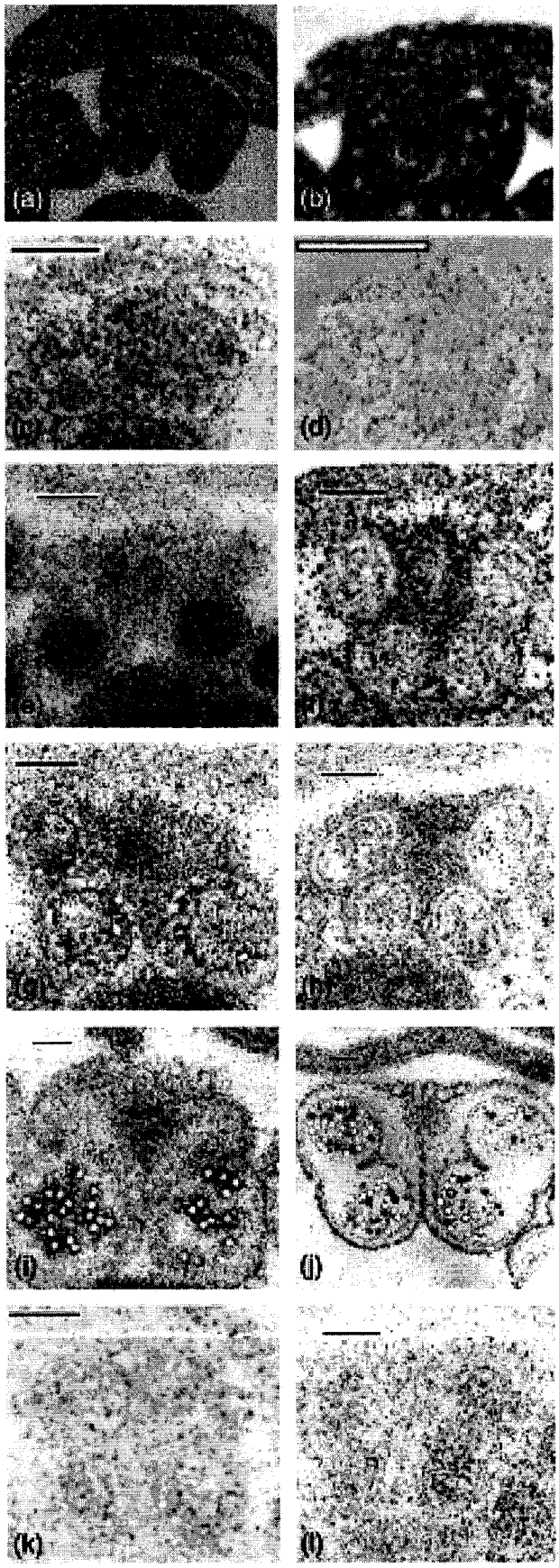
Mitochondrial transcript expression in CMS and restored buds

In situ hybridization was employed to address the role played by the expression of CMS-associated genes in the developmental abnormalities observed in sterile buds. Our approach was to generate probes from those novel portions of the ORFs (i.e. portions that are not derived from *atp8* or *rps3* sequences) that are expressed preferentially in CMS versus restored buds (L'Homme *et al.*, 1997; Menassa *et al.*, 1999; Singh and Brown, 1991). In the case of *orf*224, this region is quite small and we were unable to derive such a probe that hybridized with sufficient specificity as to allow its use in *in situ* hybridization experiments. In the case of *orf*222, the entire novel region is preferentially expressed in CMS plants (L'Homme *et al.*, 1997) and Northern analysis confirmed that the probe used here hybridized in an identical manner to that used previously. We therefore restricted our *in situ* analysis to *orf*222.

Floral tissue sections of *nap* CMS and fertility-restored buds representing most of the stages of anther development described above were hybridized to *orf*222 antisense and sense probes. In sterile buds, the antisense *orf*222 probe hybridizes uniformly over the undifferentiated tissue of the primordial anther (Figure 2.3c). In subsequent early stages of development, *orf*222 transcript abundance is

Figure 2.3 *In situ* hybridization of floral bud tissue with *orf222* sequences.

(a, c, e, g, i, k) *nap* CMS, (b, d, f, h, j, l) fertility-restored *nap*.
(a, b) Primordial anther stage, bright field image of (c) and (d), respectively.
(c-l) *In situ* hybridization, dark field images
(c, d) Primordial anther stage,
(e, f) Microspore mother cell stage anther.
(g, h) Tetrad stage anther.
(i, j) Pre-dehiscent pollen grain anther.
(a-j) Antisense *orf222* probe.
(k, l) Sense *orf222* probe, microspore mother cell stage.
Size bars = 100µm.


(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
(k)
(l)

considerably higher at sites of sporogenous tissue formation and in the vasculature of the anther than in the surrounding connective (Figure 2.3e). After the emergence of the sporogenous cell mass, there is a much lower abundance of transcripts in the areas where sporogenesis failed to take place. At the tetrad stage and throughout the subsequent formation of pollen, levels of *orf222* transcripts remain higher in gametophytic and vascular tissues than in surrounding anther cells, although the degree of expression difference is not as great as at earlier developmental stages (Figure 2.3g,i). High levels of *orf222* transcripts are observed in petals and carpels throughout development (not shown), with particularly high expression in developing ovules. Only low levels of expression are observed in sepals.

In contrast to what is observed in early sterile anther development, *orf222* transcript levels in primordial, fertility-restored anthers are very low, barely detectable above background levels (Figure 2.3d). In the transition to sporogenous tissue formation, *orf222* transcripts accumulate at low levels in the connective, but are notably absent in vascular, tapetal, and sporogenous tissues (Figure 2.3f). By the tetrad stage, levels of *orf222* transcripts in restored buds decrease, and by the final stages of pollen development the pattern of *orf222* transcript accumulation is similar in CMS and restored buds. Differences in *orf222* expression between CMS and restored buds are also observed in other floral organs. Initially, *orf222* transcripts are abundant in sepals, but not in carpels of fertility-restored buds. Later in development, expression in carpel tissue begins to increase and by the end of bud development the pattern of *orf222*

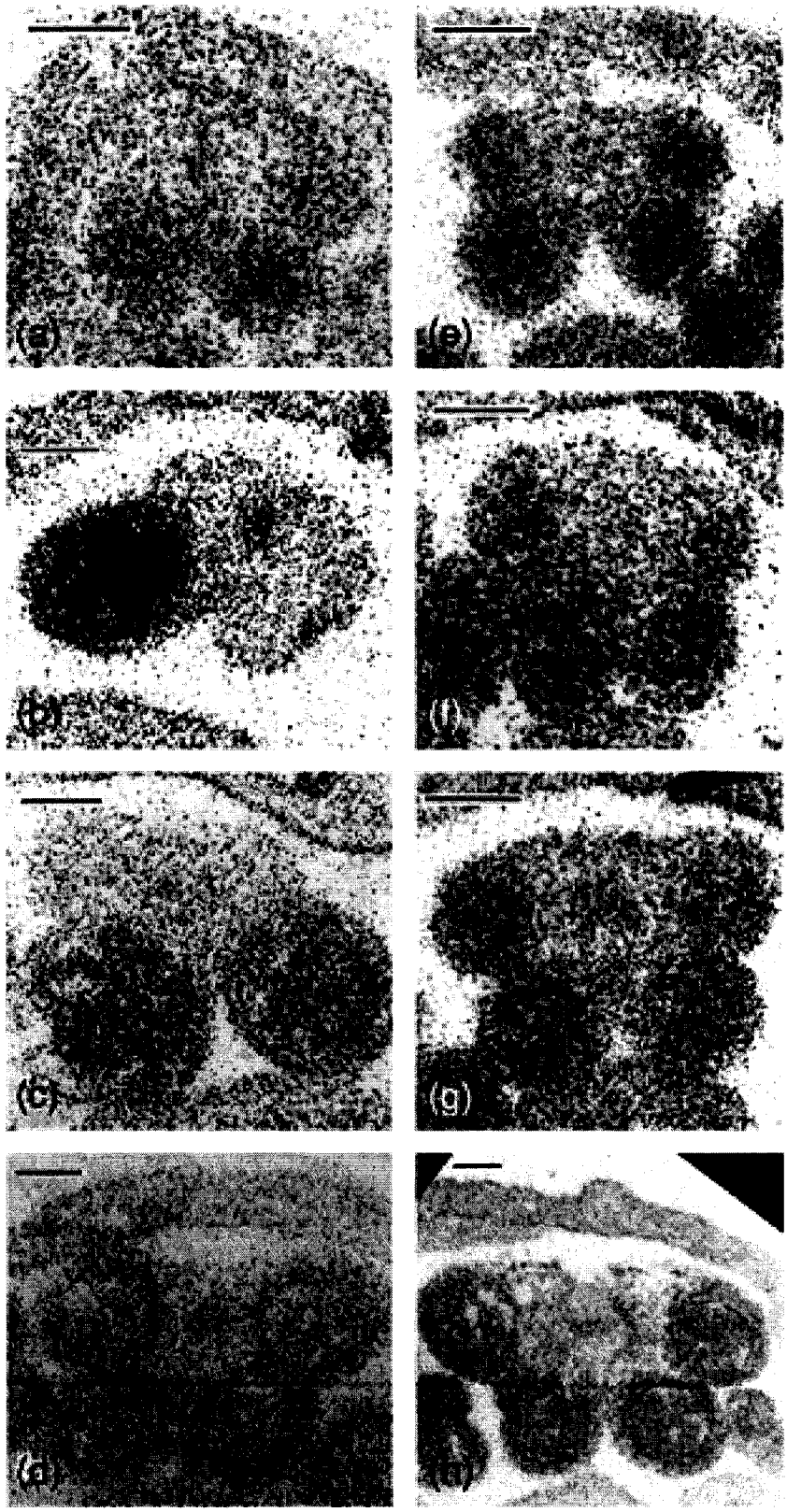
transcript distribution is identical to that found in *nap* CMS: levels are low in sepals, moderate in petals and carpels, and very high in ovules.

To compare the expression of *orf*222 to that of normal mitochondrial genes, corresponding sections of sterile and fertile buds were hybridized with antisense probes to the *B. napus cob* and *atp6* genes, which encode cytochrome *b* and subunit 6 of the ATP synthase, respectively. Interestingly, both genes show the same pattern of expression as *orf*222 in sterile buds (compare Figures 2.3e and 2.4a,b). Unlike *orf*222, however, *atp6* and *cob* show an identical pattern of expression in CMS and restored buds (Figure 2.4a,b,e,f). A similar expression pattern of these genes is observed for *pol* CMS and restored buds (not shown). This indicates that *orf*222 transcript accumulation in sterile plants resembles that of mitochondrial transcripts in general, and that the nuclear restorer gene *Rfn* acts to specifically downregulate *orf*222 transcript accumulation in sporogenous, tapetal and vascular cells as these tissues form during early anther development. Similar to our observations, Smart *et al.* (1994) found that mitochondrial gene transcripts accumulate preferentially in young meiotic cells of sunflower anthers, and that nuclear restoration specifically downregulates the accumulation of the CMS-associated gene, *orf*522.

Balk and Leaver (1998) found that transcripts of the nuclear-encoded subunit F1 beta subunit of the ATP synthase accumulated in tapetal cells but not in meiocytes in developing sunflower anthers, while transcripts of the mitochondrially encoded alpha subunit accumulated in both tissue types (Smart *et al.*, 1994). To determine if a similar lack of congruence in the pattern of

Figure 2.4 *In situ* hybridization of floral bud tissue with antisense cytochrome *b*, *atp6* and ATP synthase gamma subunit probes.

(a) CMS anther-*atp6* probe.
(b) CMS anther-*cob* probe.
(c) CMS anther-ATP synthase gamma subunit probe.
(d) Fertility-restored anther (meiocyte stage)-*cob* probe.
(e) Fertility restored anther-*atp6* probe.
(f) Ferility restored anther-*cob* probe;
(g) Ferility restored anther-ATP synthase gamma subunit probe.
(h) Fertility-restored anther (meiocyte stage)-ATP synthase gamma subunit probe.
(a-f) Microspore mother cell stage anthers.
Scale bar = 100μm.


(a)
(e)
(b)
(f)
(c)
(g)
(d)
(h)

transcript accumulation between nuclear and mitochondrial-encoded subunits occurs during anther development in *B. napus*, sections of sterile and fertile buds were hybridized to a probe for the nuclear-encoded ATP synthase gamma subunit gene. In contrast to what was observed for sunflower beta subunit transcripts, the *Brassica* gamma subunit transcripts (Figure 2.4c,g,h) accumulate in patterns indistinguishable from transcripts of the mitochondrial genes *atp6* and *cob* at both early and later (Figure 2.4d,h) stages of anther development in both sterile and fertile plants.

Altered regulation of the APETALA3 promoter in CMS versus fertile B. napus

The structures of both petals and stamens are altered in *nap* and *pol* CMS flowers (Fan and Stefansson, 1986; Figure 2.2, above) and this correlates with the accumulation of *orf222* transcripts during the early development of both organ types. In fertility-restored plants, *orf222* transcripts accumulate to significant levels only at the later stages of development in these tissues. Significantly, carpel development and ovule formation are not affected in CMS plants, although *orf222* transcript levels remain moderate to high in these tissues throughout their development. B function MADS box genes specify the identity of petals and stamens, but not sepals or carpels (Irish, 1999; Jack, 2004). It seemed possible therefore, that *orf222* expression might act to alter stamen and petal development in CMS plants by influencing the expression of B function genes. To test this possibility, we examined the activity of a promoter element that drives expression of the Arabidopsis B function gene *APETALA3* (*AP3*) gene, in fertile and CMS flower buds. A region of the Arabidopsis genome extending from the

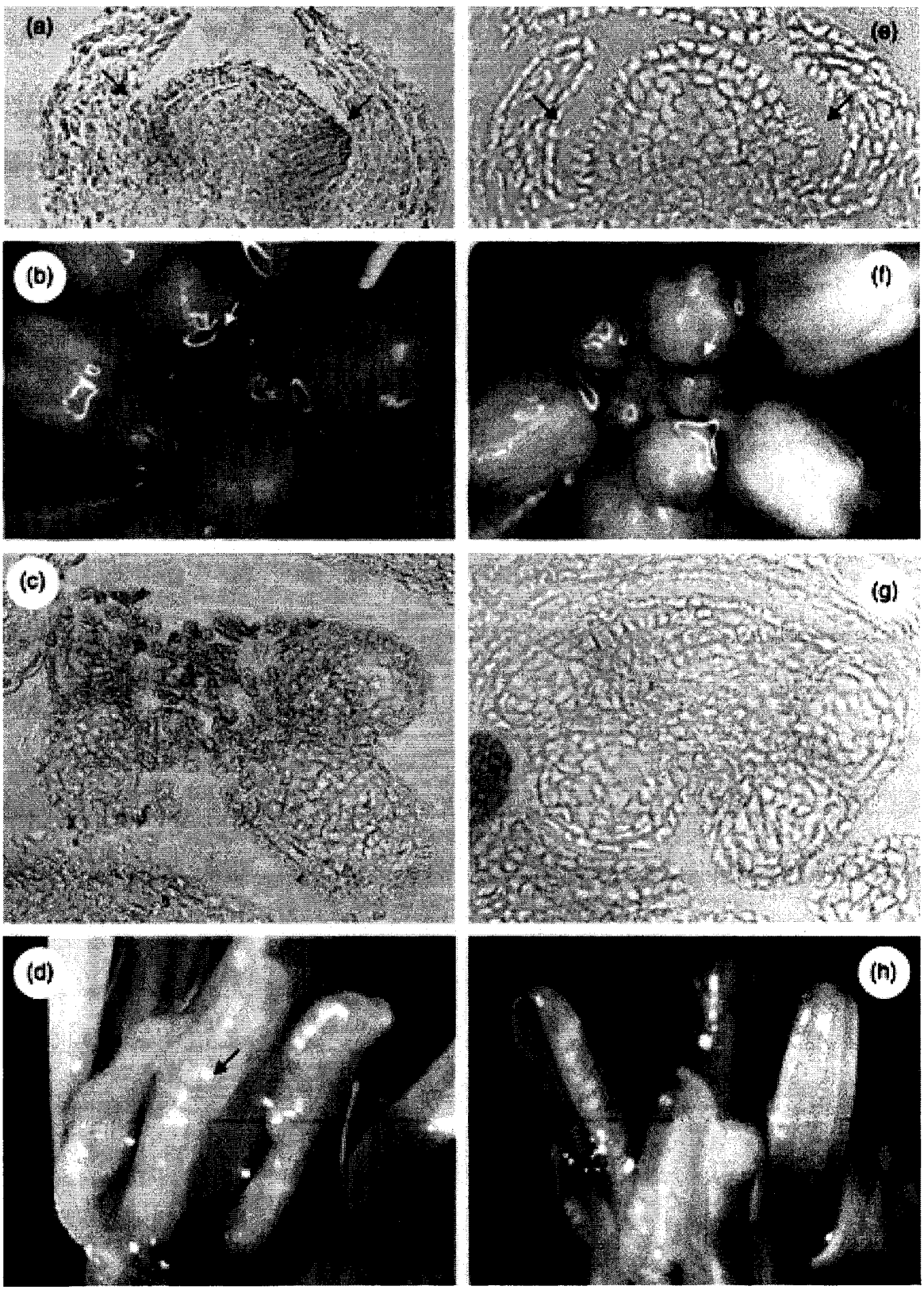
start of the *AP3* coding sequence to a site 458 bp upstream (Jack *et al.*, 1992; Tilly *et al.*, 1998) was fused upstream of the *Escherichia coli* beta-glucuronidase gene (the AP3::GUS construct) and introduced into fertile and CMS *B. napus* plants. Bud sections taken at different stages of floral development were examined for GUS expression by histochemical staining.

At the earliest stages of flower formation, as the floral meristem differentiates from the inflorescence, no *AP3*-driven GUS expression could be observed in either sterile or fertile plants. Shortly after this (stages 3 and 4 of Smyth *et al.*, 1990), GUS staining could be observed in and around the stamen primordia of sterile buds (Figure 2.5a), but not of fertile buds (Figure 2.5e). Differences in early GUS expression could be seen in whole buds as well as in sections (Figure 2.5b,f). In fertile plants, GUS expression could be detected in the petal and stamen primordia but only at a subsequent stage (stage 5 of Smyth *et al.*, 1990), after these primordia have differentiated from the floral meristem.

Analysis of cross sections of young (sporogenous cell stage) anthers revealed additional differences in *AP3* promoter activity between sterile and fertile anthers. In sporogenous cell stage sterile anthers, prior to the onset of meiosis, GUS staining was observed within most or all developing locules of the anthers (Figure 2.5c). This staining tended to be more intense in the abaxial region, in which locules are less likely to develop; in some cases, GUS staining was reduced or absent in the larger adaxial locule regions that tend to develop into microsporangia (Figure 2.5c). Staining was also observed in the region destined

Figure 2.5 *AP3* promoter activity in CMS and male-fertile flowers as measured by expression of an AP3::GUS construct.

(a) Longitudinal section through a young CMS floral bud (stage 4, Smyth *et al.*, 1990); arrows indicate the stamen primordia; blue staining results from GUS expression.
(b) Young CMS inflorescence; significant GUS activity can be observed at the site developing stamen formation (arrow).
(c) Cross section through a sporogenous cell stage CMS anther; GUS activity is observed within the sporogenous cells of the developing locules.
(d) Late-stage CMS flower showing pre-dehiscent anthers; a low level of GUS activity can be seen in the connective between the anther locules (arrow).
(e) Longitudinal section through a stage 4 (Smyth *et al.*, 1990) male fertile anther; no GUS activity is observed in the developing stamen primordia (arrows).
(f) Young male fertile inflorescence; little GUS activity is evident.
(g) Cross section through a sporogenous cell stage fertile anther; GUS activity is low or absent in the sporogenous cells of the developing locules and reduced in the region that will develop into the vasculature of the mature anther.
(h) Late-stage male fertile flower showing pre-dehiscent anthers. A high level of GUS activity can be seen in the connective between the anther locules (arrow).


(a)
(e)
(b)
(f)
(c)
(g)
(d)
(h)

to develop into the vascular system of the sterile anther. In contrast, in fertile anthers there is little if any GUS staining in the developing locules and GUS staining is reduced in the developing vasculature (Figure 2.5g). This pattern stands in marked contrast to the distribution of *orf222* transcript levels, which are high in developing locules and vasculature of sterile anthers, but are very low in the corresponding tissues of fertile anthers.

At later stages of floral development, high levels of GUS expression were observed in petals and stamens of sterile and fertile plants (not shown), while during the same period GUS expression was reduced in stamens. Shortly before anthesis GUS expression was evident in the connective of fertile and sterile stamens (Figure 2.5d,h) with the intensity of GUS staining in late-stage anthers always stronger in fertile than in sterile plants.

No GUS activity was detected in sterile or fertile sepals or early-stage carpels, or at any time in control (non-transformed) plants. These differences in activity of the promoter element were consistent across multiple sterile and fertile transgenic plants, and were heritable through successive generations. These observations suggest that the timing and level of *AP3* promoter activity differs between CMS and fertile *Brassica* plants. Moreover, in sterile buds, the *AP3* promoter is active in the developing vasculature and in regions that are unlikely to develop into microsporangia; in fertile buds this activity is markedly lower in the developing microsporangia and vasculature. These latter observations further indicate that at a crucial stage in pollen development, the domains of *AP3* promoter activity differ between CMS and fertile anthers.

Discussion

Novel morphological aspects of CMS in Brassica

The abnormalities associated with the *nap* and *pol* forms of CMS in *B. napus* appear very early in anther development: the microsporocyte and other locule features at some, and occasionally all sites within CMS anthers, fail to form. We note here several additional and stereotypic characteristics of the *Brassica* CMS phenotype including a disruption of the symmetry of anther development and a polarity to the developmental abnormalities, such that locules proximal to the carpel are significantly more likely to form than locules distal to the carpel. Although Theis and Röbbelen (1990) noted the failure of locules to form in *Brassica* CMS anthers, to our knowledge, the additional features of *Brassica* CMS that we describe have not been thus far described for any other CMS system.

Spatial and temporal regulation of mitochondrial transcript accumulation

The finding that *orf*222 transcripts accumulate prior to any morphological abnormality in CMS buds is consistent with a role for this gene, and by extension, the related *orf*224 gene, in specifying male sterility. The additional finding that the tissue-specific accumulation of these transcripts is downregulated through restorer gene action is consistent with this interpretation.

All the mitochondrial gene transcripts we examined other than *orf*222 in restored plants displayed similar patterns of cell-specific accumulation in developing

sporogenous, meiotic and tapetal tissues. This suggests that the increased accumulation is due to enhanced activity of a function that affects mitochondrial transcripts as a collective unit, such as mitochondrial RNA polymerase or an associated transcription factor. In this respect, the lack of a similar increase in *orf*222 transcripts in restored plants likely results from a corresponding enhancement in the activity of the *Rfn* restorer gene product that downregulates *orf*222 transcript accumulation.

In CMS plants, the global increase in mitochondrial transcripts in specific tissues of the anther extends to the transcripts of the CMS-associated gene *orf*222. The N-terminal coding portion of *orf*222 is derived from the *atp8* gene, which encodes a subunit of the mitochondrial ATP synthase. As suggested by Sabar *et al.* (2003), it is possible that expression of a CMS-associated gene product that resembles a normal mitochondrial protein can cause mitochondrial dysfunction by interfering with the function of the normal counterpart. The ATP8 protein is thought to play an important role in ATP synthase assembly (Devenish *et al.*, 2000), and it is possible, therefore, that in *Brassica* CMS, defective assembly of the synthase during periods of active mitochondrial biogenesis in sporogenous tissue results in a degree of mitochondrial dysfunction that is sufficient to alter normal anther development.

Smart *et al.* (1994) similarly observed enhanced expression of mitochondrial genes during microsporogenesis in sunflower anthers, suggesting that this is common in plants, and likely reflects very active mitochondrial biogenesis in these cells. We have, in addition, found that transcripts of the nuclear-encoded F1

ATP synthase gamma subunit accumulate in a tissue-specific manner during anther development that precisely parallels that of the corresponding mitochondrial encoded subunit 6 of F0. In contrast, Balk and Leaver (1998) have found that transcripts of the sunflower nuclear-encoded F1 beta subunit, like those of the mitochondrial-encoded alpha subunit, accumulate preferentially in pre-meiotic cells and meiocytes, but only in tapetal cells in the later stages of anther/pollen development. Similarly, Ribichich *et al.* (2001) found patterns of cytochrome *c* transcript accumulation in sunflower florets that largely parallel those described by Balk and Leaver (1998) for beta subunit transcripts. Taken together, these observations suggest that the expression of nuclear genes encoding respiratory chain components is coordinated with the expression of their mitochondrial counterparts during anther formation, but that variation may be observed in the degree to which this coordination is evident in patterns of transcript accumulation.

The investigations of Conley and Hanson (1994) suggest that patterns of mitochondrial gene expression during anther development may show more significant differences at the protein than at the transcript levels. These authors found that the protein products of the mitochondrial genes *atpA* and *pcf*, a CMS-associated gene, as well the nuclear-encoded mitochondrial alternative oxidase, accumulate in a manner resembling the nuclear and mitochondrial transcripts examined here. In contrast, they observed a dramatically different pattern of accumulation for the cytochrome oxidase subunit II protein. These observations further indicate that all mitochondrial proteins do not accumulate in the same

tissue-specific pattern during anther development and that some of these proteins may show quite different patterns of expression.

orf222 transcript accumulation and the CMS phenotype

Two related observations suggest that *orf222* expression blocks microsporangium development at a very early and specific stage. First, there is no evidence for very early events in sporogenesis at those sites in CMS anthers where locules fail to form. Secondly, at those anther sites where sporogenesis takes place, *orf222* transcripts, along with those of other mitochondrial genes, accumulate to high levels. Interestingly, this latter observation further suggests that *orf222* expression does not have a significant effect on male gametogenesis once an early critical stage is passed, as functional pollen grains form within those locules that develop in CMS anthers.

orf222 expression appears to be evenly distributed over the developing CMS anther primordia and only later becomes localized at sites of sporogenesis. When microsporangia appear in CMS anthers, they do so only at sites where they form in fertile anthers. It is conceivable that early *orf222* expression could lead to the formation of zones within the developing anther in which the initiation of sporogenesis cannot be triggered and these zones would arise more frequently at sites of abaxial locule initiation. In some cases, this effect might be sufficiently restrictive so as to completely prevent sporogenesis from taking place, whereas in other cases, the restriction might simply delay the onset of sporogenesis such that later forming microsporangia and locules would be smaller than those forming

earlier. This would explain the asynchronous and asymmetric formation of locules observed for CMS anthers.

B function genes such as *AP3* play crucial roles in specifying the identity and developmental patterning of petals and stamens (Jenik and Irish, 2001). In particular, *AP3* function is required for proper formation of second and third whorl organs throughout the course of their development (Bowman *et al.*, 1989). Our findings indicate that the timing of the activity of an *AP3* promoter element is altered in CMS versus fertile stamens and petals. This, in turn, suggests that the diverse aspects of the CMS phenotype that we describe here, including both stamen and petal alterations, may result from altered expression of *AP3* and/or other genes such as *PISTILLATA*, which require *AP3* for optimal expression. Interestingly, the *AP3* promoter construct used in these experiments lacks an upstream element that is required for maximal early expression (Hill *et al.*, 1998; Tilly *et al.*, 1998). It is possible, therefore, that in CMS plants, the requirement for this element is reduced, possibly through the enhanced activity or expression of transcription factors controlling *AP3* expression.

The possibility of a direct relationship between *orf*222 expression, *AP3* expression and the CMS phenotype is strengthened by the differences seen in the domains of *AP3* promoter activity between sterile and fertile anthers prior to the onset of meiosis. Fertile locules, which possess very reduced levels of *orf*222 transcripts, show little or no *AP3* promoter activity in comparison with the surrounding connective. In contrast, in locules of sterile anthers, which possess abundant levels of *orf*222 transcripts, this activity appears similar to that of the surrounding

connective tissue. This suggests that mitochondrial dysfunction triggered by *orf222* expression may interfere with the downregulation of *AP3* expression.

In *Arabidopsis*, *AP3* expression is established by multiple regulatory pathways that involve the transcription factors LEAFY and APETALA1 and F-box proteins such as UFO, which are thought to be components of a complex that targets proteins for ubiquitin-dependent degradation (Jack, 2004). It is possible that mitochondrial dysfunction resulting from expression of CMS-associated ORFs early in stamen and petal development affects the establishment or function of one of these pathways.

In *Arabidopsis*, the B-function is defined by *AP3* and *PISTILLATA* (*PI*); the product of the latter forms a dimer with the AP3 protein. *AP3* activity is required for the expression of *PI* and once established, *AP3* and *PI* together are required for the expression of a limited number genes that play as yet poorly understood roles in stamen/petal development (Zik and Irish, 2003). It is conceivable that mis-expression of *AP3* and/or genes regulated by *AP3* such as *NAP* (*NAC- LIKE, ACTIVATED BY AP3/PI*; Sablowski and Meyerowitz, 1998) early in the development of CMS anthers may block the transition to later, sporogenic stages. If this were to occur preferentially at specific sites in the anther, it could explain the observed aberrant pattern of microsporangium formation. It is conceivable, in fact, that the downregulation of *AP3* promoter activity that we have observed in fertile anthers is required for further differentiation of the microsporangium, and that the failure of this to occur in CMS anthers is related to the observed disruption of locule development. This possibility is supported by the finding

that little or no *AP3* promoter activity is observed in the large adaxial lobes of CMS anthers which differentiate into microsporangia.

The phenomenon by which mitochondrial dysfunction leads to altered expression of nuclear genes is termed retrograde regulation; it affects a large number of genes and is known to occur in fungi, metazoans and plants (Butow and Avadhani, 2004; Yu *et al.*, 2001). Several other cases where the expression or function of floral patterning genes may be altered in CMS, apparently as a result of retrograde regulation have been described (Zubko, 2004). Zubko *et al.* (2001) found that a B function gene similar to the snapdragon *PI* ortholog *GLOBOSA* was downregulated in a tobacco alloplasmic CMS. Similarly, Linke *et al.* (2003) found that in a carrot CMS characterized by a transformation of stamens into petal-like structures, expression of homologs of *GLOBOSA* and the snapdragon *AP3* ortholog *DEFICIENS* are reduced while Murai *et al.* (2002) have reported downregulation of an *AP3*-like gene in a wheat alloplasmic CMS in which stamens are transformed into pistil-like structures. Bereterbide *et al.* (2002) found that an alloplasmic CMS phenotype in *Nicotiana* can be suppressed by ectopic expression of the *Arabidopsis SUPERMAN* gene, which controls the formation of the boundary between whorl 3 and whorl 4 organs. These findings suggest that alterations in the expression or functions of genes that govern organ identity and patterning may be characteristic of those CMS systems in which the phenotype extends to floral characteristics other than anther or pollen development *per se*.

For many examples of CMS, the abnormalities associated with male sterility are restricted to late anther development, and do not extend to other floral organs

such as petals or even to the stamen filament. Such cases are generally characterized by a premature degeneration of the tapetum during meiocyte meiosis, resulting in the death of immature microspores. For one system of this type, sunflower PET1-CMS, Balk and Leaver (2001) showed that premature tapetal degeneration is characterized by several hallmarks of programmed cell death or apoptosis, including cytochrome *c* release from mitochondria and oligonucleosomal cleavage of nuclear DNA. Thus male sterility in this system is associated with an alteration in the timing of tapetal apoptosis. This is analogous to the alteration in the timing of expression *AP3* as suggested above to explain the *Brassica* CMS phenotype.

Our observations indicate that mitochondrial gene transcript accumulation is subject to a high degree of temporal and spatial regulation, not only during anther formation, but also during the formation of other floral organs as well. Although previous work (Menassa *et al.*, 1999; Singh and Brown, 1993) indicated that *Brassica* mitochondrial gene expression and restorer gene action was subjected to developmental regulation, the extent of this regulation, as evident in these *in situ* hybridization studies, is considerably greater than earlier work indicated.

Our observations suggest that the cell-specific expression of *orf*222 early in stamen development blocks initiation of sporogenesis at localized sites and according to a stereotypical pattern in CMS anthers. Our results further suggest that this effect may be mediated through an effect *orf*222 expression may have on the expression of nuclear genes that control anther development such as *AP3*. The

mechanisms through which this is achieved and identification of the specific genes involved represent interesting challenges for future studies.

References

Alexander, M.P. (1969) Differential staining of aborted and nonaborted pollen. Stain Technol. **41**, 117-122.

Balk, J. and Leaver, C.J. (1998) Cell-specific differences in the expression of a nuclear and mitochondrial transcript of ATP synthase during anther development. In *Plant Mitochondria: From Gene to Function* (Møller, I.M., Gardeström, P., Glimelius, K. and Glaser, E., eds). Leiden, The Netherlands: Backhuys Publishers, pp. 57-61.

Balk, J. and Leaver, C.J. (2001) The PET1-CMS mitochondrial mutation in sunflower is associated with premature programmed cell death and cytochrome c release. Plant Cell, **13**, 1803-1818.

Bereterbide, A., Hernould, M., Farbos, I., Glimelius, K. and Mouras, A. (2002) Restoration of stamen development and production of functional pollen in an alloplasmic CMS tobacco line by ectopic expression of the *Arabidopsis thaliana* SUPERMAN gene. Plant J. **29**, 607-615.

Bowman, J.L., Smyth, D.R. and Meyerowitz, E.M. (1989) Genes directing flower development in Arabidopsis. Plant Cell, **1**, 37-52.

Brown, G.G. (1999) Unique aspects of cytoplasmic male sterility and fertility restoration in *Brassica napus*. J. Hered. **90**, 351-356.

Butow, R.A. and Avadhani, N.G. (2004) Mitochondrial signaling: the retrograde response. Mol. Cell, **14**, 115.

Carpenter, J.L., Ploense, S.E., Snustad, D.P. and Silflow, C.D. (1992) Preferential expression of an alpha-tubulin gene of Arabidopsis in pollen. Plant Cell, **4**, 557-571.

Conley, C.A. and Hanson, M.R. (1994) Tissue-specific protein expression in plant mitochondria. Plant Cell, **6**, 85-91.

Datla, R.S., Hammerlindl, J.K., Panchuk, B., Pelcher, L.E. and Keller, W. (1992) Modified binary plant transformation vectors with the wild-type gene encoding NPTII. Gene, **122**, 383-384.

Devenish, R.J., Prescott, M., Roucou, X. and Nagley, P. (2000) Insights into ATPase assembly and function though the molecular genetic manipulation of subunits of the yeast mitochondrial enzyme complex. Biochim. Biophys. Acta, **1458**, 428-442.

Fan, Z. and Stefansson, B.R. (1986) Influence of temperature on sterility of two cytoplasmic male-sterility systems in rape (*Brassica napus* L.). Can. J. Plant Sci. **66**, 221-227.

Fan, Z., Stefansson, B.R. and Sernyk, J.L. (1986) Maintainers and restorers for three male-sterility inducing cytoplasms in rape (*Brassica napus* L.). Can. J. Plant Sci. **66**, 229-234.

Gabay-Laughnan, S., Zabala, G. and Laughnan, J.R. (1995) S-type cytoplasmic male sterility in maize. In *The Molecular Biology of Plant Mitochondria* (Levings, C.S. and Vasil, I.K., eds). Dordrecht: Kluwer Academic Publishers, pp. 395-432.

Hanson, M. and Bentolila, S. (2004) Interactions of mitochondrial and nuclear genes that affect male gametophyte development. Plant Cell, **16 (Suppl.)**, S154-S169.

Hanson, M.H. and Conde, M.F. (1985) Functioning and variation of cytoplasmic genomes: Lessons from cytoplasmic-nuclear interactions affecting male fertility in plants. Int. Rev. Cytol. **94**, 213-267.

Hill, T.A., Day, C.D., Zondlo, S.C., Thackeray, A.G. and Irish, V.F. (1998) Discrete spatial and temporal *cis*-acting elements regulate transcription of the Arabidopsis floral homeotic gene APETALA3. Development, **125**, 1711-1721.

Irish, V.F. (1999) Patterning the flower. Dev. Biol. **209**, 211-220.

Jack, T. (2004) Molecular and genetic mechanisms of floral control. Plant Cell, **16 (Suppl.)**, S1-S17.

Jack, T., Brockman, L.L. and Meyerowitz, E.M. (1992) The homeotic gene *APETALA3* of Arabidopsis thaliana encodes a MADS box and is expressed in petals and stamens. Cell, **68**, 683-697.

Jean, M., Brown, G.G. and Landry, B.S. (1997) Genetic mapping of nuclear fertility restorer genes for the 'Polima' cytoplasmic male sterility in canola (*Brassica napus* L.) using DNA markers. Theor. Appl. Genet. **95**, 321-328.

Jenik, P.D. and Irish, V.F. (2001) The Arabidopsis floral homeotic gene *APETALA3* differentially regulates intercellular signaling required for petal and stamen development. Development, **128**, 13-23.

Kemble, R.J. (1987) A rapid, single leaf, nucleic acid assay for determining the cytoplasmic organelle complement of rapeseed and related Brassica species. Theor. Appl. Genet. **73**, 364-370.

Kemble, R.J. and Barsby, T.L. (1988) Use of protoplast fusion systems to study organelle genetics in a commercially important crop. Biochem. Cell. Biol. **66**, 665-676.

Kofer, W., Glimelius, K. and Bonnett, H.T. (1991) Modifications of mitochondrial DNA cause changes in floral development in homeotic-like mutants of tobacco. Plant Cell, **3**, 759-769.

L'Homme, Y. and Brown, G.G. (1993) Organizational differences between cytoplasmic male sterile and male fertile Brassica mitochondrial genomes are confined to a single transposed locus. Nucleic Acids Res. **21**, 1903-1909.

L'Homme, Y., Stahl, R.J., Li, X.-Q., Hameed, A. and Brown, G.G. (1997) Brassica *nap* cytoplasmic male sterility is associated with expression of a mtDNA region containing a chimeric gene similar to the *pol* CMS-associated *orf224* gene. Curr. Genet. **31**, 325-335.

Li, X.Q., Jean, M., Landry, B.S. and Brown, G.G. (1998) Restorer genes for different forms of brassica cytoplasmic male sterility map to a single nuclear locus that modifies transcripts of several mitochondrial genes. PNAS, **95**, 10032-10037.

Linke, B., Nothnagel, T. and Borner, T. (2003) Flower development in carrot CMS plants: mitochondria affect the expression of MADS box genes homologous to *GLOBOSA* and *DEFICIENS*. Plant J. **34**, 27-37.

Meierhoff, K., Felder, S., Nakamura, T., Bechtold, N. and Schuster, G. (2003) HCF152, an Arabidopsis RNA binding pentatricopeptide repeat protein involved in the processing of chloroplast psbB-psbT-psbH-petB-petD RNAs. Plant Cell, **15**, 1480-1495.

Menassa, R., L'Homme, Y. and Brown, G.G. (1999) Post-transcriptional and developmental regulation of a CMS-associated mitochondrial gene region by a nuclear restorer gene. Plant J. **17**, 491-499.

Moloney, M.M., Walker, J.M. and Sharma, K.K. (1989) High-efficiency transformation of *Brassica napus* using *Agrobacterium* vectors. Plant Cell Rep. **8**, 238-242.

Murai, K., Takumi, S., Koga, H. and Ogihara, Y. (2002) Pistillody, homeotic transformation of stamens into pistil-like structures, caused by nuclear-cytoplasm interaction in wheat. Plant J. **29**, 169-181.

Ribichich, K.F., Tioni, M.F., Chan, R.L. and Gonzalez, D.H. (2001) Cell-type-specific expression of plant cytochrome c mRNA in developing flowers and roots. Plant Physiol. **125**, 1603-1610.

Robison, M.M. and Wolyn, D.J. (2002) Complex organization of the mitochondrial genome of petaloid CMS carrot. Mol. Genet. Genomics, **268**, 232-239.

Sabar, M., Gagliardi, D., Balk, J. and Leaver, C.J. (2003) ORFB is a subunit of F1F0-ATP synthase: insight into the basis of cytoplasmic male sterility in sunflower. EMBO Rep. **4**, 381-386.

Sablowski, R.W. and Meyerowitz, E.M. (1998) A homolog of *NO APICAL MERISTEM* is an immediate target of the floral homeotic genes *APETALA3/PISTILLATA*. Cell, **92**, 93-103.

Sanders, P.M., Bui, A.Q., Weterings, K., McIntire, K.M., Hsu, Y.-C., Lee, P.Y., Truong, M.T., Beals, T.P. and Goldberg, R.B. (1999) Anther developmental defects in *Arabidopsis thaliana* male-sterile mutants. Sex. Plant Reprod. **11**, 297-322.

Singh, M. and Brown, G.G. (1991) Suppression of cytoplasmic male sterility by nuclear genes alters expression of a novel mitochondrial gene region. Plant Cell, **3**, 1349-1362.

Singh, M. and Brown, G.G. (1993) Characterization of expression of a mitochondrial gene region associated with the Brassica 'Polima' CMS: developmental influences. Curr. Genet. **24**, 316-322.

Small, I.D. and Peeters, N. (2000) The PPR motif a TPR-related motif prevalent in plant organellar proteins. Trends Biochem. Sci. **25**, 46-47.

Smart, C.J., Monéger, F. and Leaver, C.J. (1994) Cell-specific regulation of gene expression in mitochondria during anther development in sunflower. Plant Cell, **6**, 811-825.

Smyth, D.R., Bowman, J.L. and Meyerowitz, E.M. (1990) Early flower development in Arabidopsis. Plant Cell, **2**, 755-767.

Theis, R. and Röbbelen, G. (1990) Anther and microspore development in different male sterile lines of oilseed rape (*Brassica napus* L.). Angew. Botanik, **64**, 419-434.

Tilly, J.J., Allen, D.W. and Jack, T. (1998) The CArG boxes in the promoter of the Arabidopsis floral organ identity gene *APETALA3* mediate diverse regulatory effects. Development, **125**, 1647-1657.

Yu, J., Nickels, R. and McIntosh, L. (2001) A genome approach to mitochondrial-nuclear communication in Arabidopsis. Plant Physiol. Biochem. **39**, 345-353.

Zik, M. and Irish, V.F. (2003) Global identification of target genes regulated by *APETALA3* and *PISTILLATA* floral homeotic gene action. Plant Cell, **15**, 207-222.

Zubko, M.K. (2004) Mitochondrial tuning fork in nuclear homeotic functions. Trends Plant Sci. **9**, 61-64.

Zubko, M.K., Zubko, E.I, Ruban, A.V., Adler, K., Mock, H.P., Misera, S., Gleba, Y.Y. and Grimm, B. (2001) Extensive developmental and metabolic alterations in cybrids *Nicotiana tabacum* (+*Hyoscyamus niger*) are caused by complex nucleo-cytoplasmic incompatibility. Plant J. **25**, 627-639.

CHAPTER 3

CELL-SPECIFIC ANTISENSE TRANSCRIPTS OF A MITOCHONDRIAL GENE ENCODING A SUBUNIT OF THE ATP SYNTHASE COMPLEX IN *BRASSICA NAPUS*

As part of studies on CMS-associated genes in *Brassica napus* during stamen development (Chapter 2), expression patterns of additional transcripts were analyzed. Since it is thought that expression of CMS inducing genes leads to an impairment of energy production, this analysis was performed on mitochondrial genes encoding proteins of the mitochondrial electron transport chain. Sense probes were used to confirm the absence of antisense transcripts and to rule out non-specific probe binding.

This chapter was written by Rachel Geddy and edited by Dr. Gregory G. Brown.

Introduction

Most of the proteins in the mitochondria are coded nuclear-encoded and synthesized on cytoplasmic ribosomes, but a smaller number, 30-40 in angiosperms (Lang *et al.*, 1999), are encoded in mitochondrial DNA and synthesized on mitochondrial ribosomes. Thus, mitochondrial biogenesis requires the coordinated operation of two different genetic systems. In most organisms, our understanding of the mechanisms that function to ensure this coordinated regulation is limited. It is clear, however, that post-transcriptional mechanisms play important roles in controlling the production of individual proteins within the mitochondria and these can act by modifying mRNA processing, stability, translation and/or protein stability (Lang *et al.*, 1999). Of particular interest are the mechanisms that control the production of respiratory complexes I, III, IV and V, in which nuclear and mitochondrially encoded subunits occur in fixed stoichiometric amounts (Lang *et al.*, 1999).

Cytoplasmic male sterility (CMS) is a deficiency in pollen production caused by novel, often chimeric, mitochondrial genes. It is thought that expression of these novel genes leads to a dysfunction in energy production, which in turn leads to male sterility (Brown, 1999). The CMS phenotype can be suppressed by nuclear restorer of fertility (Rf) genes which generally act to down-regulate the expression of CMS-associated mitochondrial genes. CMS may be a useful system with which to study nuclear-mitochondrial genome interactions.

Recent years have witnessed growing appreciation for the role played by antisense transcripts in gene regulation (Lavorgna *et al*, 2004). In plants, antisense RNAs are known to trigger down-regulation of sense transcripts through a variety of mechanisms (Wassenegger and Pelissier, 1998; Meister, 2004), including RNA-induced transcriptional silencing and mRNA degradation (RNA interference or RNAi) (Baulcombe, 2004). In bacteria, antisense RNAs play critical roles in controlling plasmid copy number, transposition frequency and conjugation (Wagner and Flardh, 2002).

Antisense RNAs have also been found in organelles. The mammalian mitochondrial genome is completely symmetrically transcribed with individual mRNA species generated though the processing of polycistronic precursors. In chloroplasts, antisense transcripts have been proposed to stabilize transcripts by protecting them from 3′ to 5′ exonuclease activity (Nishimura *et al.*, 2004) and to reduce the efficiency of RNA editing (Hegeman *et al.*, 2005). With respect to plant mitochondria, Akagi *et al.* (1994) found antisense transcripts of a region downstream of the rice *atp6* gene accumulated to higher levels in CMS than in fertility restored plants, and suggested that the antisense transcripts may be associated with the CMS trait.

We have employed *in situ* hybridization to analyze the expression of genes encoding mitochondrial proteins during stamen development in CMS and fertility restored *Brassica napus* (Geddy *et al.*, 2005). We have made the unexpected finding that antisense transcripts of the mitochondrial *atp6* gene, which encodes subunit 6 of the F_o component of the ATP synthase complex

(complex V), accumulate in a tissue-specific manner comparable to that observed for the sense transcripts. These antisense transcripts can be detected by Northern and RT-PCR analyses as well. In contrast, only sense transcripts of another mitochondrial gene were detected.

Materials and Methods

Plant material

Brassica napus cytoplasms are designated according to the convention of Kemble and Barsby (1988) and are indicated by italicized designations. The different *B. napus* strains used in this study have been described previously (L'Homme *et al.*, 1997). Plants were grown in the McGill Phytotron growth chambers under normal growth conditions as described (Singh and Brown, 1991).

Nucleic acid extraction

Total RNA was extracted from 8 day old seedlings and unopened clusters of floral buds using the RNA isolation reagent RNAwizTM by Ambion (Austin, TX, USA) according to the supplier's instructions. Mitochondrial RNA was extracted using a method similar to Singh and Brown (1991) with the following changes. Plant tissue was homogenized in 5 volumes of high ionic strengh buffer as described by Singh and Brown (1991), with the omission of beta-mercaptoethanol, and with the addition of a second grinding of the homogenate caught in the double layer miracloth filter in an additional 5 volumes of buffer. Mitochondria were lysed at room temperature for 10 min with gentle shaking. Precipitation of RNA was carried out at 4°C overnight and centrifuged at 13 000 *g* for 15 min before being resuspended in water and stored at 70°C.

Hybridization probes

Primers were designed to amplify *B. napus* mitochondrial genes by the polymerase chain reaction, using purified *B. napus* mtDNA as a template. These primers were then used to amplify the corresponding sequence from total cellular *B. napus* DNA. The sequences of the various primers employed for amplification are as follows:

atp6: 5'ATATGCGGCCGCTTGTCCCATTGATTCCTA3',
5'ATCGTCTAGAAGAAAATCTCATTCATACATAGCA3;
cytochrome *b* (cob): 5'GTTACTGAAGGGTTGTCGACTG3',
5'GGGCGAATTAAGCGTCGACG3'.

These primers amplified the following segments of the corresponding genes, with nucleotide 1 being the first nucleotide of the respective coding sequence; *atp6*: nucleotides 50 through 650; *cob*: nucleotides 73 through 281.

PCR products were cloned into the vector pCR 2.1 (Invitrogen, Carlsbad, CA, USA). These inserts were then cut at restriction enzymes designed to fall within the primer sequences, and the fragments obtained were gel-purified and recloned into pBluescript SK+ (Stratagene, La Jolla, CA, USA). These clones were sequenced to confirm their identity, and Northern and Southern analyses were performed to confirm the probes hybridized in expected patterns before use in *in situ* hybridization. RNA probes were made via *in vitro* transcription from the T3

or T7 promotors located within the vector using an *in vitro* transcription kit provided by Promega (Madison, WI, USA).

Northern blot analyses and In situ *hybridization*

Northern blots and hybridizations were performed according to the supplier's instructions using Hybond-N+ membrane by Amersham Biosciences (Piscataway, NJ, USA). *In situ* hybridization was carried out as in Geddy *et al.*, 2005.

RT-PCR

Total and mitochondrial RNA samples from both seedling tissue and floral buds were subjected to DNase I treatment (Invitrogen, Carlsbad, CA, USA), then checked for residual DNA by polymerase chain reaction (PCR). This procedure was repeated until no DNA contamination remained. Reverse-transcriptase polymerase chain reaction (RT-PCR) was performed using the One-step RT-PCR kit (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions with one alteration. A single primer was added to the mixture at the beginning of the reverse transcription reaction, corresponding to the amplification of either the sense or the antisense transcript. After the inactivation of the reverse-transcriptase, at the start of the PCR reaction, the second primer was added to permit amplification of only those double stranded products which had already been formed. The reaction contained 1xreaction mix, 1.8mM $MgSO_4$, 1µL RT/Taq DNA polymerase, 10mM of each primer and 1µg RNA. Reverse transcription was

performed at 45°C for 30 minutes with a 10 minute heat inactivation of reverse transcriptase at 94°C. PCR was carried out with 40 cycles of 94°C for 1 minute, 52°C for 1 minute and 72°C for 1 minute, followed by a final extension of 10 minutes at 72°C.

Results

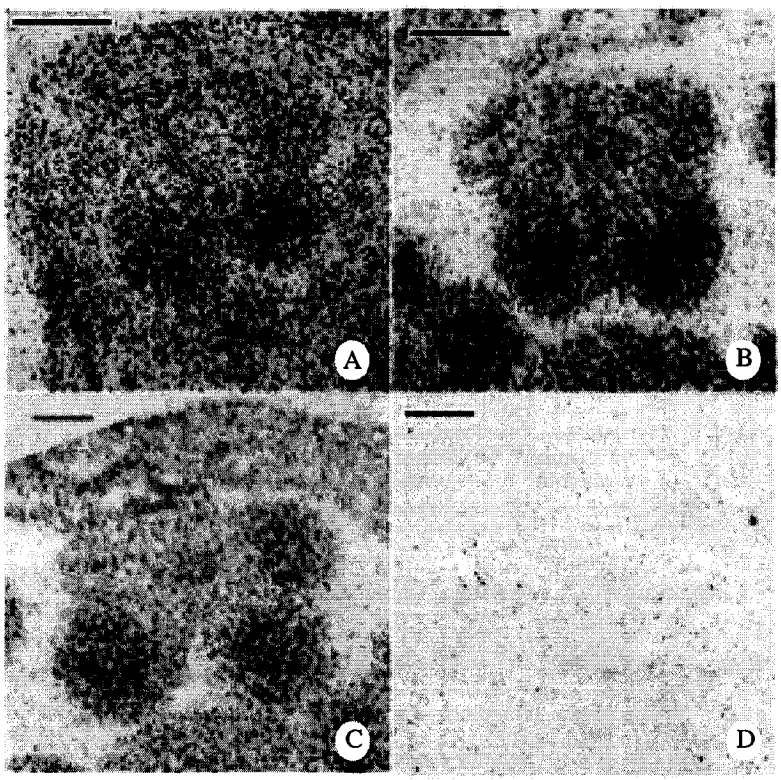
Sense and antisense transcripts of a mitochondrially-encoded subunit of the ATP synthase are expressed in a cell-specific manner during floral development

Probes composed of sense (coding) and antisense *in vitro* transcripts of the mitochondrially-encoded cytochrome *b* (*cob*) and *atp6* genes of *Brassica napus* were generated and used for *in situ* hybridization on sections of *nap* CMS and male fertile *B. napus* floral buds. The *nap* CMS is characterized by a failure of microsporogenic locules to develop in some, but usually not all four anther lobes (Geddy *et al.*, 2005). As shown in figure 3.1C, the *cob* antisense probe (which detects sense transcripts) hybridized most intensely to CMS anther sections at sites of locule development. A similar pattern of expression was observed for fertile flowers, except in this case hybridization in all four anther lobes, corresponding to sites of locule development occurred (not shown). This pattern of expression is consistent with that observed previously for transcripts encoding *cob* and other subunits of electron transfer chain subunits (Smart *et al.*, 1994, Geddy *et al.*, 2005). As expected, hybridization of the corresponding sense *cob* probe yielded only low, background levels of hybridization (figure 3.1D).

The *atp6* antisense probe displayed a pattern of hybridization similar to that of the *cob* antisense probe. (figure 3.1A). Unexpectedly, the *in vitro* transcript probe composed of the sense strand of the *atp6* gene hybridized efficiently to anther sections in a cell-specific pattern resembling that observed for the antisense probe of *atp6* (figure 3.1 B). The presence of *atp6* antisense transcripts was confirmed by

Figure 3.1 *In situ* hybridization of *Brassica napus nap* CMS floral bud tissue with *atp6* and cytochrome-*b* sequences in the microspore mother cell stage.

(A) *atp6* antisense probe.
(B) *atp6* sense probe.
(C) cytochrome *b* antisense probe.
(D) cytochrome *b* sense probe.
Size bars = 100μm.


A
B
C
D

RT-PCR (data not shown). The tissue specific pattern of expression for these probes, as well as the absence of specific sequence similarity between sense and antisense probes for this gene suggested that the observed hybridization of the sense probe was not an artifact but rather reflected the accumulation of antisense *atp6* transcripts in the same cell and tissue types as the sense transcripts of this gene.

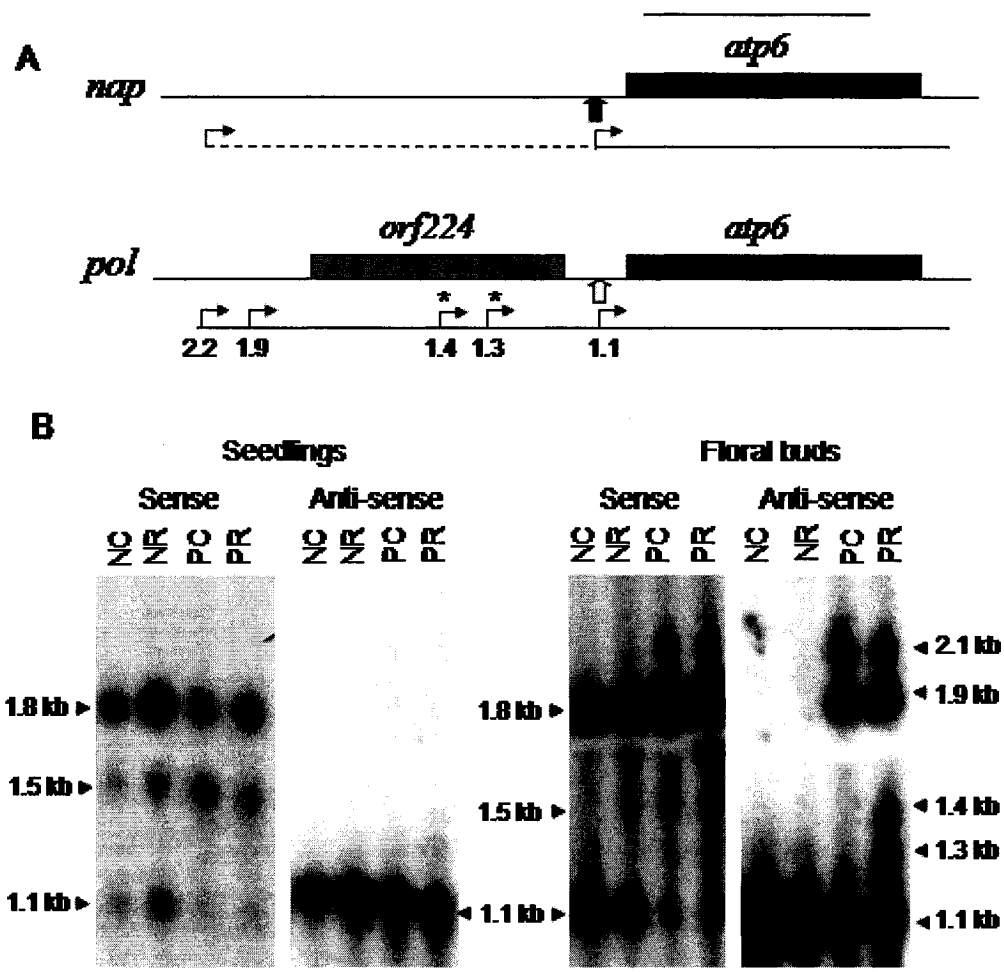
Antisense atp6 *transcripts are qualitatively different from the corresponding sense transcripts.*

RNA gel blot experiments were performed to further characterize the antisense RNA detected by *in situ* hybridization. In both CMS and fertility-restored plants with *nap* mtDNA, *atp6* transcription is initiated from a site approximately 1 kb upstream of the *atp6* initiation codon (Figure 3.2A; Menassa *et al.*, 1999). It has been proposed that the primary *nap atp6* transcript is processed through internal cleavage at a tRNA-like element located just upstream of the *atp6* coding sequence (filled block arrow Fig. 3.2A; Singh and Brown, 1991). This releases the stable, mature 1.1 kb *atp6* mRNA from the 3′ portion of the primary transcript; the 5′ non-coding portion of the primary transcript is then quickly degraded (Menassa *et al.*, 1999; Brown, 1999). In *pol* mtDNA, the tRNA-like element upstream of the *atp6* gene is partially disrupted (open block arrow, Fig. 3.2A; Singh and Brown, 1991) through a recombination event leading to the juxtaposition of *atp6* with the upstream, co-transcribed CMS-associated gene *orf224*. This results in a prevalence of dicistronic *orf224/atp6* transcripts of 2.2 and 1.9 kb in floral tissues of CMS plants. In addition, limited

Figure 3.2 Gel blot analysis of mitochondrial RNA with sense and antisense transcripts of the *atp6* gene.

A. Schematic depicting the organization and expression of sense *atp6* transcripts in *nap* and *pol* mitochondria of *B. napus*. Upper lines depict gene organization, lower lines depict transcripts. Filled boxes indicate the locations of *orf224* (gray) and *atp6* (black) coding sequences. The bar above the *atp6* coding sequence indicates the region used to prepare sense and antisense probes. Arrows beneath the upper lines indicate the position of tRNA pseudogene elements used as transcript processing sites. Arrows on the lower line indicate 5′ transcript termini. The dotted line depicts a highly unstable portion of the *nap atp6* transcript. Asterisks indicate transcript termini specific to fertility-restored plants.

B. Gel blot hybridization of seedling and bud mitochondrial RNA with sense and antisense *atp6* probes. NC, *nap* CMS plant; NR, fertility restored *nap* cytoplasm plant; PC, *pol* CMS plant; PR, fertility restored *pol* cytoplasm plant. Filled triangles indicate the estimated sizes of detected transcripts.


A
atp6
nap
orf224
atp6
pol
2.2
1.9
1.4
1.3
1.1
B
Seedlings
Floral buds
Sense
Anti-sense
Sense
Anti-sense
1.8 kb
1.5 kb
1.1 kb
1.1 kb
1.8 kb
1.5 kb
2.1 kb
1.9 kb
1.4 kb
1.3 kb
1.1 kb

cleavage at the disrupted tRNA-like element leads to the production of monocistronic 1.1 kb *atp6* transcripts. In nuclear fertility-restored plants, additional processing events within *orf224* lead to the production of restored-specific monocistronic *atp6* transcripts of 1.4 and 1.3 kb and a concomitant reduction in the levels of 2.2 and 1.9 kb *orf224/atp6* dicistronic transcripts (Singh and Brown, 1991; Menassa *et al.*, 1999).

The same *in vitro* sense and antisense *atp6* RNAs used for *in situ* hybridization were used to probe gel blots of mtRNA extracted from floral bud and seedling tissues of CMS and fertility-restored *nap* and *pol* cytoplasm *B. napus* plants. As expected, the antisense RNA probe detected a only a single 1.1 kb transcript in fertile and CMS plants with *nap* mtDNA; in buds of *pol* cytoplasm plants, the additional 2.2 and 1.9 kb *orf224/atp6* transcripts as well as the restored specific 1.4 and 1.3 kb *atp6* transcripts are detected (Figure 3.2B). In seedling tissue, again only a single, 1.1 kb nap *atp6* transcript was detected, and, consistent with earlier findings (Singh and Brown, 1993), the levels of *pol atp6* transcripts other than the 1.1 kb RNA were sharply reduced to the point where they were barely detectable.

A qualitatively different pattern of transcripts in both tissue types was detected by the sense *atp6* probe. In both floral buds and seedlings, the major antisense transcript in both *pol* and *nap* mitochondria is 1.8 kb long. Interestingly, mitochondria of all cytoplasmic types and tissues were also found to contain an antisense transcript indistinguishable in size from the 1.1 kb sense transcript; levels of this antisense transcript appeared markedly lower in *pol* than in *nap* mitochondria. A third antisense transcript of 1.5 kb was evident in seedling tissue

of all cytoplasmic types, regardless of fertility status. An antisense transcript of similar size could be detected in fertility-restored *pol* floral buds, and at somewhat lower levels in *pol* CMS and fertility restored *nap* floral buds. These results indicate that, in general, the antisense *atp6* transcripts differ from their sense counterparts both in their sizes and relative abundance in the different plant tissues analyzed. Like the sense transcripts, the *atp6* antisense transcripts differed among the different plant nuclear and cytoplasmic genotypes.

Discussion

Natural antisense transcripts (NATs) are found in all eukaryotes and have been extensively characterized in humans (Shendure and Church, 2002; Yelin *et al.*, 2003; Chen *et al.*, 2004), mice (Shendure and Church, 2002; Kiyosawa *et al.*, 2003), rice (Osato *et al.*, 2003) and *Arabidopsis* (Yamada *et al.*, 2003). It has been estimated that as many as 20% of genes in humans produce some form of antisense transcript (Chen *et al.*, 2004). In *Arabidopsis* this number may be as high as 30%, representing about 7600 annoted genes (Yamada *et al.*, 2003).

NATs can arise when both strands of a DNA sequence are transcribed into RNA (*cis*-NATs) or when single-stranded RNA is used as a template for the formation of a complementary transcript through the action of an RNA-dependent RNA polymerase (RdRP) (Baulcombe, 2004). In addition, antisense transcripts can be generated from a paralogous copy of a gene (*trans*-NATs). Many types of sense-antisense pairings have been categorized in genomic database searches (Yamada *et al.*, 2003; Osato *et al.*, 2003). Antisense transcripts often do not begin at the 3′ terminus of the sense transcript. Instead, the amount of overlap may be quite small, and can be interrupted by introns or other processing features. As a result, more than one antisense transcript may arise, depending on the available transcriptional start sites surrounding, or within, the gene in question. In *Arabidopsis*, a high proportion of NATs are derived from convergent overlapping gene pairs (COPs), overlapping genes in anti-parallel configurations, which give rise to convergent, overlapping transcripts (Jen *et al.*, 2005; Wang *et al.*, 2005).

It is not clear how the antisense transcripts of the ATP synthase subunit genes were generated. Transcription of a paralogous copy cannot account for the *atp6* antisense transcripts, since this gene occurs in only a single copy in the *Arabidopsis* mitochondrial genome (Singh and Brown, 1991; Handa *et al.*, 2003). It is noteworthy that the patterns of antisense *atp6* transcripts observed show some correlation with the patterns of sense transcripts. For example, in floral buds, levels of both the sense and antisense 1.1 kb *atp6* transcripts are lower in plants with *pol* cytoplasm than in plants with *nap* cytoplasm. This observation is consistent with the premise that the antisense transcripts may be derived from the sense transcripts through the action of an RNA-dependent RNA polymerase. In this respect, we note that Marienfeld *et al.* (1997) found open reading frames in the *Arabidopsis* and *Vicia faba* mitochondrial genomes that could potentially specify RdRPs and that one of these *Arabidopsis* open reading frames (ORF251) is virtually completely conserved over most of its length in the *Brassica* mitochondrial genome (personal observations). This suggests that plant mitochondria may possess functional RdRPs that generate antisense transcripts such as those described. In addition, because only a limited number of promoters may be employed to transcribe plant mitochondrial genomes, there is extensive transcription on non-coding DNA in plant mitochondria, and thus a high potential for the convergent transcription of complementary strands of coding and non-coding sequences.

In general, antisense transcripts regulate the corresponding sense transcripts, either by inducing their degradation (RNA interference or RNAi) or by inducing post-transcriptional gene silencing. Double stranded RNA arising, for example,

from inverted repeat structures or the action of an RNA-dependent RNA polymerase, can trigger the action of a Dicer-like protein which cleaves the double-stranded RNA into short segments. These short RNAs can lead to chromatin remodeling, RNA cleavage or translational repression, all of which can lead to gene silencing (Meister and Tuschl, 2004). Other suggested post-transcriptional regulatory mechanisms for antisense transcripts include the enhancement of RNA stability, and the control of mRNA translation (Kiyosawa *et al.*, 2003; Terryn and Rouze, 2000).

Patterns of accumulation of antisense transcripts of *atp6* closely match that of their sense transcripts in *in situ* hybridization experiments. This may indicate that the antisense transcripts are generated in a tissue-specific manner and may interact with the sense transcripts as a type of post-transcriptional regulatory mechanism. This situation appears similar to that of the maize *Bz2* gene, where antisense transcripts are observed only in tissues expressing high levels of sense transcripts (Schmitz and Theres, 1992).

References

Akagi, H., Sakamoto, M., Shinjyo, C., Shimada, H. and Fujimura, T. (1994) A unique sequence located downstream from the rice mitochondrial *atp6* gene may cause male sterility. Current Genet. **25**, 52-58.

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

Chen, J., Sun, M., Kent, W.J., Huang, X., Xie, H., Wang, W., Zhou, G., Shi, R.Z. and Rowley, J.D. (2004) Over 20% of human transcripts might form sense-antisense pairs. Nucleic Acids Res. **32**, 4812-4820.

Brown, G.G. (1999) Unique aspects of cytoplasmic male sterility and fertility restoration in *Brassica napus*. J. Heredity **90**, 351-356.

Geddy, R., Mahe, L. and Brown, G.G. (2005) Cell-specific regulation of a *Brassica napus* CMS-associated gene by a nuclear restorer with related effects on a floral homeotic gene promoter. Plant J. **41**, 333-345.

Handa, H. (2003) The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis* thaliana. Nucleic Acids Research **31**, 5907-5916.

Hegeman, C.E., Halter, C.P., Owens, T.G. and Hanson, M.R. (2005) Expression of complementary RNA from chloroplast transgenes affects editing efficiency of transgene and endogenous chloroplast transcripts. Nucleic Acids Res. **33**, 1454-1464.

Jen, C.-H., Michalopoulos, I., Westhead, D.R. and Meyer, P. (2005) Natural antisense transcripts with coding capacity in *Arabidopsis* may have a regulatory role that is not linked to double-stranded RNA degredation. Genome Biol. **6**, R51.

Kemble, R.J. and Barsby, T.L. (1988) Use of protoplast fusion systems to study organelle genetics in a commercially important crop. Biochem. Cell. Biol. **66**, 665-676.

Kiyosawa, H., Yamanaka, I., Osato, N., Kondo, S., RIKEN GER Group 1 and GSL Members and Hayashizaki, Y. (2003) Antisense Transcripts With FANTOM2 Clone Set and Their Implications for Gene Regulation. Genome Res. **13**, 1324–1334.

Lang, B.F., Gray, M.W. and Burger, G. (1999) Mitochondrial genome evolution. Ann. Rev. Genet. **33**, 351-397.

Lavorgna, G., Dahary, D., Lehner, B., Sorek, R., Sanderson, C.M. and Casari, G. (2004) In search of antisense. Trends in Biochem. Sci. **29**, 88-94.

L'Homme, Y., Stahl, R.J., Li, X.-Q., Hameed, A. and Brown, G.G. (1997) Brassica *nap* cytoplasmic male sterility is associated with expression of a mtDNA region containing a chimeric gene similar to the *pol* CMS-associated *orf224* gene. Curr. Genet. **31**, 325-335.

Marienfeld, J.R., Unseld, M., Brandt, P. and Brennicke, A. (1997) Viral nucleic acid transfer between fungi and plants. Trends in Genet. 13, 260-261.

Menassa, R., L'Homme, Y. and Brown, G.G. (1999) Post-transcriptional and developmental regulation of a CMS-associated mitochondrial gene region by a nuclear restorer gene. Plant J. **17**, 491-499.

Meister, G. and Tuschl, T. (2004) Mechanisms of gene silencing by double-stranded RNA. Nature **431**, 343-349.

Nishimura, Y., Kikis, E.A., Zimmer, S.L., Komine, Y. and Stern, D.B. (2004) Antisense transcript and RNA processing alterations suppress instability of polyadenylated mRNA in chlamydomonas chloroplasts. Plant Cell **16**, 2849-2869.

Osato, N., Yamada, H., Satoh, K., Ooka, H., Yamamoto, M., Suzuki, K., Kawai, J., Carninci, P., Ohtomo, Y., Murakami, K., Matsubara, K., Kikuchi, S. and Hayashizaki, Y. (2003) Antisense transcripts with rice full-length cDNAs. Genome Biol. **5**, R5.1-R5.10.

Schmitz, G. and Theres, K. (1992) Structural and functional analysis of the Bz2 locus of *Zea mays*: characterization of overlapping transcripts. Mol. Gen. Genet. **233**, 269-277.

Shendure, J. and Church, G.M. (2002) Computational discovery of sense-antisense transcription in the human and mouse genomes. Genome Biol. **3**, 1-14.

Singh, M. and Brown, G.G. (1991) Suppression of cytoplasmic male sterility by nuclear genes alters expression of a novel mitochondrial gene region. Plant Cell **3**, 1349-1362.

Singh, M., and Brown, G.G. (1993) Characterization of expression of a mitochondrial gene region associated with the *Brassica* "Polima" CMS: developmental influences. Curr Genet. **24**, 316-322.

Smart, C.J., Moneger, F. and Leaver, C.J. (1994) Cell-specific regulation of gene expression in mitochondria during anther development in sunflower. Plant Cell **6**, 811-825.

Terryn, N. and Rouze, P. (2000) The sense of naturally transcribed antisense RNAs in plants. Trends in Plant Sci. **5**, 394-396.

Wagner E.G. and Flardh, K. (2002) Antisense RNAs everywhere? Trends in Genet. **18,** 223-226.

Wang, X-J., Gaasterland, T. And Chua, N.-H. (2005) Genome-wide prediction and identification of cis-natural antisense transcripts in *Arabidopsis* thaliana. Genome Biol. **6,** R30.

Wassenegger, M., and Pelissier, T. (1998) A model for RNA-mediated gene silencing in higher plants. Plant Molecular Biology **37**, 349-362.

Yamada, K., Lim, J., Dale, J.M., Chen, H., Shinn, P., Palm, C.J., Southwick, A.M., Wu, H.C., Kim, C., Nguyen, M., Pham, P., Cheuk, R., Karlin-Newmann, G., Liu, S.X., Lam, B., Sakano, H., Wu, T., Yu, G., Miranda, M., Quach, H.L., Tripp, M., Chang, C.H., Lee, J.M., Toriumi, M., Chan, M.M., Tang, C.C., Onodera, C.S., Deng, J.M., Akiyama, K., Ansari, Y., Arakawa, T., Banh, J., Banno, F., Bowser, L., Brooks, S., Carninci, P., Chao, Q., Choy, N., Enju, A., Goldsmith, A.D., Gurjal, M., Hansen, N.F., Hayashizaki, Y., Johnson-Hopson, C., Hsuan, V.W., Iida, K., Karnes, M., Khan, S., Koesema, E., Ishida, J., Jiang, P.X., Jones, T., Kawai, J., Kamiya, A., Meyers, C., Nakajima, M., Narusaka, M., Seki, M., Sakurai, T., Satou, M., Tamse, R., Vaysberg, M., Wallender, E.K., Wong, C., Yamamura, Y., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R. (2003) Empirical analysis of transcriptional activity in the *Arabidopsis* genome. Science **302**, 842-846.

Yelin, R., Dahary, D., Sorek, R., Levanon, E.Y., Goldstein, O., Shoshan, A., Diber, A., Biton, S., Tamir, Y., Khosravi, R., Nemzer, S., Pinner, E., Walach, S., Bernstein, J., Savitsky, K. and Rotman G. (2003) Widespread occurrence of antisense transcription in the human genome. Nature Biotech. **21**, 379-386.

CHAPTER 4

TOWARDS IDENTIFICATION OF THE RESTORER OF FERTILITY FOR POLIMA CMS

Following studies on the expression patterns of CMS inducing genes and the effects of the cytoplasmic male sterile condition on *Brassica napus* flower buds, the next piece of the puzzle to be investigated was the nature and effect of the restorers of fertility. Studies that had already begun to locate *Rfp*, the restorer of Polima CMS, were continued to map and characterize that gene. Strategies used and results obtained in this mapping project are presented in this chapter.

Initial experimental work, presented in the results section under the heading "Background work", was performed by Natasa Formanova and staff at DNA Landmarks (St-Jean-sur-Richelieu). Experiments performed during the course of this thesis are included in the results section under the heading "Current Work", and was assisted by Céline Frechina, Nadim Maghzal and Ryan Nantel-Smith.

Introduction

The cytoplasmic male sterile phenotype in plants is reversed by the action of nuclearly encoded restorer genes that down-regulate the expression of specific mitochondrial transcripts. In *Brassica napus*, the restorer of *nap* fertility, *Rfn*, both cleaves *orf222/nad5c/orf139* co-transcripts and reduces their abundance. *Rfp*, the *pol* restorer mediates cleavage of transcripts of the *atp6* gene region, reducing the size of sterility inducing *orf224* transcripts (Li *et al.*, 1998).

Studies of *Rfn* and *Rfp* have shown that the two restorers are very closely linked and may, in fact, be alternate "alleles" or haplotypes of a single nuclear locus (Li *et al.*, 1998). Thus, in identifying the *pol* restorer of fertility, the discovery of the corresponding *nap* restorer should be a relatively simpler task. The similarity of *orf222* and *orf224*, and the close proximity of the two restorers leads to the hypothesis that *Rfn* and *Rfp* are likely to be similar to each other. Therefore, subtle differences in sequence and/or structure which may be found between *Rfn* and *Rfp* may yield clues to the mechanism of target recognition and the specificity of action which allows the two restorers to distinguish between two sterility inducing transcripts that share 85% nucleotide sequence homology.

Once *Rfp* and *Rfn* are sequenced, this may also allow us to examine how Rf systems evolve and adapt to new sterility inducing genes and the molecular mechanisms governing the generation of new restorer genes.

Many technical challenges are associated with map-based cloning approaches in *Brassica*. The *Brassicaceae* are closely related to the model organism *Arabidopsis thaliana* (Meyerowitz and Pruitt, 1985). As a result of an ancient whole genome duplication, many *Arabidopsis* genes are present in two or more diverged, paralogous copies (Arabidopsis Genome Initiative, 2000). Through additional genomic duplications that have occurred since the divergence of the Brassicaceae from the *Arabidopsis* lineage, modern-day haploid *Brassica* species contain about three orthologous loci for every *Arabidopsis* locus (Lagercrantz *et al.*, 1996; Scheffler *et al.*, 1997; Cavell *et al.*, 1998). Some evidence suggests that triplication of the complete genome occurred (Lagercrantz and Lydiate, 1996; Scheffler *et al.*, 1997; Cavell *et al.*; 1998, Parkin *et al.*, 2002) while other evidence favors a more complex explanation that may include aneuploidy and chromosomal rearrangement (Quiros, 2001; Li *et al.*, 2003). It is generally agreed, however, that the genomic expansion seen in *Brassica* is a result of duplications, and not due to the accumulation of intergenic DNA sequences (Jackson *et al.*, 2000).

The presence of multiple paralogous sequences for most genes within *Brassica* complicates gene cloning and identification approaches. For map-based cloning approaches, a means of circumventing such difficulties is by ensuring that clones that are used to initiate chromosome walks contain a genetic polymorphism that is tightly genetically linked to the gene of interest. The fully sequenced genome of *Arabidopsis* is a valuable tool for the discovery of linked polymorphic markers in *Brassica* where synteny between the two species is maintained (Schmidt *et al.*, 2001). However, genome rearrangements and sequence divergence in *Brassica* can limit its utility (O'Neill and Bancroft, 2000; Ryder *et al.*, 2001).

The radish restorer gene, *Rfo*, was cloned with the aid of synteny between radish and *Arabidopsis* (Brown *et al.*, 2003). The sequenced genome of *Arabidopsis* provided a useful set of ordered genes from which to base chromosome walking once the region of interest was anchored by a marker (Brown *et al.*, 2003). We have embarked on the use of this approach to clone *Rfp*, the restorer of *pol* CMS of *Brassica napus*. The work presented describes the significant progress that has been made towards refining the genomic interval in which *Rfp* is located and the initial steps made towards cloning this gene.

Materials and Methods

Primer design and product amplification

Primers were designed according to sequences in the *Arabidopsis* nuclear genome (Arabidopsis Genome Initiative, 2000) corresponding to loci in the region of *Rfp*, using the online primer design software Primer3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). These primers were used to amplify the corresponding sequence from *Brassica napus* total DNA using the polymerase chain reaction (PCR) with annealing temperatures varying depending on the degree of homology between primers and their corresponding *Brassica* sequence. Primer pairs chosen for cosmid selection showed a clear hybridization pattern on Southern blots of bulk DNA from fertile and sterile plants (see below). Primer pair sequences are shown below in 5′-3′ orientation.

Primers designed during Current Work	Primers from Background Work
2650 F:TGTCGACATGAAAGGAGCTG R:TTTGTTAGCACGAGCACGAC	P2 F:TTTCAGATTATATGTTTTCAGATTCG R:AGTTAAGGGGTGGGGTTTTG
2740 F:GAGAACCCTCGGAAGAATCC R:TTCCCGGTTTCTCAGAATTG	3 F:GCAACGTGACAACCCGAGGAT R:CCTCCTTGTGGCGCGTCTTC
2770 = PPR3-EX2-1 F:AACCGAATGAGGCAAGAAGA R:GAGAATGGCACAGTTGCAGA	6 F:GAAGTTCGGGTCGGTTAGGAT R:GCACGCTGCTCAGAGTACCTC
2790 F:CCCCCTGATCTCATCTTCAA R:CACTCCACGAGAGTGAGCAA	9 F:TGATAAGGAGCATTCGACGAT R:CACCAGCAGTGTACGCAAGA
2820 F:TCGAATCATCAGCGAGTTTG R:TGGTCCAGTGAGATCCACAA	12 F:AAACCGTATGTCCGCTCTCGT R:CGACAAATGGGTCTGCGAAGT
2840 F:TCTAAAAGACTCCGGCCAAA R:TACTGGAAAAGCGAGGAGGA	15 F: CTGAACCGCGCTCTTAGCAAG R: TGACGAGATTTGCGATGGGTC

Southern blot analysis

Agarose gel electrophoresis and gel blotting were performed as described in Chapter 1. Southern hybridization was performed using digoxygenin labelled probes that were created by PCR amplification on unlabelled PCR products (Roche: Laval, Quebec, Canada) using the primers designed as described above. Southern hybridization and all wash steps were performed in hybridization bottles in a rotisserie oven according to the product instructions. One change to the protocol was to make the High Stringency Wash Buffer with 0.2x SSC and 0.1% SDS. Cosmid blots were exposed to chemiluminescent detection film (Kodak: Rochester, New York, USA) for 5-30 minutes, while genomic blot hybridizations were exposed to film overnight.

Colony lift procedure

The subcloned genomic library in pOCA18 (Formanov *et al.*, 2006) was plated on large (150mm) plates containing Luria Bertani agar medium with 25µg/mL ampicillin. The cells were plated directly onto a nylon membrane laid on the top of the plate medium at a density of approximately 10000 cfu/plate and grown overnight at 37°C. After this growth phase the nylon membranes were removed from the plates and replicated by the following procedure:

Two layers of wet Whatman paper (VWR: Mississauga, ON, Canada) are laid on a piece of aluminum foil. The membrane containing the colonies is laid on this stack face up, then a second membrane is placed over it and centered. Two more layers of wet Whatman paper are placed on top of this and the stack is rolled

over with a disposable plastic pipette several times. Next, a glass plate is placed over the stack and hard pressure is applied downward for a few seconds. The stack is disassembled to the membrane layers, then a unique pattern of pinholes is poked through the double membrane layers for later identification and alignment. The two membrane layers are carefully and slowly separated, marked on the edge in pencil and each is placed face up on a separate LB/ampicillin plate. The original plate is placed at 4°C, the duplicate is re-grown at 37°C for 4 hours. The duplicate membrane is then lysed and prepared for hybridization as follows. The membrane is placed face up for two minutes on 5 mL of 0.2M NaOH then dried briefly on fresh dry Whatman paper. This step is repeated once. Then, the membrane is placed face up for two minutes on 5 mL of 1M tris pH 7.5 then dried briefly on fresh dry Whatman paper. This step is repeated once. The DNA is then cross-linked to the membranes with ultraviolet light. Labeled probes are hybridized to membranes by the same procedure as that used for Southern blots.

Colony purification

Bacterial colonies were purified from nylon membranes in several rounds according to the following procedure:

In the first round, the hybridized membrane is lined up with the corresponding exposed and developed film and reference pinholes are transferred to the film. The film is reversed and the original membrane containing live colonies is laid over the film, again lining up reference pinholes; water is first sprayed between the film and the membrane to prevent membrane drying. Sections of membrane

corresponding to hybridizing colonies are cut out and suspended in LB medium. This is then diluted and re-plated on membranes at a much lower density for a second round of hybridization, as described above. After this second round, if separate single hybridizing colonies can be resolved, they are picked, grown and cosmid DNA is prepared. If single colonies cannot be picked, a third round of purification may be performed as described above. Mini-preps of cosmid DNA are digested, run on an agarose gel, and blotted onto membranes which are then hybridized with the associated probe to confirm that the cosmid contains a hybridizing genomic fragment. Glycerol stocks of each positive single colony are stored at -80°C.

Amplified Fragment Length Polymorphism (AFLP) analysis

Adapters were synthesized from primers designed to self-anneal, and which each contained a single restriction site for either *Mse*I, *Eco*RI or *Hin*dIII. Two infrared (IR) detectable primers were designed to amplify from either the *Eco*RI or *Hin*dIII adapters, and one unlabelled *Mse*I primer was also synthesized. Qiagen purified minipreps of cosmid DNA were digested with 5U of *Mse*I and 5U of either *Eco*RI or *Hin*dIII (enzymes from New England Biolabs: Ipswich, MA, USA) to completion in the appropriate 1x buffer provided by the manufacturer (in this case NEB2), plus RNase and BSA in a total volume of 10µL overnight at 37°C. Enzymes were heat inactivated at 65°C for 20 minutes and adapters designed against (*Mse*I and one of *Eco*RI or *Hin*dIII) were then ligated overnight at room temperature. Polymerase chain reactions were then carried out using the unlabeled *Mse*I primer and one of the labeled primers. These PCR products were

then diluted, denatured and run on a Li-COR sequencing gel and bands were detected by infrared light by the Li-COR laser.

Shearing and subcloning of cosmid DNA

4 µg of Qiagen purified midiprep DNA in a volume of 300 µL was sheared down to 5-8 kb using a hydroshear (GeneMachines: Ann Arbor, MI, USA). The resulting DNA sample was subcloned into the pMOSblue cloning vector (Amersham Biosciences Corp.: Piscataway, NJ, USA) using the pMOSblue blunt ended cloning kit according to the manufacturer's instructions. Ligated DNA was then transformed into Top10 chemically competent cells (Invitrogen: Burlington, ON, Canada) using blue-white selection with LBamp 100µg/mL and X-gal according to the manufacturer's instructions.

Sequencing and sequence assembly

Sequencing and sequence assembly was performed by DNALandmarks (St-Jean-sur-Richelieu) (Brown *et al.*, 2003) and Genome Quebec (Montreal) using the Applied Biosystems 3730XL DNA analyzer for capillary sequencing and the Phred/Phrap programs for sequence assembly (Ewing and Green 1998).

Results

Background work

Extensive work has been done to isolate the restorer of fertility for Polima CMS, *Rfp*. The *Rfp* gene was mapped to linkage group 18 in *Brassica napus* using restriction fragment length polymorphism (RFLP) markers (Jean *et al.*, 1997). Subsequently, the *Rfp* gene and the *pol* cytoplasm were introduced from *B. napus* into *B. rapa* by means of an inter-specific cross and subsequent backcross with *B. rapa*, and a doubled-haploid *pol* cytoplasm *B. rapa* population segregating for the *Rfp* gene was generated (Formanov *et al.*, 2006). The introgressed *Rfp* gene was shown to map to an orthologous chromosomal site in *B. rapa*, and a cosmid library from an individual homozygous for *Rfp* was generated (Formanov *et al.*, 2006).

The comparatively simpler genome of *B. rapa* reduces the complexity of the DNA set from which to select cosmids. However, there still exist multiple copies or paralogs of most genes due to extensive internal genome duplications, such that for each probe used more than one region will be selected. Thus, selected cosmids must be anchored in the *Rfp* region by means of a genetically linked marker to confirm their position with reference to the restorer gene.

Comparison of the sequences of RFLP probes mapping to the region in which *Rfp* is localized with that of the *Arabidopsis* genome indicated that in the *Rfp* region the *Brassica* genome is generally syntenic to the closely related crucifer

Arabidopsis, whose genome has been fully sequenced. Synteny was further established using probes from the *Arabidopsis* region homologous to *Brassica*. These probes detected RFLPs which mapped close to *Rfp* and the order of these mapped probes was conserved between the two genomes. This synteny has been exploited in the design and use of probes to extract contiguous cosmids from the library.

Further mapping of the *Rfp* region in *Brassica napus* was conducted by initially screening a large *pol* CMS x Westar Rf BC1 backcross population with *B. napus* PCR-based markers that flank *Rfp* to identify individuals in which recombination events near the *Rfp* gene had occurred. A fine map of this region was then constructed by probing individuals possessing chromosomes in which recombination close to *Rfp* has occurred with *Arabidopsis*-derived RFLP markers. In this manner, the *Rfp* gene was found to map in *Brassica* between RFLP markers designated 9.6BB and 2.7HH (Figure 4.1), which are paralogous to sequences located on the left arm of chromosome 1 of *Arabidopsis* where they are separated by at least 123 kb. In *Arabidopsis* there are three PPR genes located within this region.

Current Work

Another polymorphism was detected with the primer set called P2, designed during "Background work", which amplifies a PPR-encoding gene homologous to the *Arabidopsis* locus At1g12775. Using these primers to amplify DNA from fertile plants yields two PCR bands while DNA from sterile plants only yields

Figure 4.1 Cartoon depicting the placement of probes with reference to the syntenic region of *Arabidopsis*.

Probe names are enclosed within arrows; italicised probe names indicate probes designed prior to the current study; probe names in bold indicate primers designed during the course of this study. Four-number probe names are indicative of the last four digits of the *Arabidopsis* gene locus from which they are derived. For example "2650" corresponds to At1g12650. RFLP markers 2.7HH and 9.6BB enclose the region. Position of sequenced cosmids are indicated by heavy lines with the cosmid name underneath and are positioned with respect to those probes to which they hybridize; those indicated with an asterix (*) are not yet anchored in the region. Scale given (**), as well as probe placement, is based on *Arabidopsis* syntenic region. Note that cosmid inserts are not to scale with respect to *Arabidopsis*, but are all approximately 20-25kb in length.

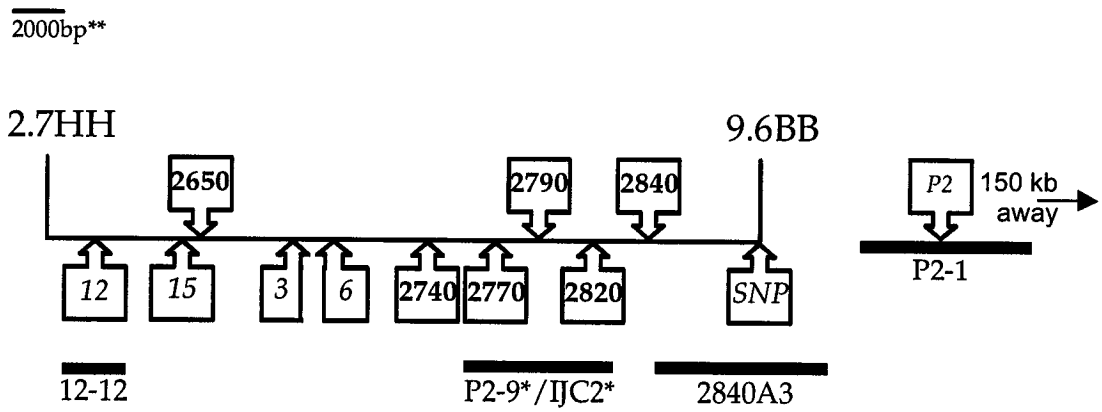
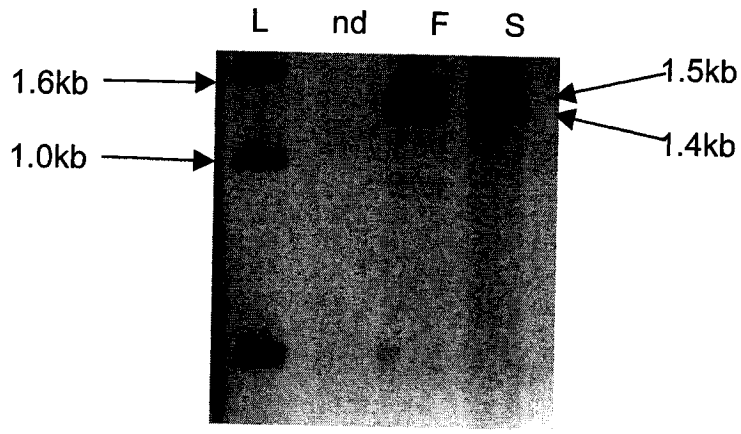

2000bp**
2.7HH
9.6BB
2650
2790
2840
P2
150 kb
away
12
15
3
6
2740
2770
2820
SNP
P2-1
12-12
P2-9*/IJC2*
2840A3

Figure 4.2 Polymorphism P2 of the *Rfp* region used to identify linked cosmids from a *Brassica rapa* library introgressed with *Rfp*

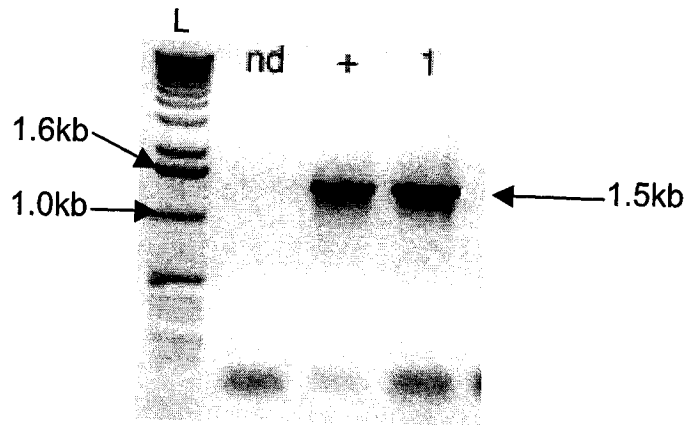
(a) Polymorphism of the marker P2 is shown by polymerase chain amplification using P2 primers. L=DNA ladder, nd = no DNA control, F = bulk DNA from fertile plants, S = bulk DNA from Sterile plants

(b) P2 amplification of the positive cosmid selected with the P2 probe L= DNA ladder, nd = no DNA control, + = reamplification of probe DNA, 1 = amplification of P2 sequence from cosmid P2-1.

(a)



(b)



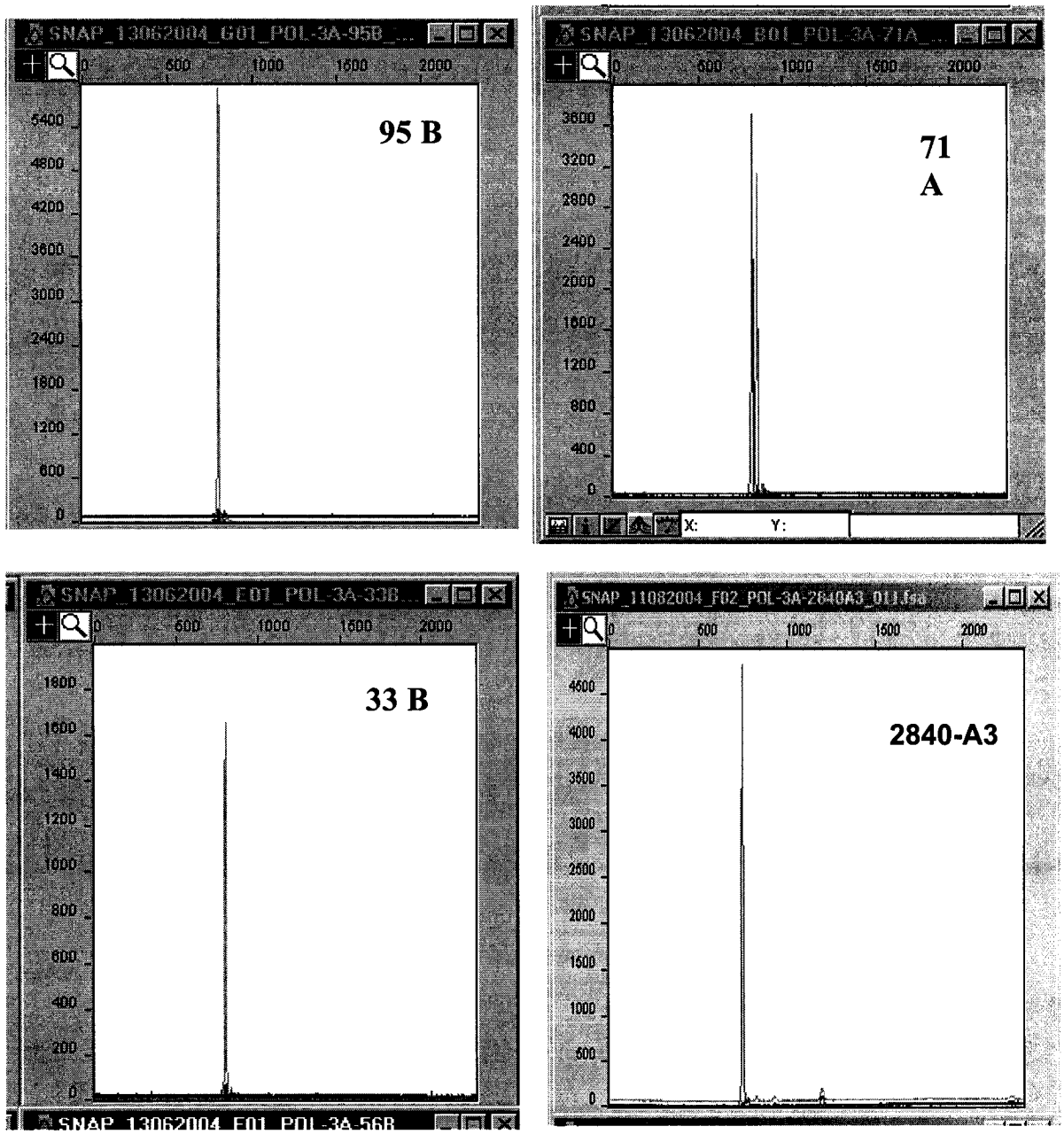
one band (Figure 4.2a). The sequence unique to fertile plants was used as a probe, but when the resultant clone, P2-1, (Figure 4.2b) was sequenced it showed that the P2 PPR-encoding gene is located outside the marker-delimited region, about 150kb to the right.

The region in question was further narrowed to a region syntenic to 105 kb of chromosome 1 of *Arabidopsis* by the use of a single nucleotide polymorphism (SNP) marker. Among recombinant individuals of the *pol* CMS x Westar Rf BC1 population from "background work", only one individual "33B" had recombination between the SNP and *Rfp* (Figure 4.3). This plant was fertile but showed a homozygous sterile SNP polymorphic genotype indicating that recombination had occurred to segregate the fertile SNP allele away from the *Rfp* allele. Several clones were selected using the amplicon containing the SNP, and from these, clones anchored in the *Rfp* region were identified as possessing the *Rfp*-linked polymorphism (Figure 4.3a). One such cosmid, 2840A3, contained the fertility associated SNP nucleotide and was completely sequenced. It contains genes from the syntenic *Arabidopsis* region between At1g12880 and At1g12950. In addition, it contains sequences homologous to At1g12840 and to the probe 2840, which have translocated out of their syntenic location in *Arabidopsis*. No PPR-encoding genes were found within the 2840A3 sequences. The 105 kb *Arabidopsis* syntenic region between the left-hand extremity of this cosmid and the marker 2.7HH still encloses all three *Arabidopsis* orthologous PPR genes mentioned above.

Figure 4.3 Single Nucleotide Polymorphism of the *Rfp* region used to identify linked cosmids from a *Brassica rapa* library introgressed with *Rfp*

(a) Images of fluorescence signal expressed by the 3′ primer annealing to the location of the SNP marker. The green signal is generated by a T/A nucleotide (sterile SNP allele), while the blue signal is indicative of the C/G nucleotide characterizing the fertile SNP allele. 95B is a sterile individual showing the characteristic homozygous green signal of the sterile SNP polymorphism. 71A is a fertile individual showing both the blue and green signals of a heterozygous SNP locus. 33B is a fertile individual showing the sterile SNP genotype, indicating that a recombination event occurred between the SNP and *Rfp*. 2840A3 shows the characteristic fertile SNP allele. (b) Sequence of the SNP polymorphism of 2840A3. The polymorphic nucleotide (C) is highlighted in green.

(a)



(b)

AATGAGAAGCTCTGAGCTGATACTCCGTTAGTGCGTCTCGGTTGGAACAA
GCAGGATCCGAGGTACATGGCGACTATCATCATGGACAGTGCTAAGGTT
GTAGTGCTTGACATTCGTTTTCCGGCGCTTCCTGTGGTGGAGCTTCAGAGG
CATCAGGCTAGTGTTAATGCTATCGCTTGGGCTCCTCATAGCTCGTCTCAT
ATCTGTACCGCGGGGGATGATTCTCAGGCTTTGATTTGGGATATTTCGTCG
ATGGGGCAGCAGGTTGAAGGTGGGCTTGATCCGATTCTTGCGTACACTGC
TGGT

Additional probes were designed using the sequence of the *Arabidopsis* genome between the primer pair corresponding to At1g12840 and the RFLP marker 2.7HH. Primers that were able to successfully amplify products from *Brassica* DNA were then selected for further analysis. Those PCR products were made into probes and tested on *Brassica* genomic Southern blots for specificity. Probes displaying specificity were sequenced and used to probe the *Brassica rapa* library containing *Rfp*. Several sets of colony lifts were performed, and each probe was tested on at least 300,000 colonies representing at least 10-fold coverage.

From the left-hand side of the region, we were unable to produce a probe corresponding to the RFLP marker 2.7HH with which to anchor the region. However, using the next closest probe "12", located to the right of 2.7HH (Figure 4.1), we recovered a cosmid which displayed a polymorphic *Hind*III fragment characteristic of fertile DNA (Figure 4.4). The sequence of this cosmid revealed only two small islands of homology to *Arabidopsis*, one of which corresponded precisely to the probe sequence.

Difficulties encountered in our attempts to reliably anchor the region to the 2.7HH RFLP marker led to a refocusing of efforts in other directions. To this end, the probing of the library was continued with an emphasis on the right-hand side of the region, anchored by the SNP marker. More than 120 independently selected cosmids were selected from the library using 12 different probes, 72 of which were chosen for further analysis based on restriction analysis and hybridization patterns (Figure 4.1, Table 4.1). The insert size of this cosmid library averages between 20 and 25 kb. It was found that restriction analysis was

Figure 4.4 Restriction Fragment Length Polymorphism (RFLP) of the *Rfp* region used to identify linked cosmids from a *Brassica rapa* library introgressed with *Rfp*

(a) RFLP seen between fertile (F) and sterile (S) genomic DNA digested with *Eco*RI and hybridized with the probe "12"
(b) the same band seen with *Eco*RI digest of positive clone 12-12; sequencing confirmed that the sequences corresponding to the probe "12" were found between two *Eco*RI sites 6000 bp apart.

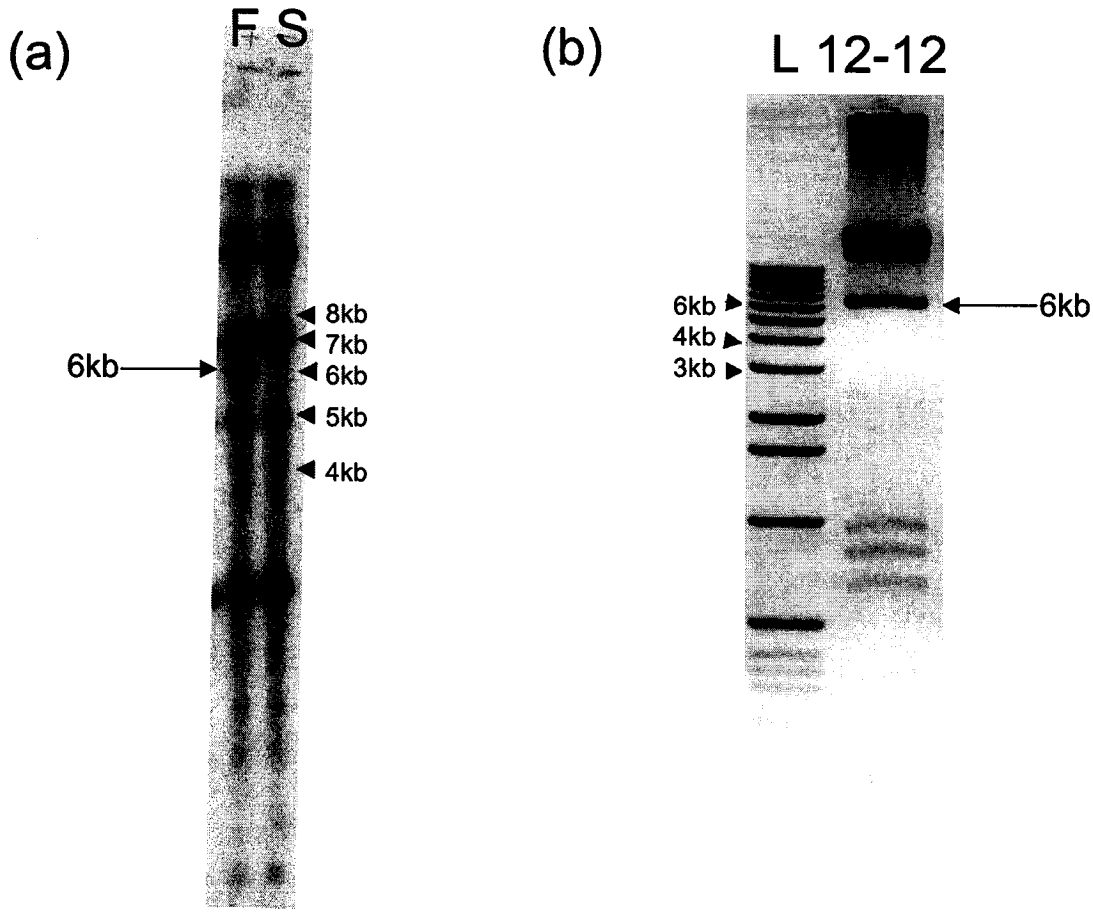

(a)
F S
8kb
7kb
6kb
5kb
4kb
6kb
(b)
L 12-12
6kb
4kb
3kb
6kb

Table 4.1: Cosmid clones recovered with their respective probes from the *B. rapa* library introgressed with *Rfp* and analysed by the Amplified Fragment Length Polymorphism (AFLP) technique.

Probe	Clones
12	12B, 12D, 12D1, 12E1, 12E2, 12-4, 12-5, 12-12, 12-14 (9 cosmids)
15	2630-6, 2630-1B/D (2 cosmids)
2650	None selected for AFLP analysis
3	3E1, 3E2, 3D1 (3 cosmids)
P2	IJC2, P2-9, P2-12, P2-16, IJD1, IJC1 (6 cosmids)
6	2710-4B, -5B, -51A, 6D30, 3/6-15-2, 3/6-13-4, 3/6-5-3, lib6C1.1, lib6A2.5 (9 cosmids)
2740	4w/4-1, -A, -B, -D, -G, -L, -N, -R, -S, 2740-3A, 2740B1, 2740-8A*, -9A, -11C, -20C, -(1)4, -(1)11, -(2)18 (18 cosmids)
2770	3-2-1E, 3-2-1A (2 cosmids)
2790	4w/9-1, -4, -7, -11, -12, -14, 2790-6A, -10, -14, -15, -16, -20, -24, -26, -33, -37. -39, 2630-1A (18 cosmids)
2820	2820-7, 2820-2A, 2820-7A, 2820-1, 2820-12A (5 cosmids)
2840	2840A3, 2840C2 (2 cosmids)

insufficient on such small fragments to unambiguously assign different selected cosmids to specific contigs.

Amplified fragment length polymorphism (AFLP) was performed as an alternative method to line up contiguous DNA sequences of positive clones. The AFLP technique involved the digestion of DNA with *Mse*I (a four-base cutter) and either *Eco*RI or *Hind*III (both six-base cutters). While *Mse*I cuts frequently, statistically every 256 bp, the six-base cutters will cut statistically only every 4096 bp. This means that in a 20kb cosmid there are about 10 *Mse*I-*Eco*RI and 10 *Mse*I-*Hind*III fragments. PCR carried out on these fragments with a labeled probe homologous to the six-base cutter end of the fragment can be resolved and visualized on an acrylamide gel with a resolution down to 1 base pair. Changes within the genome of *Brassica* can include both base substitution and base insertion or deletion, thus potentially introducing variability in length of fragments and/or sequence which could create or destroy enzyme cut sites. Therefore, if four fragments representing two six-base cutting sites and four four-base cutting sites can be found in common between adjacent cosmids, they can be said to overlap and form a contig. However, when AFLP analysis was performed on the cosmids isolated none of the contigs formed could be conclusively shown to overlap with the anchoring cosmid, 2840A3. A sample AFLP gel is shown in Figure 4.5. Cosmids grouped into contigs are listed in Table 4.2.

Two contigs, 1 and 3, contain the most overlapping cosmid clones, while the other contigs are made up of fewer than 6 clones each. The utility of these contigs is that once an overlapping clone can be found to link with the anchoring

Figure 4.5 Amplified fragment length polymorphism (AFLP) of selected cosmids.

Cosmids analysed side-by-side on this AFLP gel can be sorted into contiguous sets of cosmids (contigs) by matching banding patterns. In this example, those cosmids in lanes marked with an asterisk (*) fall into one contig, while those marked with a cross (+) fall into a second contig. The size marker at the left indicates the size in basepairs.

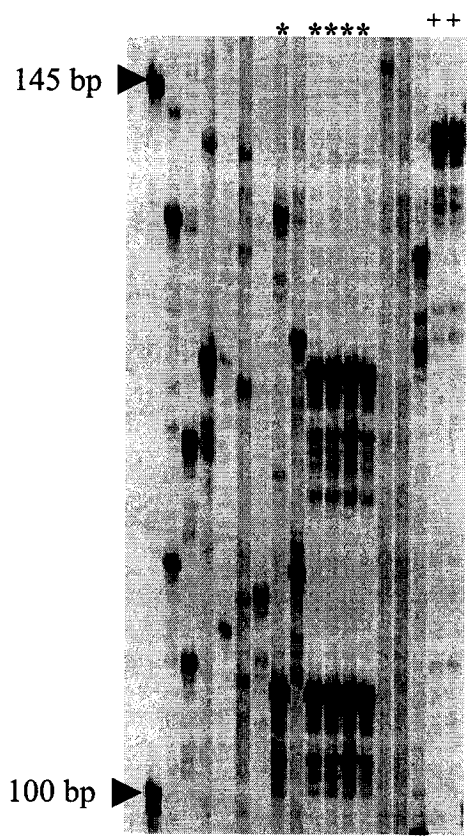

* ****
++
145 bp
100 bp

Table 4.2: Putative contigs formed among cosmids analysed by the amplified fragment length polymorphism (AFLP) technique.

Contig 1	Contig 2	Contig 3	Contig 4	Contig 5
2840C2	2630-6	12D	12B	2710-51A
4w/9-1	2630-1B/D	12D1	12-5	
4w/9-4	P2-9	12E1		
4w/9-11	P2-16	12E2		
4w/9-12		12-4		
4w/9-14		12-14		
2790-6A		6D30		
2790-33		3/6-15-2		
3-2-1A		3/6-13-14		
3-2-1E		3/6-5-3		
P2-12		lib6C1.1		
IJC1		lib6A2.5		
IJD1		3E1		
IJC2		3E2		
All 2740 cosmids		3D1		
2710-5B				
2710-4B				

Contig 6	Contig 7	Contig 8	Contig 9	Contig 10
4w/9-7	2790-16	2790-10	2790-14	2840A3
2630-1A	2790-20	2790-24	2790-26	2820-7
2790-15				2820-7A
				2820-1
				2820-12A

2840A3-containing contig (10), it should be possible to link this new cosmid to one of the other contigs, thereby extending further into the region where *Rfp* is found. For example, a clone overlapping both contig 10 and contig 3 would effectively cover the entire region in question. Otherwise, a stepwise procedure of clone selection and contig matching using the AFLP technique will close the gap.

Two PPR containing cosmids were selected using the PPR-gene specific probe, P2, that was generated using primers designed during "Background Work" according to proprietary *Brassica* sequences. When these two cosmids were analyzed via AFLP alongside other cosmids picked with the same probe they were deemed representative of the two distinct banding patterns within the group. This group of cosmids showed an unusual AFLP pattern in which there was a complete absence of higher molecular weight bands greater than 110 base pairs. This differs significantly from other cosmids, which show the greatest density of bands at 500bp and lower. This could be a result of the repetitive nature of PPR gene structure, leading to a higher cutting frequency in PPR-encoding DNA. Upon sequence analysis two of the PPR-containing cosmids, IJC2 and P2-9, were found to contain sequences similar to *Arabidopsis* orthologous genes At1g12760 through At1g12820. If one of these cosmids can be anchored in the *Rfp*-containing region, one of the PPR genes may be the restorer of fertility for *pol* CMS.

Discussion

The size and complexity of the genome structure of *Brassica* species presents significant challenges to gene mapping studies. Although the restorer gene *Rfp* was introgressed into *Brassica rapa* from *Brassica napus* to reduce the complexity of the genome to be analysed, there are still about three copies of each genomic region found in the haploid genome of *B. rapa*. The presence of paralogous copies of each gene means that marker-based anchoring of each clone must be performed to create a valid contig of overlapping clones.

The high degree of sequence conservation between *Brassica* and *Arabidopsis* allows the exploitation of synteny between the two in the design of primers based on *Arabidopsis* orthologous genes which are then used to amplify *Brassica* sequences. However, if less well conserved sequences are inadvertently picked, or the 3′ nucleotide of the primer differs from the *Brassica* DNA sequence, this may result in an inability to amplify DNA from *Brassica*. Alternately, one may amplify a paralogous copy of the *Brassica* gene of interest, thus creating a probe specific for one of the three paralogous copies not found in the region of interest. Other differences, such as loss of conservation of gene number, order or placement, can be a limiting factor in the utility of *Arabidopsis*-syntenic probes. In addition, sequences near PPR-encoding genes are more likely to diverge from synteny with *Arabidopsis* (see Chapter 5) and are less likely to yield useful clones.

Clones isolated in this study have been effective in reducing the region in which *Rfp* is found. Many pentatricopeptide repeat structures have also been identified

in two of the clones sequenced, though it remains to be determined if one of these clones can be anchored in the region where *Rfp* is found.

The next step in the isolation of *Rfp* is to create more specific, more effective probes based on these findings. The creation of probes for intergenic sequences of anchored clones will likely yield the most specific results and reduce the time and effort required to isolate positive cosmids. This method was not used initially since it is impractical to design intron-based or intergenic probes based on *Arabidopsis* sequence data, as these sequences are likely to be more highly diverged from *Brassica* than the intron sequences. An attempt to link other cosmid clones to the anchored clone 2840A3 is being performed by end-sequencing of positive clones. However, due to the positioning of 2840A3 further to the right of the region than anticipated, an overlap is unlikely. The gap between it and the next cosmids will have to be closed by another library screen using probes developed from 2840A3 sequences.

References

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

Brown, G.G., Formanov , N., Jin, H., Wargachuk, R., Dendy, C., Patil, P., Laforest, M., Zhang, J., Cheung, W.Y. and Landry, B.S. (2003) The radish *Rfo* restorer gene of Ogura cytoplasmic male sterility encodes a protein with multiple pentatricopeptide repeats. Plant J. **35**, 262-272.

Cavell, A.C., Lydiate, D.J., Parkin, I.A.P, Dean, C. and Trick, M. (1998) Collinearity between a 30-centimorgan segment of *Arabidopsis thaliana* chromosome 4 and duplicated regions within the *Brassica napus* genome. Genome **41**, 62-69.

Ewing, B. and Green, P. (1998) Basecalling of automated sequencer traces using phred. II. Error probabilities. Genome Research 8:186-194.

Formanov , N., Li, X.-Q., Ferrie, A.M.R., DePauw, M., Keller, W. A., Landry, B. and Brown, G.G. (2006). Towards positional cloning in *Brassica napus*: generation and analysis of doubled haploid *B. rapa* possessing the *B. napus pol* CMS and *Rfp* nuclear restorer gene. Plant Mol. Biol., in press.

Jackson, S.A., Cheng, Z., Wang, M.L., Goodman, H.M. and Jiang, J. (2000) Comparative Fluorescence *in situ* hybridization mapping of a 431-kb *Arabidopsis thaliana* bacterial artificial chromosome contig reveals the role of chromosomal duplications in the expansion of the *Brassica rapa* genome. Genetics **156**, 833-838.

Jean, M., Brown, G.G. and Landry, B.S. (1997) Genetic mapping of nuclear fertility restorer genes for the 'Polima' cytoplasmic male sterility in canola (*Brassica napus* L.) using DNA markers. Theor. Appl. Genet. **95**, 321-328.

Lagercrantz, U., Puterill, J., Coupland, G. and Lydiate, D.J. (1996) Comparative mapping in *Arabidopsis* and *Brassica*, fine scale genome collinearity and congruence of genes controlling flowering time. Plant J. **9**, 13-20.

Lagercrantz, U. and Lydiate, D. (1996) Comparative genome mapping in *Brassica*. Genetics **144**, 1903-1910.

Li, X.-Q., Jean, M., Landry, B.S. and Brown, G.G. (1998). Restorer genes for different forms of *Brassica* cytoplasmic male sterility map to a single nuclear locus that modifies transcripts of several mitochondrial genes. PNAS **95**, 10032-10037.

Li, G., Gao, M., Yang, B. and Quiros, C.F. (2003) Gene for gene alignment between the *Brassica* and *Arabidopsis* genomes by direct transcriptome mapping. Theor. Appl. Genet. **107**, 168-180.

Meyerowitz, E.M and Pruitt, R.E. (1985) *Arabidopsis thaliana* and plant molecular genetics. Science **229**, 1214-1218.

O'Neill, C.M. and Bancroft, I. (2000) Comparative physical mapping of segments of the genome of *Brassica oleracea var. alboglabra* that are homeologous to sequenced regions of chromosomes 4 and 5 of *Arabidopsis thaliana*. Plant J. **23**, 233-243.

Parkin, I.A.P., Lydiate, D.J. and Trick, M. (2002) Assessing the level of collinearity between *Arabidopsis thaliana* and *Brassica napus* for *A. thaliana* chromosome 5. Genome **45**, 356-366.

Quiros, C.F. (2001) DNA-based marker maps of *Brassica*. In: *DNA-based markers in plants* (Phillips, R.L. and Vasil J.K eds). Kluwer Academic Publishers, London, 201-238.

Ryder, C.D., Smith, L.B., Teakle, G.R. and King, G.J. (2001) Contrasting genome organisation: two regions of the *Brassica oleracea* genome compared with collinear regions of the *Arabidopsis thaliana* genome. Genome **44**, 808-817.

Scheffler, J.A., Sharpe, A.G., Schmidt, H., Sperling, P., Parkin, I.A.P., Luhs, W., Lydiate, D.J. and Heinz, E. (1997) Desaturase multigene families of *Brassica napus* arose through genome duplication. Theor. Appl. Genet. **94**, 583-591.

Schmidt, R., Acarkan, A. and Boivin, K. (2001) Comparative structural genomics in the Brassicaceae family. Plant Physiol. Biochem. **39**, 253-262.

CHAPTER 5

PENTATRICOPEPTIDE REPEAT ENCODING GENE LOCATION IN PLANTS IS NOT CONSERVED IN ORTHOLOGOUS AND PAROLOGOUS COMPARISONS

Several researchers have noted clustering of related pentatricopeptide repeat genes alongside true restorers of fertility in other plant systems. Further examination of these clusters of PPRs in radish (Brown *et al.*, 2003) showed that these PPR genes are asyntenically located in comparison with sequences of *Arabidopsis thaliana,* although flanking genes maintain synteny. Further investigation of the asynteny of PPR-encoding genes in various plants is presented in this chapter.

This work is being prepared as a manuscript that will be submitted for publication. The manuscript presented is co-authored by Dr. Gregory G. Brown in his capacity as research supervisor.

Introduction

The pentatricopeptide repeat (PPR) peptide motif, as first described by Small and Peeters (2000), is a degenerate 35 amino acid sequence, closely related to the tetratricopeptide (TPR) motif, which is often found tandemly repeated in a widespread family of eukaryotic proteins. Each PPR domain is thought to contain two distinct alpha helices. In proteins with tandemly repeated PPRs, it is thought that each PPR domain forms a separate superhelical turn, in which the two distinct alpha-helices are arranged in an antiparallel configuration (Small and Peeters, 2000). This superhelix is predicted to enclose a central spiral groove with a positively charged ligand binding surface capable of holding and binding a single RNA chain (Small and Peeters, 2000). PPR proteins are known to mediate specific RNA processing events including RNA editing (Kotera *et al.*, 2005) and transcript processing (Nakamura *et al.*, 2004).

Although the PPR gene family is widespread among eukaryotes, the numbers of such genes in both animal and fungal genomes is relatively small. In plants, however, the size of this gene family is greatly expanded. In *Arabidopsis thaliana* there are 441 identified PPR genes (Lurin *et al.*, 2004), a number which is predicted to exceed the total number of PPR genes yet to be identified in rice (International Rice Genome Sequencing Project, 2005). At present, 655 PPR proteins have been predicted in rice (Lurin *et al.*, 2004).

Analysis of the PPR gene content of the *Arabidopsis* genome by Lurin *et al.* (2004) elucidated several categories and subcategories of PPR genes. The largest

category encodes proteins that are composed of tandem repeats of the "classical" 35 amino acid PPR motif initially described by Small and Peeters (2000) and now referred to as the P-type repeat. Lurin *et al.* (2004) were able to differentiate three additional PPR-related motifs found in PPR-encoding genes. Two of these motifs, S and L1, are tandemly arrayed with the classical P-type motif in a repeated P-L1-S (PLS) pattern, with the third motif, L2, replacing L1 in the last repeat pattern at the C-terminal end of the protein. Their analyses also showed that the PLS subfamily of PPR-encoding genes is unique to plants and not found in other systems.

Four subgroups of PPR proteins from the PLS subfamily differ in the structure of their C-terminal domains. Although two of the subgroups, E and E+, are highly degenerate in their C-terminal sequences, the DYW subgroup shows some conservation of amino acid residues. It has been suggested that this C-terminal domain may function as a catalytic domain for these PPR proteins (Lurin *et al.*, 2004). One PPR gene belonging to the PLS subfamily is *Emb175*, a gene essential for plant embryogenesis. *EMB175*, like many PPR proteins is targeted to the plastid (Cushing *et al.*, 2005).

Another major group of plant-specific PPR genes are the restorer of fertility (Rf) genes. These nuclear-encoded genes act to suppress male sterility associated with expression of mitochondrial encoded sterility-associated genes. Rf genes identified thus far in petunia, rice and radish belong to the P subfamily of PPR genes (Bentolila *et al.*, 2002; Kazama and Toriyama, 2003; Brown *et al.*, 2003, respectively).

Expansion of the complement of PPR genes within plant genomes may have occurred through gene duplication. In *Arabidopsis*, ancient large-scale genome duplication events have resulted in multiplication of loci and regions of synteny, where gene number and location are conserved as paralogous copies. Gene duplication can also arise from tandem and segmental gene duplication, creating clusters of identical genes that diverge over time (Cannon *et al.*, 2004; Leister, 2004). PPR genes in *Arabidopsis* are generally not clustered, except in one location on chromosome 1 (Lurin *et al.*, 2004). However, tandem clusters of PPR genes have been observed in Petunia (Bentolila *et al.*, 2002), radish (Brown *et al.*, 2003) and rice (Kazama and Toriyama, 2003; Komori et al., 2004; Akagi et al., 2004). The synteny observed in duplicated genomic regions within a genome or between orthologous copies in related genomes can be exploited in gene mapping. However, disruptions of synteny can occur due to gene loss, rearrangement, acquisition or duplication. The source of these structural changes can be due to tandem and segmental duplication, as discussed above, but could also be attributed to aberrant homologous recombination, selection, or changes introduced by transposition events.

The radish restorer of fertility, *Rfo*, is found in a cluster of PPR genes at a genomic site where no corresponding PPR gene is found in the syntenic region of *Arabidopsis* (Brown *et al.*, 2003). We report here that other PPR genes display a characteristic lack of synteny in comparisons of both orthologous and paralogous plant genomic regions. We show that while non-PPR genes are largely colinear in arrangement and identical in orientation between different related regions,

PPR genes are rarely maintained in the same position or orientation when two related regions are compared. We show that PPR gene family members share characteristics with plant disease resistance genes (R genes); in particular we present evidence that at least some PPR genes, as per R genes, are subject to diversifying selection which acts to multiply and distribute copies of the genes. Our results also suggest that the Birth-and-Death process initially described for immunoglobulin genes (Nei *et al.*, 1997), and adapted by Michelmore and Meyers (1998) for R genes, is affecting the duplication and divergence of PPR genes.

Materials and Methods

Sequence analysis

Sequencing and sequence assembly was performed by DNALandmarks (St-Jean-sur-Richelieu) (Brown *et al.*, 2003) and Genome Quebec (Montreal) using the Applied Biosystems 3730XL DNA analyzer for capillary sequencing and the Phred/Phrap programs for some of the sequence assembly. Additional sequences obtained via shotgun sequencing were assembled using **CodonCode Aligner v.1.3.4** (www.codoncode.com). Sequences were analysed using **ORF finder** (Tatiana Tatusov and Roman Tatusov; www.ncbi.nlm.nih.gov/gorf/gorf.html) to detect ORFs and **Genscan** (genes.mit.edu/GENSCAN.html) to detect promotor regions, introns/exons and polyA signals. **Blast** and **Blast2Sequences** (www.ncbi.nlm.nih.gov/blast) were used for data mining from nucleotide and protein databases and for aligning pairs of sequences. Tree building was performed with **TreeTop** (www.genebee.msu.su/services/phtree_reduced.html) using the Blosum62 matrix and phylip tree building software. Multiple sequence alignments were performed using ClustalW online (http://www.ebi.ac.uk/clustalw/). The output was shaded using Boxshade online (http://www.ch.embnet.org/software/ BOX_form.html). Subcellular targeting predictions were made using online programs Mitoprot (http://ihg.gsf.de/ihg/mitoprot.html) and Predotar (http://genoplante-info.infobiogen.fr/predotar/predotar.html).

Criteria for choosing pairs of duplicated genes included online annotation mapping (www.tigr.org; www.ncbi.nlm.nih.gov) and BLAST sequence alignment of the entire CDS and protein sequences with a cutoff expect value of 1e-20 and bit score of >100.

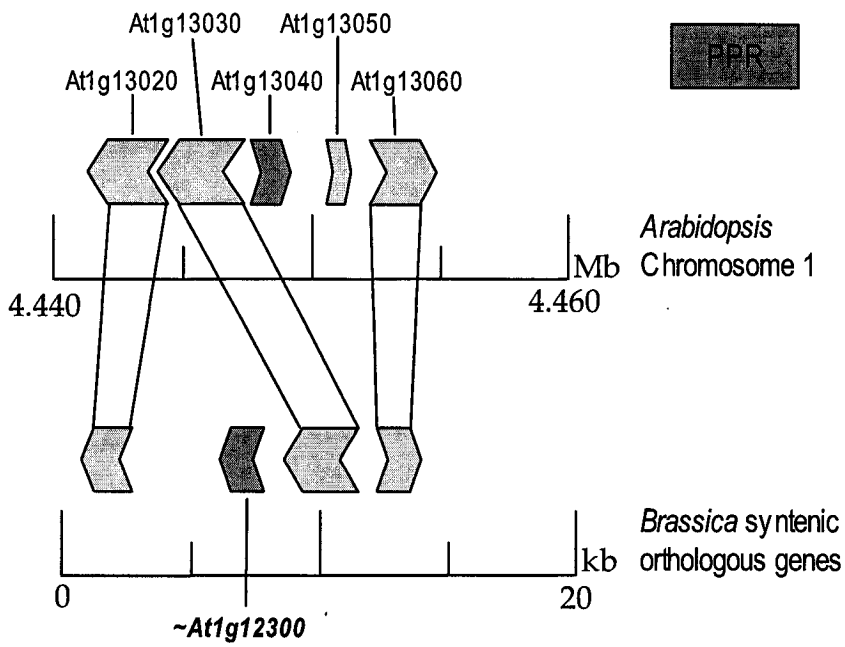
Results

Locations of PPR genes are highly variable between co-linear regions of Arabidopsis and Brassica *or* Raphanus *genomes*

We have sequenced *Brassica* genomic regions in an effort to identify and characterize the fertility restorer gene *Rfp* of *Brassica napus*. One such region contained four predicted protein-coding genes, one of which could potentially specify a protein with nine PPR domains. Comparison with the *Arabidopsis* genome sequence revealed a related region on chromosome one, spanning five predicted genes, At1g13020 through At1g13060 (Figure 5.1). Three of these genes, At1g13020, At1g13030 and At1g13060 are co-linear in arrangement and identical in orientation with their *Brassica* genome counterparts, indicating that synteny is preserved between these chromosomal regions of the two species. A high degree of sequence similarity/identity (47-76%/47-78%) is observed between the proteins encoded by these *Arabidopsis* genes and their *Brassica* counterparts. As is commonly observed in genomic comparisons between *Arabidopsis* and *Brassica*, one of the *Arabidopsis* genes, At1g13050, has no apparent counterpart in the *Brassica* sequence (O'Neill and Bancroft, 2000, Lukens *et al.*, 2003, Parkin *et al.*, 2005). Most segments of the *Arabidopsis* genome are represented at multiple sites in *Brassica* genomes and the resulting high level of genetic redundancy in *Brassica* may lead to the loss of coding sequences in one or more such sites.

Figure 5.1 Predicted genes in *Brassica* and the syntenic segment of the *Arabidopsis* genome

Protein-coding sequences in the *Brassica* genome (bottom) as predicted by Genscan. Unfilled regions indicate predicted introns in *Brassica*. Dark filled arrows indicate PPR-encoding sequences. Protein-coding sequences in the *Arabidopsis* genome reflect the Munich Information Center for Protein Sequences (MIPS) annotation. In this Figure, predicted exons are not indicated. The direction of the arrows indicates the direction of transcription/translation. The connecting lines indicate orthologous coding regions in the two genomic sequences. Distances in megabasepairs in *Arabidopsis* are indicative of the location of sequences on chromosome 1. The PPR encoding gene in *Brassica* is asyntenic with respect to the *Arabidopsis* sequences shown here. In addition, it shares homology with *Arabidopsis* At1g12300, and not with it's most closely linearly aligned partner At1g13040.

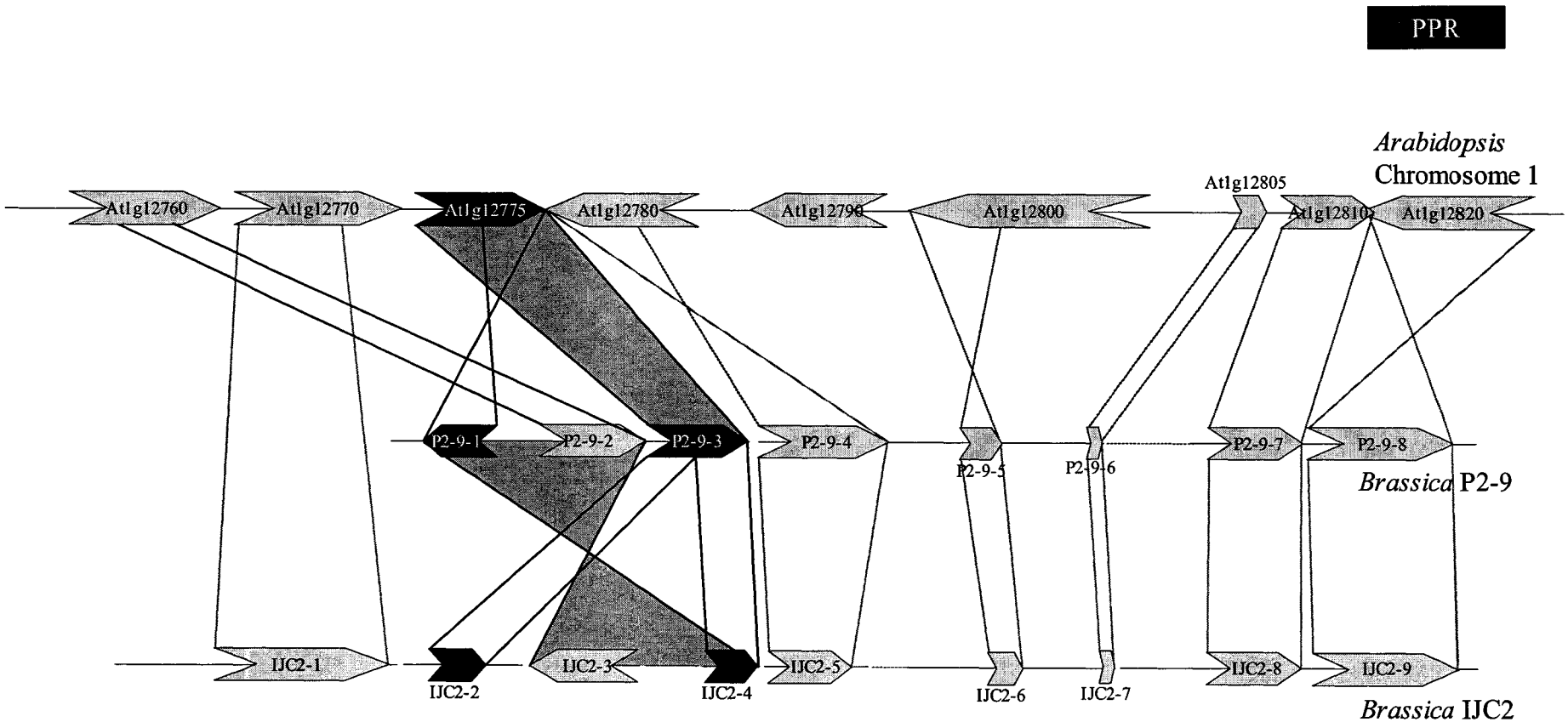

At1g13030
At1g13050
At1g13020
At1g13040
At1g13060
PPR
Arabidopsis
Chromosome 1
Mb
4.440
4.460
Brassica syntenic
orthologous genes
kb
0
20
~At1g12300

The syntenic region of the *Arabidopsis* genome also contains a predicted gene (At1g13040) that could potentially specify a protein with six PPR domains. In contrast with At1g13020, At1g13030 and At1g13060, this protein, however, has little similarity with PPR protein encoded by the PPR gene in the co-linear *Brassica* region (26% identity [I], 46% similarity [+]). Moreover, its location, between At1g13030 and At1g13050, is different from that of the *Brassica* PPR gene, which is positioned between orthologs of At1g13020 and At1g13030; its transcriptional orientation, with respect to the co-linear genes of the region, also differs. Interestingly, this *Brassica* PPR protein does show a high degree of sequence similarity with PPR proteins present at distinct sites on chromosome one. In particular, it possesses 69% identity and 81% similarity with the protein encoded by At1g12300, a PPR gene located in a cluster of such genes near the 4.3 megabase (Mb) mark of chromosome one. Thus, we observe a preservation of synteny for most genes between these *Arabidopsis* and *Brassica* genomic regions but an apparent lack of synteny for the PPR genes. This suggests that the function and order of the non-PPR genes in the region has been conserved during the evolution of the Brassicaceae, but that one, and possibly both of the PPR genes in these two related chromosome regions has descended from a progenitor located at a distinct, non-syntenic chromosomal site.

The lack of synteny with respect to related PPR genes was evident in additional comparisons of *Brassica* and *Arabidopsis* genomic sequences. Two cosmids, P2-9 and IJC2, containing *Brassica* DNA sequences encoding PPR domains were sequenced. P2-9 and IJC2 are paralogous regions that show high structural similarity to one another (Figure 5.2). P2-9 contains eight predicted coding genes

Figure 5.2 Predicted genes in two regions of *Brassica napus* and the syntenic segment of the *Arabidopsis* genome.

Protein-coding sequences in the *Brassica* genome fragments P2-9 and IJC2 (as predicted by Genscan). For clarity, *Brassica* predicted genes are numbered. Protein-coding sequences in the *Arabidopsis* genome reflect the Munich Information Centre for Protein Sequences (MIPS) annotation. Predicted exons are not indicated. The direction of arrows indicates the direction of transcription/translation. Pairs of connecting lines indicate orthologous coding sequences between *Arabidopsis* and *Brassica*; darker pairs of lines connect orthologous PPR encoding sequences and other important sequences. Shading between connecting lines is added for emphasis and clarity. Dark-filled arrows indicate PPR encoding sequences. *Brassica* sequences shown are limited by the size of fragment sequenced, thus genes that are outside these sequences are not available.


PPR
Arabidopsis
Chromosome 1
At1g12760
At1g12770
At1g12775
At1g12780
At1g12790
At1g12800
At1g12805
At1g12810
At1g12820
P2-9-1
P2-9-2
P2-9-3
P2-9-4
P2-9-5
P2-9-6
P2-9-7
P2-9-8
Brassica P2-9
IJC2-1
IJC2-2
IJC2-3
IJC2-4
IJC2-5
IJC2-6
IJC2-7
IJC2-8
IJC2-9
Brassica IJC2

and IJC2 contains nine. In both cases, two of the genes encode PPR domains (P2-9-1, P2-9-3, IJC2-2 and IJC2-4). These PPR gene sequences are 80-90% conserved with areas of higher conservation within exons. Between the two *Brassica* sequences, gene order and direction of transcription is conserved between the two cosmids except in the region surrounding PPR genes.

IJC2-2 and P2-9 both show extensive similarity to a region of *Arabidopsis* chromosome I spanning the nine genes flanked by At1g12760 and At1g12820 (Figure 5.2); one of these *Arabidopsis* genes, At1g12775, encodes a PPR domain protein. In nucleotide comparisons, 85-90% of gene sequences are conserved between *Arabidopsis* and *Brassica* in this region. Gene order is maintained between *Arabidopsis* and the two *Brassica* sequences, with a few exceptions. No counterpart of At1g12790, and only an inverted segment of At1g12800, is found in the *Brassica* cosmids.

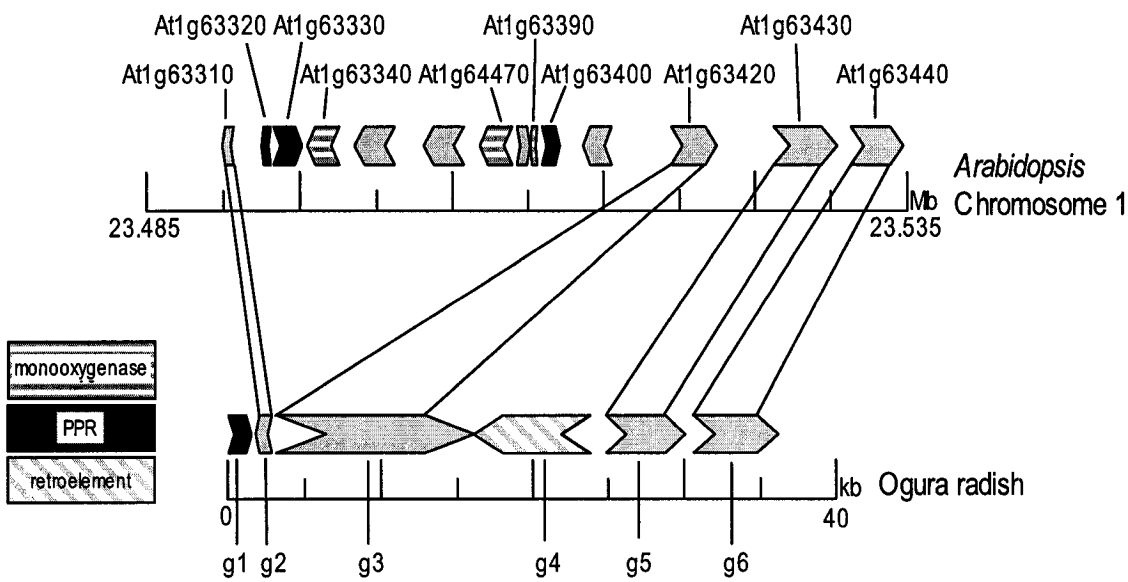
The *Brassica* gene P2-9-3 is similar in both sequence and orientation to the *Arabidopsis* PPR gene At1g12775. P2-9-1 is a duplication of the 3′ end of P2-9-3. In addition, P2-9-1 and P2-9-2 are represented in IJC2 as IJC2-4 and IJC2-3, respectively, and are inverted in orientation with respect to their P2-9 counterparts. It is likely that after the genomic duplication leading to the formation of these paralogous regions a local rearrangement occurred. This rearrangement may have excised the P2-9-1/P2-9-2//IJC2-4/IJC2-3 fragment and reinserted it into the genome, knocking out the 3′ end of IJC2-2 and replacing it with the inverted fragment. This may have occurred through the homologous recombination of P2-9-1 and the 3′ end of P2-9-3//IJC2-2.

The *Brassica* PPR-encoding ORF P2-9-1 is found in a genomic region that corresponds to sequences flanking At1g12760; no PPR domains occur in this location in the *Arabidopsis* sequence. As explained above, the duplication of the *Brassica* ortholog At1g12775, P2-9-3, likely resulted in the presence of multiple PPR sequences. The positioning of P2-9-1 to the left of the At1g12760 ortholog P2-9-9 suggests that a genome rearrangement occurred at this location after the split between *Arabidopsis* and *Brassica* which resulted in the movement of At1g12760/P2-9-2 sequences. Thus, as in other genomic comparisons, PPR encoding regions in P2-9, IJC2 and the corresponding *Arabidopsis* chromosome I segment are more highly rearranged than flanking regions encoding other types of proteins.

The segment of the radish (*Raphanus sativum*) genome encoding the *Rfo* restorer gene has been shown to possess extensive co-linearity with the *Arabidopsis* genome (Brown *et al.*, 2003). One of the radish genes close to this region, g1, encodes a protein with PPR motifs (Figure 5.3). The syntenic non-PPR encoding genes of *Arabidopsis* and radish in this region show a high degree of conservation of amino acid sequence (63-88%I/67-92%+). Radish g1, however, does not conform to the synteny with *Arabidopsis*, and is located on the opposite side of g2/At1g63310 from its *Arabidopsis* counterparts At1g63320, At1g63330 and At1g63400 with which it is most closely linearly aligned. Although g1 shares similarity to these three PPR genes (35-40%I/55-60%+), it is more similar to *Rfo* (67%I/75%+) than to any *Arabidopsis* gene. Other highly related matches to g1 include the restorers of fertility *Rf1* in rice (32%I/55%+) and *Rf-1* in petunia

Figure 5.3 Predicted genes in radish and the syntenic segment of the *Arabidopsis* genome

Protein-coding sequences in the radish genome (bottom; as predicted by Genscan). Protein-coding sequences in the *Arabidopsis* genome reflect the Munich Information Center for Protein Sequences (MIPS) annotation. Predicted exons are not indicated. The direction of the arrows indicates the direction of transcription/translation. Dark filled arrows indicate PPR encoding sequences. Arrows with horizontal stripes indicated monooxygenase encoding genes. The gene g4, indicated with diagonal stripes, encodes a non-LTR retroelement reverse transcriptase. Orthologous coding sequences in the two genomic sequences are indicated by connecting lines. Distances in megabasepairs in *Arabidopsis* are indicative of the location of sequences on chromosome 1.


At1g63320
At1g63330
At1g63390
At1g63430
At1g63310
At1g63340
At1g64470
At1g63400
At1g63420
At1g63440
Arabidopsis
Chromosome 1
Mb
23.485
23.535
monooxygenase
PPR
retroelement
kb
Ogura radish
0
40
g1
g2
g3
g4
g5
g6

(36%I/57%+). A major break in synteny exists between radish and *Arabidopsis* in this region; a stretch of DNA flanked by PPR genes and also by three flavin-containing monooxygenase-related genes is present in *Arabidopsis* but absent in radish (Figure 5.3). It is possible that this fragment was inserted in *Arabidopsis* or excised via homologous recombination from radish; the presence of PPR genes flanking the monooxygenase genes may be of significance in the excision of this region from radish or the introduction of it into *Arabidopsis*, although the mechanism responsible for this change remains unknown at present.

Variation in PPR gene location between two paralogous gene regions on chromosome I of Arabidopsis

Several restorer genes from various plant species have thus far shown homology to a cluster of PPR genes found in the *Arabidopsis thaliana* genome (Brown *et al.*, 2003, Bentolila *et al.*, 2002, Kazama and Toriyama, 2003). This particular set of PPR genes is the largest grouping in the *Arabidopsis* genome of highly homologous PPR genes. This genome segment is located at about the 23 Mb mark of chromosome 1 and includes loci At1g62260 through At1g63630, encompassing 18 PPR genes and pseudogenes (Lurin *et al.*, 2004). *Rfp*, the restorer of fertility for polima CMS of *Brassica napus*, has been mapped to a genomic region that is syntenic to a portion of the *Arabidopsis* genome located near the 4.3 Mb coordinate of chromosome 1.

Previous studies of whole genome internal duplications of the *Arabidopsis* genome have not noted any homology between the 23 Mb and 4.3 Mb regions

(Ziolkowski *et al.*, 2003; Blanc *et al.*, 2000). However, when *Arabidopsis* genome regions surrounding the 4.3 and 23 Mb coordinates are closely examined, it becomes evident that there is a group of genes that are conserved between these two segments (Figure 5.4). Recently, this duplication was noted on the website of the institute for genomic research (TIGR): (http://www.tigr.org/tdb/e2k1/ath1/Arabidopsis_genome_duplication.shtml). In the 23.370-23.470 Mb region, there are 27 genes identified by the Munich Information Centre for Protein Sequences (MIPS), comprising five PPR genes, five genes encoding hypothetical proteins and 17 other genes. In the 4.285-4.365 Mb region there are 19 genes, including three PPR genes, two genes encoding hypothetical proteins and a transposase-encoding gene. The two regions share nine predicted coding sequences displaying significant sequence similarity; 56% of the non-PPR encoding genes in the 4.3 Mb region possess a paralogous counterpart in the 23 Mb region. In the 23Mb region 40% of non-PPR encoding genes share a paralogous counterpart in the 4.3 Mb region (Tables 5.1, 5.2). Eight of the nine conserved genes are maintained in the same order and transcriptional direction between the two fragments.

It is suggested that the gene At1g12770 could be aligned with At1g63250. Although the two genes share an identical coding function according to online annotation, they share no significant homology at the nucleotide level and limited segments of amino acid similarity (Table 5.2). It is possible that one of these duplicated genes was redundant and became nonfunctional. A loss of function would account for subsequent sequence divergence of At1g12770 from At1g63250. In addition, At1g63250, is located at a distance of over 35,000 base

Figure 5.4 Comparison of two *Arabidopsis* duplicated regions on chromosome 1

Protein-coding sequences in the *Arabidopsis* genome chromosome 1 reflect the Munich Information Center for Protein Sequences (MIPS) annotation. Predicted exons are not indicated. The direction of the arrows indicates the direction of transcription/translation. Dark filled arrows indicate PPR encoding sequences. Paralogous coding sequences in the two genomic regions are indicated by pairs of connecting lines. Distances in megabasepairs are indicative of the location of sequences on chromosome.

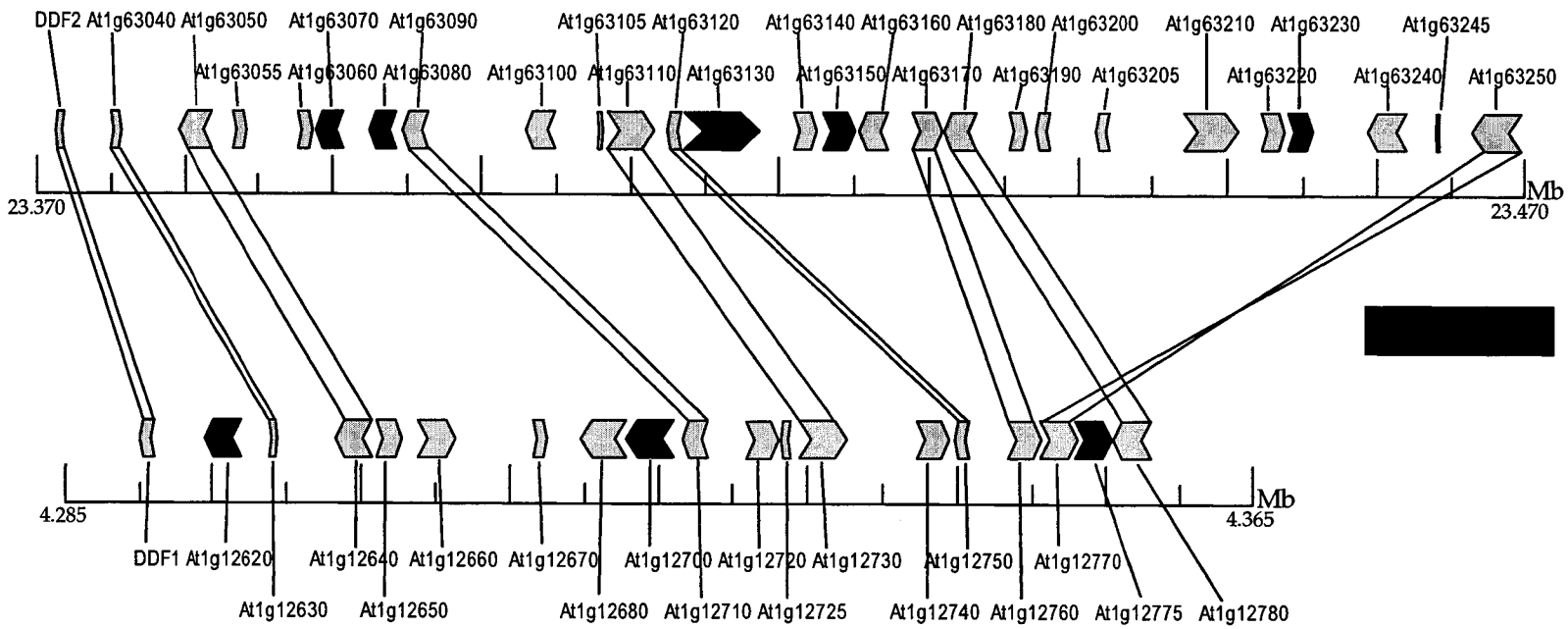

DDF2
At1g63040
At1g63050
At1g63055
At1g63060
At1g63070
At1g63080
At1g63090
At1g63100
At1g63105
At1g63110
At1g63120
At1g63130
At1g63140
At1g63150
At1g63160
At1g63170
At1g63180
At1g63190
At1g63200
At1g63205
At1g63210
At1g63220
At1g63230
At1g63240
At1g63245
At1g63250
23.370
Mb
23.470
4.285
Mb
4.365
DDF1
At1g12620
At1g12630
At1g12640
At1g12650
At1g12660
At1g12670
At1g12680
At1g12700
At1g12710
At1g12720
At1g12725
At1g12730
At1g12740
At1g12750
At1g12760
At1g12770
At1g12775
At1g12780

Table 5.1 Conservation of genes within two regions of *Arabidopsis* chromosome 1.

Percentage of conservation is based on number of genes conserved out of the total number of genes, not including PPR genes. PPR genes are discounted due to the difficulty in concretely matching PPRs by sequence comparison, and also because of the variability in number and location of PPR genes. The full 4.3 Mb region includes genes DDF1 through At1g12780. The full 23 Mb region includes genes DDF2 through At1g63250. The limited 23 Mb region includes genes DDF2 through At1g63180. Hypothetical genes are annotated according to the *Arabidopsis* annotation according to TAIR.

Location	4.3Mbase	23Mbase	23Mbase (limited)
Number of genes/ORFs	19	27	18
PPR-encoding genes	3	5	4
Conserved genes	9	9	8
Hypothetical proteins	2 (+ 1 transposon)	5	1
% conserved genes (all, discounting PPRs)	56	40	57

Table 5.2 Relatedness of genes of the 4.3 and 23 Mb syntenic regions of *Arabidopsis*.

Gene coding regions which are correlated in the duplicated regions are compared in a pairwise fashion to show their homology at the nucleotide level and their relative identity and similarity at the amino acid level. All pairwise comparisons show a high degree of likeness except for At1g12775 compared with At1g63250, although this may be due to redundancy of one copy leading to divergence of the sequence. These findings are consistent with the hypothesis that the two regions are related and are derived from a duplication event.

4.3 Mb region locus	Coding function	23 Mb region locus	Coding function	% nucleotide homology	% amino acid identity/ similarity
At1g12610 (DDF1)	DREB subfamily Transcription factor	At1g63030 (DDF2)	DREB subfamily Transcription factor	79%	77%/86%
At1g12630	DREB subfamily Transcription factor	At1g63040	DREB subfamily Transcription factor **pseudogene**	76%	65%/78%
At1g12640	O-acyl transferase protein	At1g63050	O-acyl transferase protein	84%	83%/89%
At1g12710	F-box family protein	At1g63090	F-box family protein	81%	76%/88%
At1g12730	CDC-like protein	At1g63110	CDC-like protein	83%	76%/84%
At1g12750	Rhomboid family protein	At1g63120	Rhomboid family protein	78%	65%/76%
At1g12760	Zinc finger protein	At1g63170	Zinc finger protein	80%	75%/83%
At1g12770	DEAD/DEAH box helicase family protein	At1g63250	DEAD/DEAH box helicase family protein	No significant similarity found	27%/46%
At1g12780	UDP-glucose epimerase	At1g63180	UDP-glucose epimerase	86%	89%/93%

pairs away from the rest of the conserved sequence. It seems likely that subsequent to the duplication event that gave rise to the paralogous 23 Mb and 4.3 Mb regions, a second structural change occurred, resulting in the displacement of either At1g12770 or At1g63250. If At1g63250 is disregarded, and the duplicated 23 Mb region is limited at At1g63180 on the right hand, centromeric side, the degree of gene conservation of the 23 Mb region increases from 40% to 57% (Table 5.2).

It has already been shown by Lurin *et al.* (2004) that PPR-containing genes of the *Arabidopsis* genome can be categorized by their structure and sequence similarity to each other, including their C-terminal domains. The group of PPR genes of the 23 Mb region are included in the P subfamily of PPR proteins and are highly similar to those in the 4.3 Mb region (Table 5.2). Within the 4.3 Mb region PPR genes At1g12620 and At1g12775 are 82% identical (I) and 90% similar (+) at the amino acid level, whereas they are each only about 58% identical and 74% similar to the PPR gene At1g12700. Within the 23 Mb region PPR proteins At1g63070, At1g63080, At1g63130 and At1g63150 are 68-78% identical and 78-88% similar. The outlying PPR gene, At1g63230, could theoretically encode a protein which is less similar to other PPR proteins in the 23 Mb group (52%I/67%+); however, it has been annotated as false by Lurin *et al.* (2004) due to the lack of a "plausible initiation codon". The divergent nonfunctional sequence may have arisen from the loss of a start codon, leading subsequently to lower similarity to related PPR sequences.

PPR genes in the two regions are highly similar to one another and the PPR proteins in these two regions are significantly less similar to PLS subfamily PPR proteins (ie. E, E+, PLS and DYW subgroups) encoded elsewhere in the *Arabidopsis* genome (Figure 5.5). When At1g12700 (4.3Mb cluster) is compared with At1g63130 (23Mb cluster) they are 52% identical and 69% similar. When At1g12700 is compared to PPR proteins from these other subgroups they are only 23% identical and 42% similar. Strikingly, when At1g12700 is compared to other PPR genes of the P subfamily found outside the 23 Mb cluster their proteins are only about 25% identical and 48% similar, a percentage that is not significantly different from the comparison to other PLS subfamily genes. Thus, the PPR proteins found in the duplicated regions of 4.3 Mb and 23 Mb of chromosome 1 are more similar to each other than they are to other PPR proteins, even within the same structural subfamily as described by Lurin *et al.* (2004).

When the PPR domain structure of all subfamilies and subgroups of PPR genes are considered, restorer genes and their related *Brassica*, *Arabidopsis*, and radish PPR genes shown in this study show a larger number of repeats (usually 8 or more) and a structure of repeats which are arranged in tandem. This contrasts with other unrelated PPR proteins, even from the P subfamily, which generally have fewer than 4 repeats widely spaced from each other (Table 5.3). The subcellular targeting of PPR genes, however, seems to be independent of subfamily or subgroup of PPR protein (Table 5.3).

As can be seen in Figure 5.4, the order of genes, and their direction of transcription, is conserved for the non-PPR encoding genes. PPR genes, however,

Figure 5.5 Phylogenetic relationship between PPR proteins of the 4.3 and 23 Megabase region of *Arabidopsis* chromosome 1 and PPR genes found elsewhere in the *Arabidopsis* genome.

The seven PPR proteins at the right hand side of this unrooted tree are all members of the P subfamily of PPR proteins and belong to the 4.3 or 23 megabase clusters of PPR genes. The other ten PPR proteins are a representative random sampling of PPR proteins of all 5 subfamilies and subgroups (P, E, E+, PLS, DYW, as defined by Lurin *et al.*, 2004) from chromosome 1 and another chromosome of *Arabidopsis* chosen at random. PPR proteins of the 4.3 and 23 megabase regions are more closely related to each other than they are to other P-subfamily genes or other PPR genes from other sub-groups, and form a distinct branch away from these PPR proteins.

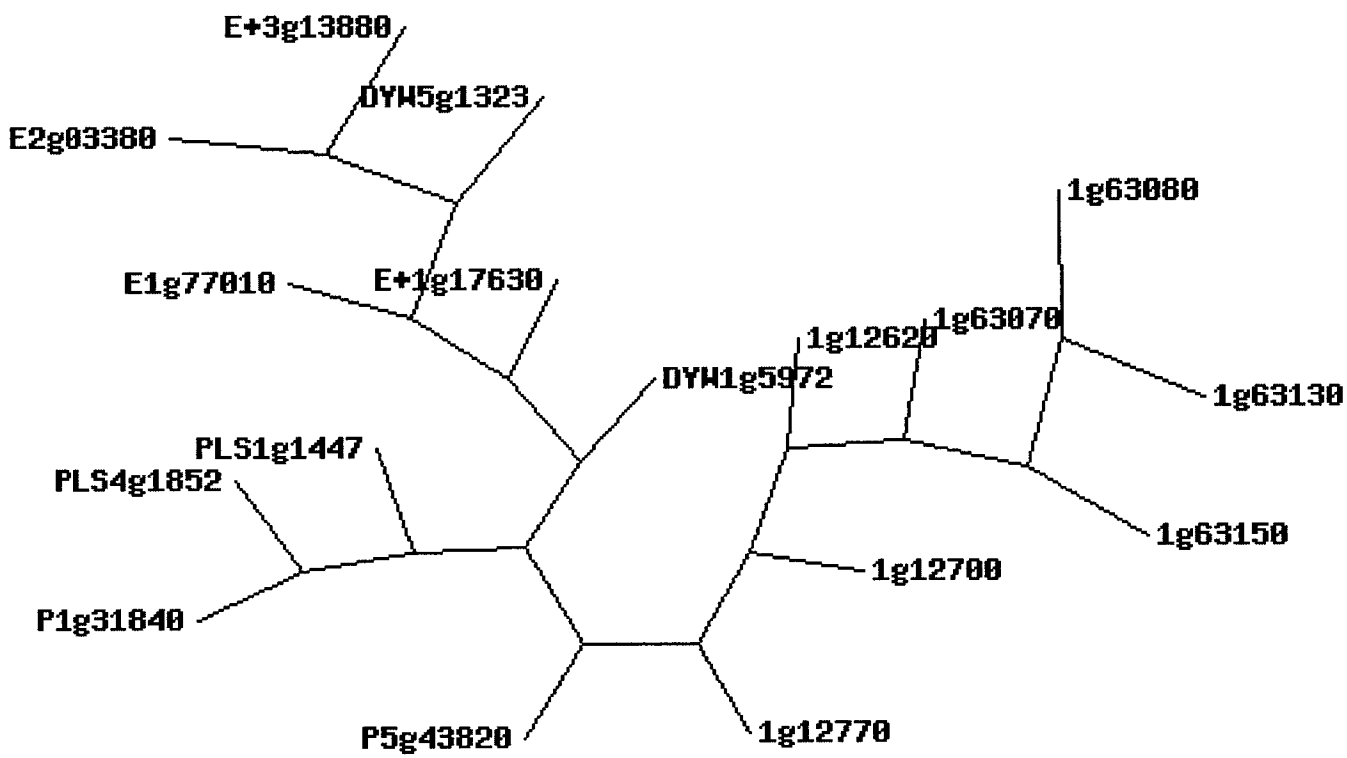
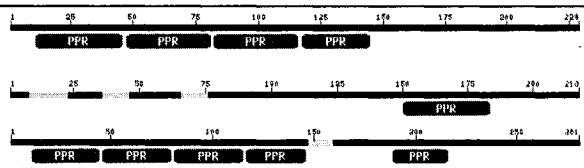
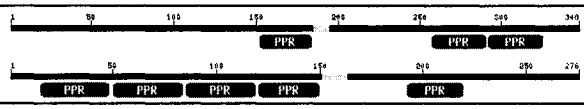
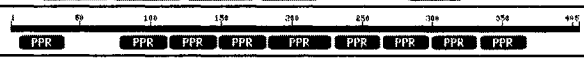
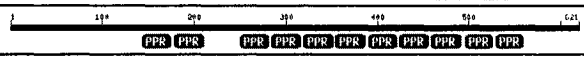
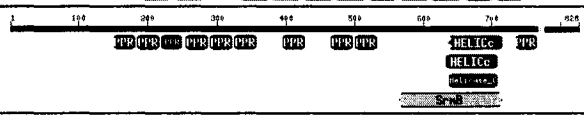
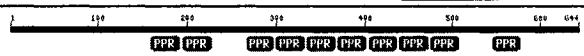
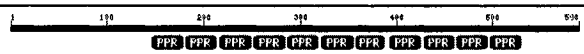
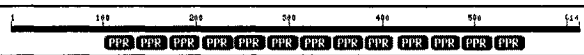
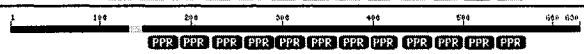
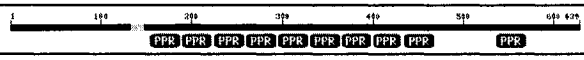
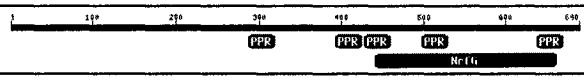
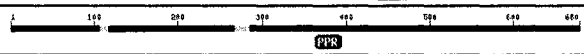
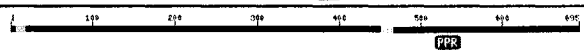
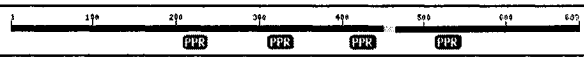
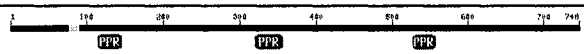
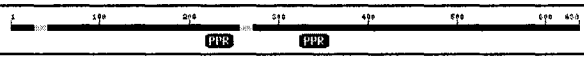
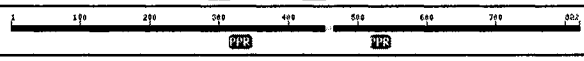
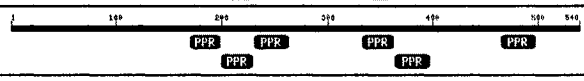
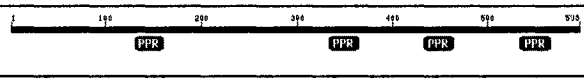

E+3g13880
DYW5g1323
E2g03380
1g63080
E1g77010
E+1g17630
1g63070
1g12620
DYW1g5972
1g63130
PLS1g1447
PLS4g1852
1g63150
1g12700
P1g31840
P5g43820
1g12770

Table 5.3 Structural annotation and targeting of PPR-encoding proteins from various plants.

Subcellular localization was predicted by Mitoprot and Predotar online programs. Structures were derived by the Conserved Domain Search (CDS) on the NCBI website. All gene entries beginning with "At" are from *Arabidopsis thaliana*.

	Gene	# PPRs	Subcellular localization	Structural Annotation
Brassica	Similar to At1g12300	9	Mitochondrial	
Brassica P2-9	Similar to At1g12775	10	None	
Brassica IJC2	Similar to At1g12775	8	None	
Radish	G1	9	None	
4.3 Mb	At1g12620	11	Mitochondrial	
	At1g12700	10	Mitochondrial	
	At1g12775	10	Mitochondrial	
23 Mb	At1g63070	11	Mitochondrial	
	At1g63080	13	Mitochondrial	
	At1g63130	12	Mitochondrial	
	At1g63150	10	Mitochondrial	
P	At1g31840	5	None	
P	At5g43820	1	Mitochondrial	
E	At1g77010	1	Mitochondrial	
E	At2g03380	4	Mitochondrial	
E+	At1g17630	3	Possibly mitochondrial	
E+	At3g13880	3	Possibly mitochondrial	
DYW	At1g59720	2	None	
DYW	At5g13230	2	Mitochondrial	
PLS	At1g14470	6	None	
PLS	At4g18520	4	Possibly mitochondrial	

appear to be distributed randomly throughout the two regions. Comparisons of the PPR sequences of the two regions do not reveal a significant correlation of homology between closely linearly aligned PPR genes. For example, At1g12700 shares 50% homology and 65% identity with At1g63070 and At1g63080, with which it is most closely linearly aligned, but this is about the same degree of similarity as is found between it and the other PPR genes in the 23 Mb region.

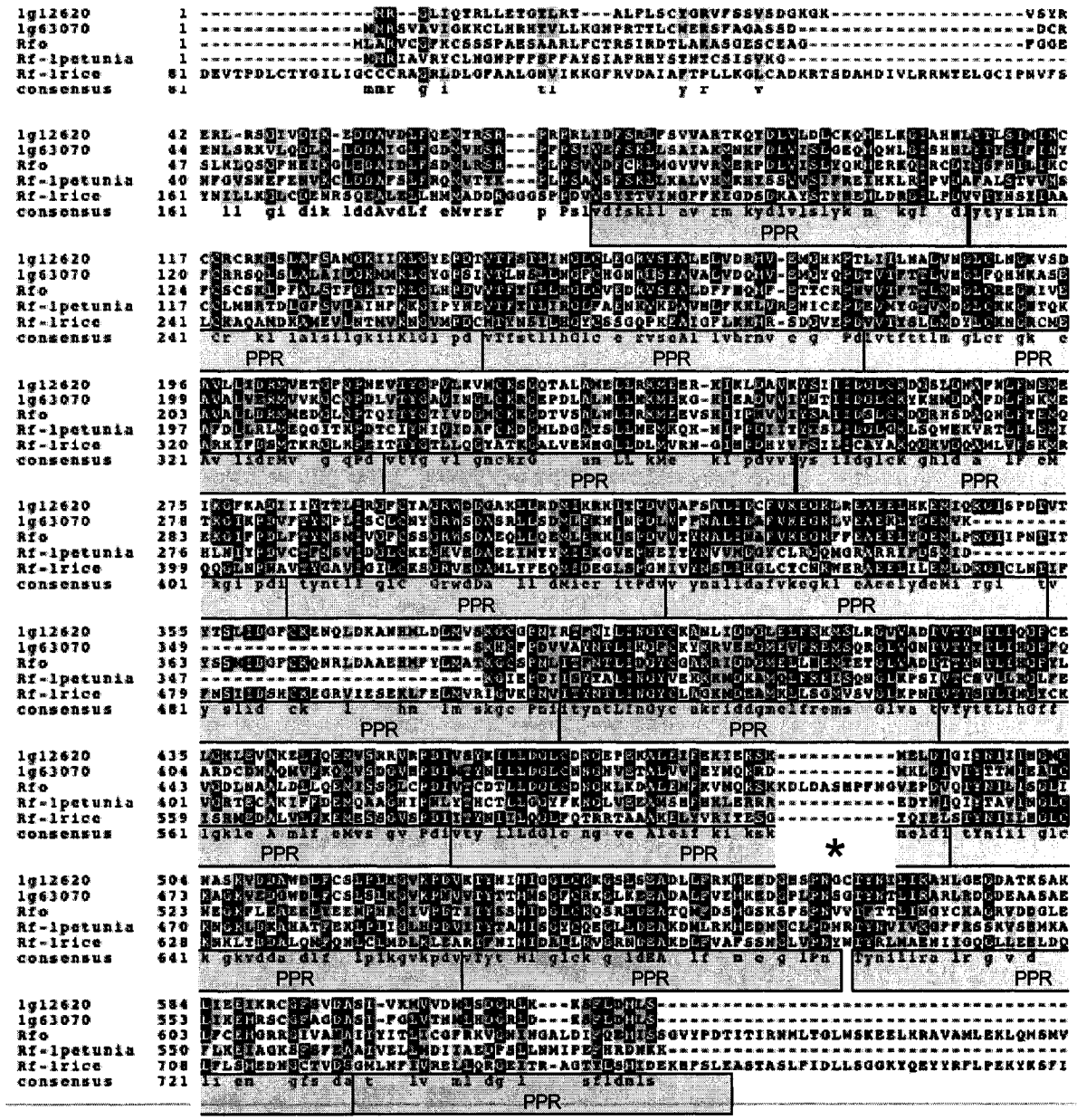
Restorer genes from diverse species have been shown to share significant similarity to the 23 Mb cluster of PPR genes. As shown in Figure 5.6, they are highly structurally related to both the 4.3 and 23 Mb PPR genes. This translates to an overall 49%I/66%+ between restorer genes and related *Arabidopsis* P subfamily PPR genes, a much higher percentage than between restorers and unrelated P and PLS subfamily PPR genes. This evidence taken together indicates that PPR-encoding restorer genes originate from the same subset of P subfamily PPR genes; no PPR-encoding restorer genes have yet been shown to originate from any other subtype of PPR gene.

PPR genes in the radish Rfo region have been subject to diversifying selection

Clustering of PPR genes, such as that seen in the 23 Mb region of chromosome 1 of *Arabidopsis* is a phenomenon also found associated with disease resistance genes (R genes). R genes are subject to diversifying selection that acts on them in a manner that causes duplication and sequence divergence of genes, thus promoting the creation of new or different genes to combat new pathogens (Baumgarten *et al.*, 2003). Brown *et al.* (2003) have shown that the PPR-encoding

Figure 5.6 Alignment of protein coding sequences of PPR encoding restorer of fertility genes with PPR genes of the 4.3 and 23 megabase regions.

Rfo from Ogura radish, *Rf-1* from petunia and *Rf-1* from rice are all highly related to the PPR genes from *Arabidopsis thaliana* chromosome 1 4.3 and 23 megabase regions. They are less highly related to other PPR genes from other subgroups and locations. It is likely that PPR genes which act as restorers all originated from the same progenitor PPR gene. As this alignment shows, the repeat structure among these genes is conserved from sequence to sequence, even between different species. The asterisk (*) indicates a sequence insertion present only in *Rfo*.


lg12620
lg63070
Rfo
Rf-lpetunia
Rf-lrice
consensus
PPR
PPR
PPR
PPR
PPR
PPR
PPR
PPR
PPR
PPR
PPR
PPR
*

restorer gene for Ogura CMS, *Rfo*, forms a mini-cluster of PPR genes in radish with the genes g24 and g27. These three PPR genes are subject to diversifying selection with a rate of non-synonymous nucleotide substitution (Ka) greater than that of synonymous nucleotide substitution (Ks, Table 5.4). Conversely, other non-PPR encoding genes of the same region are under the influence of purifying selection with a greater rate of synonymous nucleotide substitution (Table 5.4). This indicates that PPR genes are under pressure to alter their sequences, thus creating changes that will diversify the population of PPR genes as a whole. This differs from other genes which have a tendency to select against mutation and thus to conserve the sequence of functional proteins. This evidence may lend credence to the hypothesis that PPR genes act as sequence-specific binding proteins, requiring changes in their own sequence to match the sequences they will bind.

Table 5.4 Synonymous and nonsynonymous nucleotide substitution in pairwise comparisons of sequences from radish and *Arabidopsis*.

Where Ka:Ks=1 there is no selective pressure acting on the sequence; where Ka:Ks<1 purifying selection, with a minimization of mutations acts on the sequence; where Ka:Ks>1 diversifying selection, favoring of amino acid change acts on the sequence. *Rfo*, g19, g21, g23, g24 and g27 are all genes from radish, while those starting "At" are genes from *Arabidopsis*.

Genes	Synonymous nucleotide substitutions (Ks)	Nonsynonymous nucleotide substitutions (Ka)	Ka:Ks
Rfo-g24	43	84	1.95
Rfo-g27	48	76	1.58
g24-g27	43	93	2.16
g19-At1g63640	167	123	0.74
g21-At1g63650	111	107	0.96
g23-At1g63680	218	88	0.40

Discussion

Pentatricopeptide repeats (PPR) are structural motifs encoded by a large number of genes in plants and other organisms, although the PPR gene family is greatly expanded in plants. It was hypothesized that this could be due to novel functions served by PPR proteins in plants that are not required in other organisms, or that PPR proteins replace functions performed by other genes in other organisms (Lurin *et al.*, 2004). Recent evidence shows that PPR proteins can function in RNA editing via post-transcriptional conversion of cytoplasmically encoded cytosines to uracil (Kotera *et al.*, 2005), supporting the first hypothesis.

Restoration of male fertility is a plant-specific function encoded by PPR genes. Several recently identified restorers of male fertility in plants encode PPRs and are strongly related to each other at the amino acid level. *Rf1* of petunia (Bentolila *et al.*, 2002), *Rf-1* from rice (Kazama and Toriyama, 2003), and *Rfo* from radish (Brown *et al.*, 2003) are restorers of fertility that encode pentatricopeptide repeats. They are homologous to one another at the amino acid level, sharing about 50% similarity. Since the PPR-encoding restorer genes discovered thus far share sequence similarity, and arise in related gene regions (as is the case with *Rfo* from radish (Desloire *et al.*, 2003)), it seems reasonable to speculate that these genes have arisen from a small number, perhaps even a single, progenitor PPR gene or genes. It is possible that sequences similar to a restorer gene progenitor are located in the 23Mb region of chromosome I, and, in fact, that a progenitor of one these genes may have functioned as a restorer gene at some point in the evolutionary past.

We have found that through comparisons of closely related orthologous sequences as well as comparisons of at least one paralogous region within the *Arabidopsis* genome, the locations of genes encoding PPR domain proteins are highly variable relative to the locations of other types of genes. A consideration of the the most abundant type of plant disease resistance genes (R genes), NBS-LRR genes, may be useful for understanding the mechanisims underlying PPR gene diversity and evolution. PPR genes and NBS-LRR type R genes share many common features. Both types of genes encode proteins with a variable number of repetitive motifs, leucine-rich repeats (LRRs) in the case of NBS-LRR type R genes. In both cases, a single dominant gene determines the phenotype, and, in addition, it is the sequence variability within the repeats that lends specificity of action (Ellis *et al.*, 2000; Michelmore and Meyers, 1998; Richly *et al.*, 2002; Brown *et al.*, 2003).

The evolution and diversity of plant disease resistance genes is a result of tandem and segmental gene duplication, recombination, mutation and natural selection (Meyers *et al.*, 2005). Two sources of gene duplication include local chromosomal rearrangement and large scale genomic duplications (Richter and Ronald, 2000) which is consistent with the conserved synteny model of gene evolution that states that these two mechanisms are the cause of gene distribution and long-distance (ectopic) duplication of genes (Baumgarten *et al.*, 2003). However, most gene duplications are within restricted local chromosomal segments. These local events are the most recent duplications and are most evident when they interrupt the colinearity of gene order in duplicated chromosomal fragments (Baumgarten

et al., 2003). A nonconservative mechanism (i.e. a local change of location) would certainly explain the lack of conservation of synteny of paralogous PPR genes. Lurin *et al.* (2004) suggest that reverse transposition and reintegration could account for the movement of PPRs and their wide distribution among chromosomes. Genomic duplication and chromosomal rearrangement could also account for the widespread distribution of members of the different subtypes of PPR genes among the five chromosomes of *Arabidopsis thaliana.*

Paralogous PPR genes such as those found in the syntenic *Arabidopsis* 4.3 and 23 Mb regions are likely the result of genomic duplications. These duplicated regions have diverged because they are too far away from each other for intergenic exchange to occur (Leister, 2004). Instead, sequence variability and changes in copy number within these regions likely arise through interallelic recombination and diversifying selection (Leister, 2004). Over time, tandem gene duplication can occur as is seen in the radish *Rfo* region. Interestingly, this tandem duplication is not evident in *Arabidopsis* PPR gene distribution. PPR genes are for the most part found as singlets with PPR gene subgroup members evenly distributed in the *Arabidopsis* genome amongst the five chromosomes, whereas R genes are found more often in clusters of related genes (Meyers *et al.*, 2003). This may be a result of the large diversifying selective pressure exerted on disease resistance loci as plants are continually adapting to new plant pathogens (Baumgarten *et al.*, 2003).

There is only one defined cluster of PPR genes in *Arabidopsis thaliana*; it is the cluster related to restorer genes of rice, petunia and radish (Lurin *et al.*, 2004). It is

possible that this clustering indicates diversifying selection acting on PPR genes from that region as a result of plants adapting to newly emerging sterility inducing genes. The diversifying selective pressure exerted on the mini-cluster of PPR genes at the radish *Rfo* locus is one example of this effect acting on PPR genes and not on other genes of the same region. Again, the PPR genes of the *Rfo* region are out of synteny with the PPR genes of *Arabidopsis*, while the non-PPR encoding gene locations are conserved. Diversifying selection can explain why non-PPR genes are not apt to fall out of synteny with their paralogous partners. If diversifying selection is not acting on a gene then any changes are more likely to be conservative, with synonymous substitution outweighing nonsynonymous substitution and little selection for location changes (Parniske *et al.*, 1997; Ellis *et al.*, 2000).

The presence of so-called false PPR genes (Lurin *et al.*, 2004) follows the Birth-and-Death model adopted by Michelmore and Meyers (1998) for R genes, which indicates that following gene duplication due to diversifying selection some members have become redundant; mutations which cause frameshifts or premature stop codons in the coding sequences of these genes have had their function disabled. It has been noted that PPR genes contain, on average, many fewer introns than other *Arabidopsis* genes, thus increasing the likelihood that mutations will affect coding regions (Lurin *et al.*, 2004).

The potential wide ranging and diverse utility of PPR genes, in conjunction with their nomadic nature, indicates that they are a versatile gene family akin to disease resistance genes. The common features exhibited by both types of genes

has led to the hypothesis that PPR genes may function in a similar manner to R genes, that is as malleable proteins capable of alteration in response to changing environmental pressures. Our work has shown that PPR genes display the features necessary for them to be a diverse, general use family, and future work will bring to light the manner in which PPR genes achieve their specificity of action.

References

Akagi, H., Nakamura, A., Yokozekik-Misono, Y., Inagaki, A., Takahashi, H., Mori, K. and Fujimura, T. (2004) Positional cloning of the rice *Rf-1* gene, a restorer of BT-type cytoplasmic male sterility that encodes a mitochondria-targeting PPR protein. Theor. Appl. Genet. **108**, 1449–1457.

Baumgarten, A., Cannon, S., Spangler, R. and May, G. (2003) Genome-level evolution of resistance genes in *Arabidopsis thaliana*. Genetics **165**, 309-319.

Bentolila, S., Alfonso, A. and Hanson, M.R. (2002) A pentatricopeptide repeat-containing gene restores fertility to cytoplasmic male-sterile plants. PNAS **99**, 10887-10892.

Blanc, G. Barakat, A., Guyot, R., Cooke, R. and Delseny, M. (2000) Extensive duplication and reshuffling in the *Arabidopsis* genome. Plant Cell **12**, 1093-1101.

Brown, G.G., Formanova, N., Jin, H., Wargachuk, R., Dendy, C., Patil, P., Laforest, M., Zhang, J., Cheung, W.Y. and Landry, B.S. (2003) The radish *Rfo* restorer gene of Oguar cytoplasmic male sterility encodes a protein with multiple pentatricopeptide repeats. Plant J. **35**, 262-272.

Cannon, S.B., Mitra, A., Baumgarten, A., Young, N.D. and May, G. (2004) The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. BMC Plant Biol. **4**:10.

Claros, M.G. and Vincens, P. (1996) Computational method to predict mitochondrially imported proteins and their targeting sequences. Eur. J. Biochem. **241**, 770-786.

Cushing, D.A., Forsthoefel, N.R., Gestaut, D.R. and Vernon, D.M. (2005) *Arabidopsis emb*175 and other *ppr* knockout mutants reveal essential roles for pentatricopeptide repeat (PPR) proteins in plant embryogenesis. Planta **221**, 424-436.

Desloire, S., Gherbi, H., Laloui, W., Marhadour, S., Clouet, V., Cattolico, L., Faletin, C., Giancola, S., Renard, M., Budar, F., Small, I, Caboch, M., Delourme, R. and Bendahmane, A. (2003) Identification ofthe fertility restoration locus, *Rfo*, in radish, as a member of the pentatrico-peptiderepeat protein family. EMBO Rep. **4**, 588 -594.

Ellis, J., Dodds, P. & Pryor, T. (2000) The generation of plant disease resistance gene specificities. Trends in Plant Science **5**, 373-379.

International rice genome sequencing project (2005) The map-based sequence of the rice genome. Nature **436**, 793-800.

Kazama, T. and Toriyama, K. (2003) A pentatricopeptide repeat-containing gene that promotes the processing of aberrant *atp6* RNA of cytoplasmic male-sterile rice. FEBS Letters **544**, 99-102.

Komori, T., Ohta, S., Murai, N., Takakura, Y, Kuraya, Y, Suzuki, S., Hiei, Y., Imaseki, H. and Nitta, N. (2004) Map-based cloning of a fertility restorer gene, *Rf-1*, in rice (*Oryza sativa* L.). Plant J. **37**, 315-325.

Kotera, E., Tasaka, M. and Shikanai, T. (2005) A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. Nature 433, 326-330.

Leister, D. (2004) Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance genes. Trends Genet. **20**, 116-122.

Lukens, L. , Zou, F., Lydiate, D., Parkin, I. and Osborn, T. (2003) Comparison of a Brassica oleracea Genetic Map With the Genome of *Arabidopsis thaliana.* Genetics **164**, 359–372.

Lurin, C., Andre, C., Aubourg, S., Bellaoui, M., Betton, F., Bruyere, C., Caboche, M., Debast, C., Gualberto, J., Hoffmann, B., Lecharny, A., LeRet, M., Martin-Magniette, M.-L., Mireau, H., Peeters., N., Renou, J.-P., Szurek, B., Taconnat, L. and Small, I. (2004) Genome-wide analysis of *Arabidopsis* pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. Plant Cell **16**, 2089-2103.

Marchler-Bauer, A. and Bryant, S.H. (2004) CD-Search: protein domain annotations on the fly. Nucleic Acids Research **32**,W327-331.

Meyers, B.C., Kozik, a., Griego, A., Kuang, H. and Michelmore, R.W. (2003). Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. Plant Cell **15**, 809-834.

Meyers, B.C., Kaushik, S. and Nandety, R.S. (2005) Evolving disease resistance genes. Current Opinion in Plant Biology **8**, 129-134.

Michelmore, R.W. and Meyers, B.C. (1998) Clusters of resistance genes in plants evolve by divergent selection and birth-and-death process. Genome Research **8**, 1113-1130.

Nakamura, T., Schuster, G., Sugiura, M. and Sugita, M. (2004) Chloroplast RNA-binding and pentatricopeptide repeat proteins. Biochem. Soc. Trans. **32**, 571-574.

Nei, M., Gu, X. and Sitnikova, T. (1997) Evolution by the birth-and-death process in multigene families of the vertebrate immune system. PNAS **94**, 7799-7806.

O'Neill, C.M. and Bancroft, I. (2000) Comparative physical mapping of segments of the genome of *Brassica oleracea* var. alboglabra that are homoeologous to sequenced regions of chromosomes 4 and 5 of *Arabidopsis thaliana*. Plant J. **23**, 233-43.

Parkin, I.A., Gulden, S.M., Sharpe, A.G., Lukens, L., Trick, M., Osborn, T.C. and Lydiate, D.J. (2005) Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. Genetics **171**, 765-781.

Parniske, M. Hammond-Kosack, K.E., Golstein, C., Thomas, C.M. and Jones, D.A. (1997) Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the *Cf-4/9* locus of tomato. Cell **91**, 821-832.

Richly, R., Kurth, J. and Leister, D. (2002) Mode of amplification and reorganization of resistance genes during recent *Arabidopsis thaliana* evolution. Mol. Biol. Evol. **19**, 76-84.

Richter, T.E. and Ronald, P.C. (2000) The evolution of disease resistance genes. Plant Mol. Biol. **42**, 195-204.

Small, I.D. and Peeters, N. (2000) The PPR motif a TPR-related motif prevalent in plant organellar proteins. Trends Biochem. Sci. **25**, 4647.

Small, I., Peeters, N., Legeai, F. and Lurin, C. (2004) Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. Proteomics, **4**, 1581-1590.

Zielkowski, P.A., Blanc, G. and Sadowski, J. (2003) Structural divergence of chromosomal segments that arose from successive duplication events in the *Arabidopsis* genome. Nucleic acids research **31**, 1339-1350.

GENERAL CONCLUSIONS AND SUMMARY

In *Brassica napus*, Cytoplasmic Male Sterility (CMS) results in the loss of locule development in the anther as well as effects such as malformation of the petals and loss of synchronicity of anther development. In *nap* CMS, these defects are associated, with the expression of *orf222* in locules while restoration of fertility perturbs the transcript expression pattern resulting in transcript expression only outside of the locules. The pattern of *orf222* transcript accumulation in CMS anthers matches that of *atp6*, cytochrome *b* and ATP synthase gamma subunit in both CMS and fertility-restored plants. Perturbation of *orf222* transcripts in fertility restoration, therefore, suggests that *orf222* may have a role in interference with electron transport chain function. It's physical effects being restricted to whorls two and three of the flower bud, together with evidence proved by *AP3* promoter fusion to *orf222*, also indicates downstream effects of *orf222* in homeotic gene expression and, ultimately, floral patterning. Regulation of floral development may also be tied to the accumulation of antisense transcripts of some elements of the electron transport chain.

Restoration of fertility has been associated in other plant systems with genes belonging to the pentatricopeptide repeat (PPR) family. Mapping of *Rfp*, the restorer of polima CMS, has so far placed it in a genomic location syntenic to 105 kb of the 4.3 Mb region of *Arabidopsis thaliana* chromosome 1. This is a duplicated region of the 23 Mb mark of chromosome 1. This region contains PPRs most highly homologous to restorer genes of rice, petunia, and radish CMS. By

inference, it is possible that *Rfp* also encodes a PPR gene and that it will be homologous to other restorer genes discovered thus far.

PPR genes of restorer genes in various plant species are highly homologous to one another. When PPR genes and flanking genes of radish and *Brassica* are compared to orthologous genomic sequences in *Arabidopsis* synteny is maintained except for PPR genes which are asyntenically located. In addition, comparison of paralogous sequences within both *Brassica* and *Arabidopsis* also reveals asynteny of PPR genes. Expansion of the PPR gene family, as well as frequent location change and the pressure of diversifying selection acting on PPR genes are all attributes shared with disease resistance genes in plants. PPR genes may, therefore, act as malleable proteins capable of being altered in response to changing environmental pressures, such as the evolution of sterility inducing genes.

ORIGINAL SCHOLARSHIP AND DISTINCT CONTRIBUTIONS TO KNOWLEDGE

Comparison of developmental changes that occur in anther development between fertile and cytoplasmic male sterile *Brassica napus* floral buds including:

- malformation of sterile floral bud shape;
- polar loss of locule development in CMS anthers;
- loss of synchronous development of locules both within and between anthers;
- locule fusion in CMS anthers;
- tubular and thickened petal development in CMS flower buds;
- clumping of sporogenous tissue away from locule walls in CMS anthers.

Demonstration of the patterns of transcript accumulation of CMS-inducing gene *orf222* throughout anther development in both fertile and cytoplasmic male sterile *Brassica napus* floral buds including:

- early accumulation of *orf222* transcripts in CMS floral buds;
- accumulation of *orf222* transcripts in locules of CMS floral buds;
- accumulation of *orf222* transcripts in anther tissue with the exception of locules in fertility-restored buds;
- harmonization of *orf222* transcript accumulation between CMS and fertility restored anthers in the post-microspore mother cell stages.

Demonstration of the patterns of transcript accumulation of *atp6*, cytochrome *b* and ATP synthase gamma subunit in both fertile and cytoplasmic male sterile *Brassica napus* floral buds including:

- accumulation of transcripts of these genes in both CMS and fertility restored anthers in a pattern matching that of *orf222* transcripts in CMS anthers;

- accumulation of antisense transcripts of *atp6* and ATP synthase gamma in a pattern matching that of sense transcripts.

Reduction of the interval in which *Rfp* is found on one side by isolation of an anchoring fragment of DNA bearing a SNP marker.

Demonstration that AFLP is an effective technique for distinguishing paralogous sequences from the cosmid library of *Brassica rapa* introgressed with *Rfp*.

Discovery that PPR genes are not located in fixed syntenic locations in orthologous and paralogous comparisons where synteny is maintained among non-PPR genes.

Discovery that a segment of DNA at the 4.3 Mb mark of chromosome 1 of *Arabidopsis thaliana* is a duplication of a chromosomal segment located at the 23 Mb mark of the same chromosome. In addition:

- PPR genes of these two DNA segments are highly homologous to one another;
- these PPR genes show higher homology to each other than to other PPR genes of *Arabidopsis thaliana;*
- the PPR genes of this duplication are syntenically located, although synteny is maintained for non-PPR genes.

Determination that PPR genes are subject to diversifying selection, a phenomenon leading to expansion and diversification of the gene family.

McGill University Internal Radioisotope Permit

20000154

Permit Holder & Position	Building (Office)	Building (Lab) / Laboratory Classification		
GREGORY G. BROWN, PROFESSOR	STEWART BIOLOGY BUILDING	STEWART BIOLOGY BUILDING	**PLEASE POST**	Date Issued 2005/09/01
Department: BIOLOGY	Room Number(s): N5/5 — Telephone: 398-6427	Room Number(s): N5/4 — Telephone: 398-6427		Expiry Date 2006/08/31

PERSON(S) APPROVED TO WORK WITH RADIOISOTOPES

Name	Train.	Cond(s)	Class(es)	Radioisotope(s)
NATASA FORMANOVA, VISITING SCIENTIST	N	2, 4	1	P-32, P-33
GREGORY G. BROWN, PROFESSOR	N	2, 4	1	P-32
HELEN ELINA, GRAD. STUDENT	N	2, 4,	3	P-32, S-35
BENJAMIN STOREY, GRAD. STUDENT	N	2, 4	1	P-32,
LATITIA MAHE, GRADUATE STUDENT	N	2, 4	1	P-32, S-35
RACHEL GEDDY, GRADUATE STUDENT	N	2, 4	1	P-32
PRACHANT PATIL, VISITING SCIENTIST	N	2, 4	1	P-32

GENERAL LICENCE CONDITIONS (OPEN AND/OR SEALED SOURCES)

- The permit must be posted with the CNSC safety poster in the permit holder's premises.
- Radioisotope handling shall be in accordance with the McGill Radiation Safety Policy Manual.
- The permit holder must ensure that all persons mandated to work with radioisotopes be properly trained in radiation safety prior to start of work.
- Radioactive work areas must be clearly identified with radiation warning signs.
- Smoking, eating, drinking, storage of foods or drink and the application of cosmetics and contact lenses are prohibited in areas where radioisotopes are used.
- All procedures involving radioactive materials should be carried out on spill trays or on benches lined with disposable absorbent material.
- Procedures that might produce airborne radioactive contamination should be carried out in a functioning fume hood.
- When hand or clothing contamination is possible, protective gloves and clothing must be worn.
- After handling radioactive material and especially before leaving the laboratory, personnel must ensure that all parts of their cloths are not contaminated.
- Purchase and disposal of radioisotopes must be kept electronically or documented in a log book.
- For disposal of radioactive waste, consult the McGill Radiation Safety Policy Manual and/or McGill WMP.
- Wipe tests must be performed and records be kept in a log book.
- The permit must reflect the exact conditions under which radioactive material is used. If changes must be made, contact the RSO at 398-1538.
- The device(s) containing the sealed source(s) must have a radiation symbol and an identification label bearing the name and telephone number of the permit holder.
- Leak tests must be performed on sealed sources equal to or greater than 50 MBq (1.35mCi).
- Extremity dosimeter (i.e. ring or wrist badges) must be worn if 50 MBq or more of P-32, Sr-89, Sr-90 & Y-90 are used.
- Workers using I-125 or I-131 on open bench (5 MBq), in a fume hood (50 MBq) or a vented glove

Approved Unsealed Radioisotope(s) and Location(s)

Isotope	Possession Limit	Stored	Handled
H-3	< 400 MBq (11 mCi)	N5/4	N5/4
P-32	< 400 MBq (11 mCi)	N5/4	N5/4
S-35	< 400 MBq (11 mCi)	N5/4	N5/4
P-33	< 100 MBq (2.7 mCi)	N5/4	N5/4

Approved Sealed Radioisotope(s) and Location(s)

Permanently Housed Source(s)				*Accessible Source(s)*			
Isotope	Activity	Stored	Handled	Isotope	Activity	Stored	Handled

Personnel Conditions

1 .Must attend thyroid bioassays within 5 days of use if 50 MBq (1.35 mCi) of I-125 are manipulated in a fume hood.
2 .Must wear a whole-body film badge, if gamma, x-ray or high energy beta emitters are used.
3 .Must wear an extremity TLD dosimeter, if more than 50 MBz (1.35 mCi) of P-32, Sr-89, Sr-90 or Y-90 are used.
4 .Classified as Radiation User.
5 .Classified as Nuclear Energy Worker (NEW).
6 .Does not work with any radioisotopes but may be indirectly exposed.

Workload Classes

1 .Work load < 10MBq (270 uCi) of unsealed radioisotopes in open areas.
2 .Work load < 10MBq (270 uCi) of unsealed radioisotopes in a fume hood.
3 .Work load > 10MBq (270 uCi) of unsealed radioisotopes in open areas.
4 .Work load > 10 MBq (270 uCi) of unsealed radioisotopes in a fume hood.
5 .Work with sealed sources.
6 .Individual does not work with radioactive sources but normal working conditions involve presence in a room where radioactive material is used or stored.

Joseph Vincelli
RSO & Occupational Hygienist
McGill Environmental Health & Safety

Date

For: Dr. Janusz Kozinski, Chair
McGill University Laboratory Safety Committee
Associate Vice-Principal (Research)

McGill University Internal Radioisotope Permit

20010154

Permit Holder & Position	Building (Office)	Building (Lab) / Laboratory Classification		Date Issued
GREGORY G. BROWN, PROFESSOR	STEWART BIOLOGY BUILDING	STEWART BIOLOGY BUILDING	**PLEASE POST**	2005/09/01
Department	**Room Number(s) / Telephone**	**Room Number(s) / Telephone**		**Expiry Date**
BIOLOGY	N5/5 398-6427	N5/4 398-6427		2006/08/31

PERSON(S) APPROVED TO WORK WITH RADIOISOTOPES

Name	Train.	Cond(s)	Class(es)	Radioisotope(s)
NATASA FORMANOVA, RESEARCH ASSISTANT	N	2, 4	1	P-32, P-33
GREGORY G. BROWN, PROFESSOR	N	6	6	
HELEN ELINA, GRAD. STUDENT	N	2, 4,	3	P-32, S-35
BENJAMIN STOREY, GRAD. STUDENT	N	2, 4	1	P-32,, S-35
LATITIA MAHE, GRADUATE STUDENT	N	2, 4	1	P-32, S-35
RACHEL GEDDY, GRADUATE STUDENT	N	2, 4	3	P-32, S-35
PRACHANT PATIL, VISITING SCIENTIST	N	2, 4	1	P-32
RACHEL STOLLAR, GRADUATE STUDENT	N	2, 4	1	P-33, P-32
RICHARD WARGACHUK, GRADUATE STUDENT	N	2, 4	1	P-32, S-35
JINFA ZHANG, POST DOC	N	2, 4	1	P-32, S-35
PHILOMENA KAMYA, RESEARCH ASSISTANT	N	2, 4	1	P-32

GENERAL LICENCE CONDITIONS (OPEN AND/OR SEALED SOURCES)

- The permit must be posted with the CNSC safety poster in the permit holder's premises.
- Radioisotope handling shall be in accordance with the McGill Radiation Safety Policy Manual.
- The permit holder must ensure that all persons mandated to work with radioisotopes be properly trained in radiation safety prior to start of work.
- Radioactive work areas must be clearly identified with radiation warning signs.
- Smoking, eating, drinking, storage of foods or drink and the application of cosmetics and contact lenses are prohibited in areas where radioisotopes are used.
- All procedures involving radioactive materials should be carried out on spill trays or on benches lined with disposable absorbent material.
- Procedures that might produce airborne radioactive contamination should be carried out in a functioning fume hood.
- When hand or clothing contamination is possible, protective gloves and clothing must be worn.
- After handling radioactive material and especially before leaving the laboratory, personnel must ensure that all parts of their cloths are not contaminated.
- Purchase and disposal of radioisotopes must be kept electronically or documented in a log book.
- For disposal of radioactive waste, consult the McGill Radiation Safety Policy Manual and/or McGill WMP.
- Wipe tests must be performed and records be kept in a log book.
- The permit must reflect the exact conditions under which radioactive material is used. If changes must be made, contact the RSO at 398-1538.
- The device(s) containing the sealed source(s) must have a radiation symbol and an identification label bearing the name and telephone number of the permit holder.
- Leak tests must be performed on sealed sources equal to or greater than 50 MBq (1.35mCi).
- Extremity dosimeter (i.e. ring or wrist badges) must be worn if 50 MBq or more of P-32, Sr-89, Sr-90 & Y-90 are used.
- Workers using I-125 or I-131 on open bench (5 MBq), in a fume hood (50 MBq) or a vented glove

Approved Unsealed Radioisotope(s) and Location(s)

Isotope	Possession Limit	Stored	Handled
H-3	< 400 MBq (11 mCi)	N5/4	N5/4
P-32	< 400 MBq (11 mCi)	N5/4	N5/4
S-35	< 400 MBq (11 mCi)	N5/4	N5/4
P-33	< 100 MBq (2.7 mCi)	N5/4	N5/4

Approved Sealed Radioisotope(s) and Location(s)

Permanently Housed Source(s)				*Accessible Source(s)*			
Isotope	Activity	Stored	Handled	Isotope	Activity	Stored	Handled

Personnel Conditions

1 .Must attend thyroid bioassays within 5 days of use if 50 MBq (1.35 mCi) of I-125 are manipulated in a fume hood.
2 .Must wear a whole-body film badge, if gamma, x-ray or high energy beta emitters are used.
3 .Must wear an extremity TLD dosimeter, if more than 50 MBz (1.35 mCi) of P-32, Sr-89, Sr-90 or Y-90 are used.
4 .Classified as Radiation User.
5 .Classified as Nuclear Energy Worker (NEW).
6 .Does not work with any radioisotopes but may be indirectly exposed.

Workload Classes

1 .Work load < 10MBq (270 uCi) of unsealed radioisotopes in open areas.
2 .Work load < 10MBq (270 uCi) of unsealed radioisotopes in a fume hood.
3 .Work load > 10MBq (270 uCi) of unsealed radioisotopes in open areas.
4 .Work load > 10 MBq (270 uCi) of unsealed radioisotopes in a fume hood.
5 .Work with sealed sources.
6 .Individual does not work with radioactive sources but normal working conditions involve presence in a room where radioactive material is used or stored.

Joseph Vincelli
RSO & Occupational Hygienist
McGill Environmental Health & Safety

Date

For: Dr. Janusz Kozinski, Chair
McGill University Laboratory Safety Committee
Associate Vice-Principal (Research)

McGill University Internal Radioisotope Permit

20020154

Permit Holder & Position	Building (Office)	Building (Lab)	Laboratory Classification		Date Issued
GREGORY G. BROWN, PROFESSOR	STEWART BIOLOGY BUILDING	STEWART BIOLOGY BUILDING	BASIC	**PLEASE POST**	2005/09/01
Department	Room Number(s) / Telephone	Room Number(s) / Telephone			Expiry Date
BIOLOGY	N5/5 398-6426	N5/4 398-6427			2006/08/31

PERSON(S) APPROVED TO WORK WITH RADIOISOTOPES

Name	Train.	Cond(s)	Class(es)	Radioisotope(s)
NATASA FORMANOVA, RESEARCH ASSISTANT	N	2, 4	1	P-32, P-33
GREGORY G. BROWN, PROFESSOR	N	6	6	
HELEN ELINA, GRAD. STUDENT	Y	2, 4,	3	P-32, S-35
BENJAMIN STOREY, GRAD. STUDENT	N	2, 4	1	P-32,, S-35
LATITIA MAHE, GRADUATE STUDENT	Y	2, 4	1	P-32, S-35
RACHEL GEDDY, GRADUATE STUDENT	Y	2, 4	3	P-32, S-35
RICHARD WARGACHUK, GRADUATE STUDENT	Y	2, 4	1	P-32, S-35
RACHEL STOLLAR, GRADUATE STUDENT	N	2, 4	1	P-32, P-33

GENERAL LICENCE CONDITIONS (OPEN AND/OR SEALED SOURCES)

- The permit must be posted with the CNSC safety poster in the permit holder's premises.
- Radioisotope handling shall be in accordance with the McGill Radiation Safety Policy Manual.
- The permit holder must ensure that all persons mandated to work with radioisotopes be properly trained in radiation safety prior to start of work.
- Radioactive work areas must be clearly identified with radiation warning signs.
- Smoking, eating, drinking, storage of foods or drink and the application of cosmetics and contact lenses are prohibited in areas where radioisotopes are used.
- All procedures involving radioactive materials should be carried out on spill trays or on benches lined with disposable absorbent material.
- Procedures that might produce airborne radioactive contamination should be carried out in a functioning fume hood.
- When hand or clothing contamination is possible, protective gloves and clothing must be worn.
- After handling radioactive material and especially before leaving the laboratory, personnel must ensure that all parts of their cloths are not contaminated.
- Purchase and disposal of radioisotopes must be kept electronically or documented in a log book.
- For disposal of radioactive waste, consult the McGill Radiation Safety Policy Manual and/or McGill WMP.
- Wipe tests must be performed and records be kept in a log book.
- The permit must reflect the exact conditions under which radioactive material is used. If changes must be made, contact the RSO at 398-1538.
- The device(s) containing the sealed source(s) must have a radiation symbol and an identification label bearing the name and telephone number of the permit holder.
- Leak tests must be performed on sealed sources equal to or greater than 50 MBq (1.35mCi).
- Extremity dosimeter (i.e. ring or wrist badges) must be worn if 50 MBq or more of P-32, Sr-89, Sr-90 & Y-90 are used.
- Workers using I-125 or I-131 on open bench (5 MBq), in a fume hood (50 MBq) or a vented glove

Approved Unsealed Radioisotope(s) and Location(s)

Isotope	Possession Limit	Stored	Handled
P-32	< 400 MBq (11 mCi)	N5/4	N5/4
S-35	< 400 MBq (11 mCi)	N5/4	N5/4
P-33	< 100 MBq (2.7 mCi)	N5/4	N5/4

Approved Sealed Radioisotope(s) and Location(s)

Permanently Housed Source(s)				*Accessible Source(s)*			
Isotope	Activity	Stored	Handled	Isotope	Activity	Stored	Handled

Personnel Conditions

1 .Must attend thyroid bioassays within 5 days of use if 50 MBq (1.35 mCi) of I-125 are manipulated in a fume hood.
2 .Must wear a whole-body film badge, if gamma, x-ray or high energy beta emitters are used.
3 .Must wear an extremity TLD dosimeter, if more than 50 MBz (1.35 mCi) of P-32, Sr-89, Sr-90 or Y-90 are used.
4 .Classified as Radiation User.
5 .Classified as Nuclear Energy Worker (NEW).
6 .Does not work with any radioisotopes but may be indirectly exposed.

Workload Classes

1 .Work load < 10MBq (270 uCi) of unsealed radioisotopes in open areas.
2 .Work load < 10MBq (270 uCi) of unsealed radioisotopes in a fume hood.
3 .Work load > 10MBq (270 uCi) of unsealed radioisotopes in open areas.
4 .Work load > 10 MBq (270 uCi) of unsealed radioisotopes in a fume hood.
5 .Work with sealed sources.
6 .Individual does not work with radioactive sources but normal working conditions involve presence in a room where radioactive material is used or stored.

Joseph Vincelli
RSO & Occupational Hygienist
McGill Environmental Health & Safety

Date

For: Dr. Janusz Kozinski, Chair
McGill University Laboratory Safety Committee
Associate Vice-Principal (Research)

McGill University Internal Radioisotope Permit

20030154

Permit Holder & Position	Building (Office)	Building (Lab)	Laboratory Classification		Date Issued
GREGORY G. BROWN, PROFESSOR	STEWART BIOLOGY BUILDING	STEWART BIOLOGY BUILDING	INTERMEDIATE	**PLEASE POST**	2005/09/01
Department	Room Number(s) / Telephone	Room Number(s) / Telephone			Expiry Date
BIOLOGY	N5/5 398-6426	N5/4 398-6427			2006/08/31

PERSON(S) APPROVED TO WORK WITH RADIOISOTOPES

Name	Train.	Cond(s)	Class(es)	Radioisotope(s)
GREGORY G. BROWN, PROFESSOR	N	6	6	
HELEN ELINA, GRAD. STUDENT	Y	2, 4,	3	P-32, S-35
RACHEL GEDDY, GRADUATE STUDENT	Y	2, 4	3	P-32, S-35
RICHARD WARGACHUK, GRADUATE STUDENT	Y	2, 4	1	P-32, S-35

GENERAL LICENCE CONDITIONS (OPEN AND/OR SEALED SOURCES)

- The permit must be posted with the CNSC safety poster in the permit holder's premises.
- Radioisotope handling shall be in accordance with the McGill Radiation Safety Policy Manual.
- The permit holder must ensure that all persons mandated to work with radioisotopes be properly trained in radiation safety prior to start of work.
- Radioactive work areas must be clearly identified with radiation warning signs.
- Smoking, eating, drinking, storage of foods or drink and the application of cosmetics and contact lenses are prohibited in areas where radioisotopes are used.
- All procedures involving radioactive materials should be carried out on spill trays or on benches lined with disposable absorbent material.
- Procedures that might produce airborne radioactive contamination should be carried out in a functioning fume hood.
- When hand or clothing contamination is possible, protective gloves and clothing must be worn.
- After handling radioactive material and especially before leaving the laboratory, personnel must ensure that all parts of their cloths are not contaminated.
- Purchase and disposal of radioisotopes must be kept electronically or documented in a log book.
- For disposal of radioactive waste, consult the McGill Radiation Safety Policy Manual and/or McGill WMP.
- Wipe tests must be performed and records be kept in a log book.
- The permit must reflect the exact conditions under which radioactive material is used. If changes must be made, contact the RSO at 398-1538.
- The device(s) containing the sealed source(s) must have a radiation symbol and an identification label bearing the name and telephone number of the permit holder.
- Leak tests must be performed on sealed sources equal to or greater than 50 MBq (1.35mCi).
- Extremity dosimeter (i.e. ring or wrist badges) must be worn if 50 MBq or more of P-32, Sr-89, Sr-90 & Y-90 are used.
- Workers using I-125 or I-131 on open bench (5 MBq), in a fume hood (50 MBq) or a vented glove

Approved Unsealed Radioisotope(s) and Location(s)

Isotope	Possession Limit	Stored	Handled
P-32	< 400 MBq (11 mCi)	N5/4	N5/4
S-35	< 400 MBq (11 mCi)	N5/4	N5/4

Approved Sealed Radioisotope(s) and Location(s)

Permanently Housed Source(s)				*Accessible Source(s)*			
Isotope	Activity	Stored	Handled	Isotope	Activity	Stored	Handled

Personnel Conditions

1 .Must attend thyroid bioassays within 5 days of use if 50 MBq (1.35 mCi) of I-125 are manipulated in a fume hood.
2 .Must wear a whole-body film badge, if gamma, x-ray or high energy beta emitters are used.
3 .Must wear an extremity TLD dosimeter, if more than 50 MBz (1.35 mCi) of P-32, Sr-89, Sr-90 or Y-90 are used.
4 .Classified as Radiation User.
5 .Classified as Nuclear Energy Worker (NEW).
6 .Does not work with any radioisotopes but may be indirectly exposed.

Workload Classes

1 .Work load < 10MBq (270 uCi) of unsealed radioisotopes in open areas.
2 .Work load < 10MBq (270 uCi) of unsealed radioisotopes in a fume hood.
3 .Work load > 10MBq (270 uCi) of unsealed radioisotopes in open areas.
4 .Work load > 10 MBq (270 uCi) of unsealed radioisotopes in a fume hood.
5 .Work with sealed sources.
6 .Individual does not work with radioactive sources but normal working conditions involve presence in a room where radioactive material is used or stored.

Joseph Vincelli
RSO & Occupational Hygienist
McGill Environmental Health & Safety

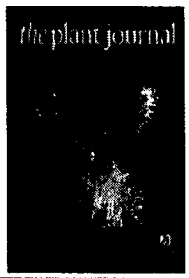
Date

For: Dr. Janusz Kozinski, Chair
McGill University Laboratory Safety Committee
Associate Vice-Principal (Research)

Home
ls Home
Journals
ed Search
es
Alerts
dex
t Index
rarians
Room
ant journal
l Home
Scope
ibe/Renew
al Board
of Contents
Sample Issue
ation / Society
ership
thors
and Services
ner Services
sions
its
ising
t Sites
ell Plant
e

the plant journal

The Plant Journal



Published in association with the Society for Experimental Biology

Edited by:
Harry Klee

ISI Journal Citation Reports® Ranking: 2004: 6/138 (Plant Sciences)
Impact Factor: 6.367

View content online
Sign up for e-tocs
Recommend to your library
Submit an article

Permissions and Copyright Information

Reprinting

Requests for permission to reprint material from this journal should be addressed to:

Journals Rights & Permissions Controller
Blackwell Publishing Ltd
9600 Garsington Road
Oxford OX4 2DQ
UK
Email: JournalsRights@oxon.blackwellpublishing.com

RECEIVED
FEB 0 2 2006
GPSO

Please be sure to include the following information:

- The title, author(s) and page extent of the article you wish to republish, or, in the case where you do not wish to reproduce the whole article, details of which section(s), with a word count where relevant.
- The journal title, volume/issue number and year of publication in which the article appeared, and the page numbers of the article.
- The title, editor(s)/author(s), price, print run, publication date and publisher of the publication in which you wish to reprint the journal material.
- If the material is being edited or amended in any way, a copy of the final material as it will appear in your publication.
- Contact details for yourself, including a postal address.

If you wish to include your own article (or an amended version of it) in a volume of which you are the editor or co-editor, prior permission is not required, but the usual acknowledgements should be made.

Photocopying

Teaching institutions with a current paid subscription to the journal may make multiple copies for teaching purposes without charge, provided such copies are not resold. In all other cases, permission should be obtained from a reproduction rights organisation (see below) or directly from the Publisher.

Permission to photocopy print versions of this journal varies by territory.

USA: Please obtain permission via the CCC:

Copyright Clearance Center
222 Rosewood Drive
Danvers, MA 01923
USA
Tel: +1(978) 750 8400
Fax: +1 (978) 750 4744

UK: This journal is registered with the Copyright Licensing Agency, and licensed institutions may make copies under the terms of their licence.

For multiple copies and coursepacks permission should be cleared through CLARCS, the CLA Rapid Clearance Service:

Education, tel: 020 7631 5560 Business and Government, tel: 020 7631 5561; email: