Fine mapping and functional analysis of the radish *Rfo* nuclear restorer locus

Richard Burns Wargachuk
Department of Biology
McGill University, Montreal

A thesis submitted to McGill University in partial fulfilment of the requirements of the degree of Master of Science.

Your file Votre référence
ISBN: 0-494-06470-6
Our file Notre référence
ISBN: 0-494-06470-6

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Abstract

Cytoplasmic male sterility (CMS) is a widespread, maternally inherited trait that results in an inability of plants to produce functional pollen. The Ogura CMS system originated in radish, but has since been transferred to, and confers male sterility on, plants in the related genus *Brassica*. A gene which restores male fertility is needed for the Ogura CMS system to be exploited commercially for hybrid seed production in oilseed species such as *Brassica napus*. The restorer gene *Rfo* is a dominant radish nuclear gene that restores the male fertility to plants with Ogura cytoplasm. This gene has been transferred into *Brassica* napus through intergeneric crosses; however the introgressed segment of radish DNA contains an unknown number of genes, some of which confer undesirable traits, such as an elevated content of seed glucosinolates, antinutritive compounds that render the seed meal unusable as animal feed. A fine scale linkage map of the region in radish containing *Rfo* was constructed, and a map-based cloning approach relying on synteny between radish and Arabidopsis was used to clone *Rfo*. A radish gene encoding a 687 amino acid protein with a predicted mitochondrial targeting presequence was found to confer male fertility upon transformation into Ogura CMS *B. napus*. This gene, codes for a pentatricopeptide repeat (PPR)-containing protein with multiple, in this case 16, PPR domains. Two similar genes that do not appear to function as *Rfo* flank this gene. A trasnscipt representing a non-functional allele (*rfo*) was detected in sterile radish plants. Comparison of the *Rfo* region with the syntenic *Arabidopsis* region indicates that a PPR gene is not present at the *Rfo*-equivalent site in *Arabidopsis,* although a smaller and related PPR gene is found about 40 kb from this site.

Résumé

La stérilité cytoplasmique masculine (SCM) une condition répandu et maternellement hérité. Les plantes SCM sont caractérisez par une incapacité de produire le pollen fonctionnel. Le système de SCM d'Ogura a été découvert dans un radis Japonais (*Raphanus sativus)*. Ce système a été transféré au *Brassica napus*. Un gène qui reconstitue la fertilité masculine est nécessaire pour que le système de CMS d'Ogura soit exploité commercialement pour la production de graine hybride. Le gène Rfo est un gène nucléaire de radis dominant qui reconstitue la fertilité masculine aux plantes avec le cytoplasme d'Ogura. Ce gène a été transféré au le *Brassica napus* par les croix intergeneric, toutefois introgressed le segment de ADN de radis contient un nombre inconnu de gènes, dont certains confèrent des traits indésirables. ne cartographie génétique détaillée de la région entourant le gène de la restauration a été construite. Une approche carte-basée de clonage se fondant sur synteny entre le radis et l'Arabidopsis a été utilisée comme moyen pour copier Rfo. Un gène de radis codant une protéine de 687 acides aminés s'est avéré pour conférer la fertilité masculine sur la transformation dans le B napus de CMS d'Ogura. Ce gène, encode pour une protéine contenant multiple (16) répétition de la domain "pentatricopeptide « (PPR). La comparaison de la région de Rfo avec la région syntenic d'Arabidopsis indique qu'un gène de PPR n'est pas présent à l'emplacement Rfo-équivalent dans Arabidopsis. Un gène dans la famille de ppr est trouvé à environ 40 kbs de l'emplacement équivalent de Rfo dans Arabidopsis.

Contribution to original knowledge

- *created a large (1082 plants) mapping population of* Raphanus sativus
- limited the genetic region containing the *Rfo* gene using RFLP and PCR markers
- designed a PCR based marker inside of the *Rfo* gene for mapping purposes
- demonstrated the pattern of expression of *Rfo by* rtPCR
- identified a non-functional allele of the restorer gene (*rfo*)

ACKNOWLEDGMENTS

First of all, I thank my supervisor Dr. Greg Brown for giving me the opportunity to work in his lab on such a cool project. I am also grateful to Dr. Candace Waddell, Dr. Wing Cheung, and Dr. Christian Hardtke for the guidance they have shown me as members of my supervisory committee.

I would also like to thank all of the people who have contributed to this project form both the Brown lab at McGill University and the many technicians and research assistants at DNA LandMarks. The members of the Brown lab Jinfa Zhang, Natasa Formanova and Prashant Patil, all of whom participated in the development of genetic markers and cloning of candidate genes. Many thanks to Mary Kalogerakis, and Evelyn Tsang at DNA LandMarks for their technical assistance with southern hybridizations used in the RFLP analysis of the *Rfo* region. I am grateful to Phytotron staff, Marc Romer, Clair Cooney and Frank Scopelleti for the excellent care that my plants received.

Dr. Jackie Vogel whose considerable patients and generosity of time in the editing and organization of this manuscript exceeded all reasonable expectations. For this I am sincerely grateful.

I must also thank my family for the endless support and encouragement. I reserve my deepest gratitude and appreciation for my parents and four sisters whose support gave me the strength to carry on through the difficult times.

INTRODUCTION

Intra-cellular signaling in plants

Intra-cellular signaling is important for all aspects of cellular function. However, few intra-cellular signaling systems have been characterized extensively in plants. For chloroplasts and mitochondria, is crucial for plant cell growth and development. The majority of mitochondrial proteins are encoded in the nucleus. Mitochondrial genes encode a minority of mitochondrial proteins. How proteins encoded by the nucleus regulate mitochondrial function remains unclear.

Many plant species exhibit defects in mitochondrial function that are suppressed by specific nuclear genes. For this reason, plants are an excellent system in which to study the interactions between the nucleus and mitochondria. One such example is cytoplasmic male sterility (CMS), a very common trait of plants (Schnable and Wise, 1998). CMS is caused by the expression of novel mitochondrial genes and results in a deficiency in pollen production. This defect can be corrected by the presence of a nuclear gene called a restorer (Rf) gene. Restorer genes are one of the few known examples of a single nuclear gene having a direct effect on mitochondrial function.

Plants with CMS have a clear phenotype, lack of functional pollen. For this reason, CMS has the potential to be exploited in agricultural crops, as it prevents self-pollination. However, it is important to be able to generate fertile strains for seed crop production, which requires the presence of a nuclear restorer gene in the fertile parent of the hybrid cross. Despite the agricultural and academic importance of understanding the function of restorer genes, very few have been cloned and characterized. This thesis describes the characterization of a restorer gene, *Rfo*, in the Japanese radish.

Cytoplasmic Male Sterility

As stated above, CMS results from abnormal mitochondrial genes, whose expression impairs the production of viable pollen without otherwise affecting the plant. CMS is a widespread trait in plants (Budar et al, 2001), and has been characterized in over 140 different species. CMS is frequently associated with large insertions or other rearrangements that result in novel, often chimeric mitochondrial open reading frames. In a number of cases, transcripts originating from these altered open reading frames are translated into unique proteins, which appear to interfere with mitochondrial function and pollen development (Schnable and Wise, 1998).

CMS/restorer systems have been identified and characterized in many crop species, including *Brassica napus* (canola), rice, sorghum sunflower, rye, onion, maize, carrot, beet, and petunia (Schnable and Wise, 1998). The phenomenon of CMS has been exploited in the commercial production of hybrid seed. CMS/Restorer systems have greatly facilitated hybrid seed production by eliminating the need for labor intensive hand emasculation and ensuring that the all of the seeds produced are the result of cross pollination.

CMS Systems in Brassica napus

CMS has been extensively studied in *Brassica* species because of its commercial importance as an agricultural crop. Three different CMS systems exist in *Brassica napus*: polima (*pol*) CMS, napus (*nap*) CMS and Ogura (*ogu*) CMS. The *pol* and *nap* cytoplasm

are endogenous to *Brassica*, while the *ogu* cytoplasm was transferred into *Brassica* through intergeneric crosses and protoplast fusion (Bannerot *et al.*, 1974, Pelletier et al., 1983).

Polima

The polima cytoplasm is believed to have arisen spontaneously in *Brassica* as a result of mitochondrial genetic recombination (Fu, 1981). Polima CMS is associated with the presence of a chimeric ORF, *orf224* (figure 1), that is situated upstream of the gene encoding ATP synthase subunit 6 (*atp6*) in the mitochondria (Singh and Brown, 1991). The first 58 codons of the *orf224* gene are highly similar to the amino-terminal coding region of the *orfB* gene, followed by 43 bp homology with the last exon 1 of *rps3*, while the rest of the sequence is of unknown origin (Singh and Brown, 1991).

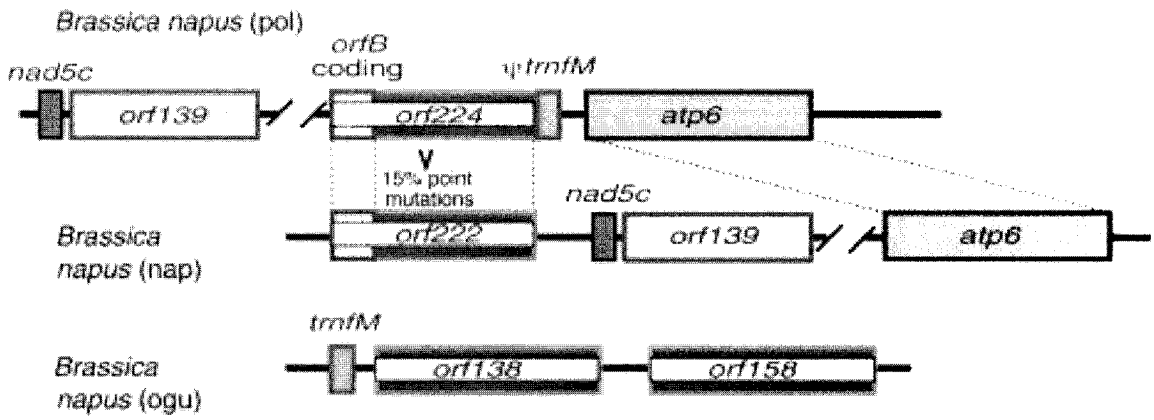

Figure 1 CMS-associated open reading frames
Chimeric cytoplasm male sterility (CMS)-associated regions in the mitochondrial genomes of *Brassica* species. CMS-associated open reading frames (ORFs) (open boxes) often include novel sequences of unknown origin (red). Other colors indicate known mitochondrial genes and are designated by standard mitochondrial gene nomenclature. Parallel dashed lines indicate sequence similarities. (Adapted from Schnable and Wise, 1998)

Napus

The *nap* CMS is associated with expression of a mitochondrial chimeric open reading frame, *orf222* (figure 1). *orf222* is co-transcribed with an exon of a trans-spliced gene, *nad5*c, and another open reading frame, *orf139* (L'Homme et al., 1997). The open reading frames responsible for causing sterility in both *pol* and *nap* CMS plants show a sequence similarity of 79% at the amino acid level (L'Homme et al., 1997).

Ogura

The Ogura cytoplasm was originally characterized in a variety of Japanese radish, *Raphanus sativus,* by Ogura (1968). The Ogura cytoplasm was transferred to *Brassica napus* by intergeneric crosses (Bannerot *et al.*, 1974), and was subsequently modified by protoplast fusion (Pelletier et al., 1983) to correct problems with chloroplast incompatibility. Ogura CMS has been associated with a 2.5 kb *NcoI* mitochondrial DNA fragment (Bonhomme *et al.*, 1992). This fragment is specific to sterile cybrids, and is lost, or rearranged in fertile revertants. Two open reading frames (figure 1) were found in this fragment, *orf138* and *orf158* (ORFB). In sterile plants, *orf138* is translated into a mitochondrial membrane polypeptide (Grelon *et al.*, 1994).

Restorer of fertility genes

Relatively little is known about the molecular features of restorer genes and the proteins they encode. *Rf2*, one of two genes required for restoration of the "Texas" type

male sterility (*cms*-T) of maize, was the first restorer gene to be characterized. The *Rf2* gene encodes a mitochondrial aldehyde dehydrogenase (Cui *et al.*, 1996). This gene does not seem to function in a similar manner to the majority of known restorer genes, which act to down regulate specific mitochondrial genes. Instead, the *Rf2* restorer seems to have a general role in mitochondrial function and has only been recruited recently to function as a nuclear restorer (Liu *et al.*, 2001).

Recently, restorer genes have been cloned in petunia, radish, and rice (Bentolila, *et al.*, 2002, Brown *et al.*, 2003, Kazama, *et al.*, 2003, Komori *et al.*, 2004). Unlike *Rf2*, these restorer genes encode proteins that are members of a single family of pentatricopeptide repeat (PPR) proteins (Small and Peeters, 2000), suggesting PPR proteins may play a conserved role in mitochondrial function. The PPR family of proteins, represents over 1% of the coding functions of the *Arabidopsis thaliana* genome, and accounts for 10% of all proteins whose specific function cannot be determined by homology (Aubourg, *et al.*, 2000). Many PPR proteins contain targeting pre-sequences, and are likely to be targeted to mitochondria/chloroplasts. Thus, the identification and characterization of CMS/PPR restorer systems could also be used to study the role of PPR proteins. For example, the identification of novel restorer genes encoding PPR proteins could shed new light on their role in intracellular signaling. It is also possible that analysis of PPR proteins may reveal restorer genes in other CMS systems.

How PPR proteins down-regulate CMS-associated mitochondrial gene expression and restore fertility to CMS plants is unclear. Two hypotheses have been proposed; first, PPR proteins may target mitochondrial proteins for degradation or alternatively that PPR proteins inhibit the expression of mitochondrial genes post-

transcriptionally. The PPR repeat is closely related to a 34 amino acid tetratricopeptide (TPR) motif involved in protein-protein interaction. For this reason, it has been suggested that PPR restorer genes may act through direct protein-protein interaction. PPR containing proteins, however, have also been found to serve as regulators of organelle gene expression in plants and fungi (Small and Peeters, 2000), and most act to modify the translation and/or processing of specific organelle-encoded mRNAs. It has been suggested that tandem PPRs may fold into a superhelix that could encircle and possibly bind an RNA strand (Small and Peeters, 2000). A PPR protein/RNA complex might inhibit translation, thereby silencing the mutation.

CMS in Petunia is encoded by an abnormal gene termed *pcf*, a chimeric mitochondrial ORF composed of portions of the coding region of ATP synthase subunit 9 fused to an ORF of unknown origin (Young and Hansom, 1987). The petunia restorer (Rf) has been found to encode a mitochondrially targeted protein that consists almost entirely of 14 repeats of PPR motifs (Bentolila, *et al.*, 2002). The petunia Rf alters both the profiles of *pcf* transcripts and reduces the levels of the encoded PCF protein (Pruitt and Hanson, 1991, Nivision and Hanson, 1989).

One particular CMS system in rice (*Oryza sativa* L.) is called the *ms-bo* type or BT type. In this BT-type CMS line, pollen abortion is initiated after pollen mitosis. This male sterility has been associated with a unique sequence (*orf79*) located downstream from mitochondrial *atp6* (Kadowaki, *et al.*, 1990). These two ORFs are co-transcribed, resulting in an abnormal *atp6-orf79* mRNA. The *Rf-1* gene encodes a 791- amino acid protein with 18 PPR repeats (Kazama and Toriyama, 2003, Komori *et al.*, 2004). The *Rf-*

1 gene product influences the processing of the altered *atp6-orf79* transcript and restores fertility of cytoplasmic male sterile rice (Iwabuchi *et al.,* 1993).

Restorer of fertility genes in Brassica napus

There are two indigenous restorer genes in *Brassica napus*, *Rfp* and *Rfn*. *Rfp* restores male fertility to *pol* CMS *Brassica,* while *Rfn* confers fertility to plants that bear the *nap* cytoplasm. Originally, it was believed that fertility in *pol* CMS plants was governed by two unlinked dominant restorer genes, derived from two separate varieties of *Brassica napus,* "Italy" and "UM2353". It is now clear that these genes map to the same locus and likely represent one single gene, *Rfp* (Jean *et al.,* 1997). Interestingly, the *Rfn* restorer maps to this genetic locus as well (Li et al, 1998). Although both restorers map to the same locus, the *pol* restorer is not capable of restoring fertility to plants that contain the *nap* cytoplasm and vice versa. Thus, it is likely that *Rfp* and *Rfn* are either different alleles or haplotypes of a single locus, a situation found thus far to be unique to CMS/restorer systems (Li et al, 1998).

The restorer gene *Rfo* is capable of restoring male fertility to plants bearing the non-indigenous Ogura cytoplasm (figure 2). The *Rfo* gene was introgressed into *Brassica napus* fom *Raphanus* through intergeneric crosses (Heyn, 1976). Krishnasamy and Makaroff (1994) demonstrated that a decrease in the amount of ORF138 protein accompanied the restoration of fertility. This reduction in ORF138 levels was detectable in western blots of mitochondrial proteins extracted from leaves and floral buds but not roots. Interestingly, mRNA levels were not reduced, suggesting the *Rfo* is acting post-transcriptionally. The *Rfo* gene decreases the amount of ORF138 protein in floral buds,

with this effect being the most dramatic in anthers at the stage of development when the sterile phenotype is normally expressed (Bellaoui, et al., 1999). Total polysome analysis of buds and anthers show that *orf138* transcripts are translated with the same efficiency in sterile and restored plants (Bellaoui et al. 1999). Based on these results, it was proposed that *Rfo* might destabilize the ORF138 protein, leading to a decrease in the accumulation of the protein and a restoration of fertility.

A radish CMS-Rf system named kosena CMS (*kos* CMS) has also been isolated from a population of the Japanese radish cultivar Kosena. The *kos* CMS system was transferred to *Brassica napus* by donor-recipient protoplast fusion (Sakai and Imamura 1992; Sakai et al. 1996). Fertile restorer lines have been identified in the same Kosena population. Sequence analysis of the Kosena mtDNA revealed that *kos* CMS carries *orf125*, which encodes a 17-kDa chimeric protein. The sequence of *orf125* is homologous to that of Ogura radish *orf138*, except for two amino acid substitutions and a 39-bp deletion in the *orf138* coding region (Grelon *et al.,* 1994, Iwabuchi et al. 1999). The accumulation of ORF125 is associated with the CMS phenotype in *B. napus* (Iwabuchi *et al.*, 1999). The restorer gene *Rfk1* restores fertility to Kosena CMS plants. It is now known that *Rfk1* is identical to *Rfo*, and both are capable of restoring fertility to male sterile plants possessing either the Kosena or Orgura cytoplasms (Koizuka *et al.,* 2000).

Map based cloning as a strategy for the isolation of the Rfo restorer gene

Although the effect of the *Rfo* restorer gene could be clearly accessed by simple a phenotypic screen, the identity of the gene responsible remained unidentified. In order

to identify the *Rfo* locus, a map-based cloning approach was chosen as it allows the identification of genes without prior knowledge of the gene product. There are two basic requirements for the use of map-based cloning: a genetic population exhibiting differences in the trait of interest, and capacity to map the trait to a chromosomal position adjacent to segments of DNA that have been or can easily be cloned.

Steps involved in map-based cloning.

1. High-resolution mapping to identify DNA markers very tightly linked to the gene in question.
2. Cloning of the DNA interval separating the regions detected by such flanking markers.
3. Identification and characterization of the DNA sequence within the interval that carries the gene specifying the trait.
4. Determination of the nucleotide and protein sequence of a candidate gene from the sequenced interval.
5. Complementation of the mutant phenotype by transformation with the candidate gene.

Genetic Markers Used In Map-Based Cloning

Agricultural traits, such as disease resistance, can be difficult to evaluate precisely. Instead, genetic markers that are linked to agronomic traits, and that can be easily scored, are used as selection tools in plant breeding programs. In addition, nucleotide polymorphisms linked to an agriculturally important trait can monitored at any time during development in any tissue. Determining the map position of a gene consists

of testing linkage with a number of previously mapped markers. Once the gene of interest has been linked to a specific marker, a fine scale map can be created by determining the linkage between markers in that region.

Restriction Fragment Length Polymorphism (RFLP)

RFLP polymorphisms result in length variations in restriction fragments that are usually the result of base pair changes in restriction endonuclease sites. RFLP markers are usually co-dominant in nature, permitting the distinction between both homozygous and heterozygous state of the alleles. DNA fragments are separated according to size by gel electrophoresis, and visualized by DNA blot hybridization (Southern) using a radioactively labeled DNA fragment corresponding to that region as a probe.

Advantages of using RFLP markers include the co-dominance of alleles and high reproducibility; membranes can be re-probed several times. Disadvantages are the large quantities of purified, high molecular weight DNA required (5-10 μg), the laborious and technically demanding construction and hybridization of membranes, and need for large quantities of radioactive isotopes to construct probes.

Cleaved Amplified Polymorphic Sequence (CAPS)

CAPS markers can also be used to visualize restriction site polymorphisms (Konieczny and, Ausubel, 1993) Like RFLPs, CAPS markers also permit co-dominant scoring of alleles. However, unlike RFLP markers, which use DNA blot hybridization, CAPS utilize PCR to visualize the polymorphisms. Sequence specific primers are created, and the region surrounding the polymorphism is amplified by PCR. Since the restriction site will only be present in one of the alleles after digestion with the

appropriate enzyme, the polymorphism can be seen as a length polymorphism using gel electrophoresis.

The main advantages to using CAPS markers are: the low quantities of template DNA required (10-100 ng per reaction), the co-dominance of alleles and the high reproducibility of results. The main disadvantages of using CAPS markers are the requirement of sequence data to construct primers, and ambiguities in scoring, due to incomplete enzymatic digestion.

Sequence Characterized Amplified Region (SCAR)

SCAR markers are PCR based markers that function similarly to CAPS markers except that they do not require endonuclease digestion (Paran and, Michelmore 1993). SCAR markers are used to detect sequence variation in a particular region. This type of marker can be dominant or co-dominant. The elimination of the requirement for enzyme digestion has the advantage of reducing ambiguous scoring caused by incomplete digestion.

The main advantages to using SCAR markers are: low quantities of template DNA are required (10-100 ng per reaction), and results may be obtained in a relatively quick and easy manner. The main disadvantage of using SCAR markers is the requirement of sequence data to construct primers.

Analysis of Flanking Markers

Flanking marker analysis was used to define a sub-population of plants, which as a result of recombination would delineate the mapping region of *Rfo*. A larger mapping population would result in a much more detailed genetic map. In a relatively small

segregating population of 100 meiotic events, a target gene can be placed within an interval of 5 to 10 cM. These defining markers can then be used to screen a larger segregating population, and identify individuals containing chromosomes in which recombination events have occurred near the gene of interest. Only such individuals are useful for more precise mapping of additional markers with respect to the target gene. This initial step of eliminating plants that could not aid in reducing the reducing the region that could contain the gene allowed for the analysis of a much larger population than otherwise possible.

The goals of this thesis work were to identify the *Rfo* locus by creating a high-resolution linkage map of the region, and to characterize its expression pattern *in vivo*.

MATERIALS AND METHODS

Plant Material and Growth Conditions

A mapping population of 1,082 Japanese radish plants (*Raphanus sativus*) was grown in climate controlled growth chambers in the McGill Phytotron facility. The growth chamber temperature was held at a constant 20°C day and 15°C night, with a relative humidity of 70% and a 16-h photoperiod. Plants were tested for recombination in the genomic region of the restorer, and were subsequently grown to maturity in a green house.

Fertility Assessment and Genotype at Restorer Loci

The presence of the *Rfo* gene was determined by examination of the phenotype of the plant. The assessment of male fertility was done by a thorough examination of the flowers, especially the anthers (Figure 2). This examination involved assessment of the

anther shape, size, and pollen production. Male sterile plants had anthers that were much smaller than normal anthers, and had a slight hook at the end. The presence of pollen on the anther was determined by touching the anther lightly with a finger to release pollen. Male fertile plants deposited yellow pollen, while plants exhibiting the male sterile phenotype did not produce pollen. The capability to visualize the pollen made the scoring of the phenotype of the plants, with regard to male fertility, unambiguous. The ability to score the genotype at the *Rfo* locus with a phenotypic screen allowed the gene to be placed in a linkage map, without prior knowledge of the gene product.

Figure 2. Two radish (*Raphanus sativus)* flowers. Both flowers contain the Ogura (*ogu*) sterile cytoplasm. However, the flower on the right contains the nuclear restorer gene *Rfo,* resulting in a fertile phenotype.

Progeny Analysis

Plants heterozygous for the restorer gene are phenotypically indistinguishable from fertile plants that are homozygous for the gene. Thus, to determine the genotype of individual fertile plants at the restorer locus, fertile plants were self-fertilized and their progeny were analyzed. A branch from the plant being crossed was selected and buds that were plump and about to open were opened using tweezers. The anthers and petals were then removed to fully expose the stigma. An anther from another flower, located on the same plant, was picked and pollen obtained. Fertilization was accomplished by applying the pollen directly on to the stigma. This procedure was repeated on 3 or 4 of the buds on the same branch, in order to obtain sufficient seeds for analysis. A white paper crossing bag was used to prevent uncontrolled pollination. The seeds produced grown to maturity and scored for their ability to produce pollen. Plants were scored as homozygous at the restorer locus if no sterile plants were obtained among the selfed progeny; the presence of sterile plants in the selfed progeny indicated that the plant was heterozygous at the restorer locus.

DNA extraction and Purification

The initial genetic screen to identify plants with recombination events in the vicinity of the restorer gene was performed with small-scale DNA preparations of Ogura radish, according to Zhang and Stewart (2000). Since the DNA content of the plant is the same in all tissues and at all stages of development, DNA can be extracted from any tissue of the plant. Cotyledons were chosen for the initial screen, as they are the first tissue to emerge. One half of a cotyledon from each plant was used for the extraction of DNA. The use of cotyledon tissue, which appears very early in development, allowed a

large number of immature plants to be screened in a limited amount of space. Plants that showed genetic events in the region of interest were grown to maturity in order to obtain sufficient quantities of DNA for RFLP analysis. The extraction of DNA for RFLP analysis was preformed as described in Cheung et al. (1997).

Genetic Markers

An F5 Japanese radish mapping population of 1,082 individuals was initially screened with two PCR based markers, 64K20-9 and 125H9-16 (Figure 3). The vast majority of (~94%) plants did not show a recombination between these two markers and are therefore extremely likely to be uniform for all other markers in this interval. The location of the restorer gene could not be determined from these plants, as it would co-segregate with all of the markers in that interval. A much smaller percentage of plants, about 6%, did show a recombination between these two markers. This drastically reduced the number of plants that needed to be scored with other markers, since only plants that showed a recombination between the two flanking markers would yield any useful information.

The 64K20-9 marker is a co-dominant PCR marker that was developed after alignment of multiple sequences derived from a BAC clone known to map near Rfo (provided by DNA LandMarks) from six sterile plants and five fertile plants. This alignment showed a 31 base pair deletion in the fertile plants. Primers were constructed (Table 1) that amplified an 181bp fragment from fertile plants, and a 212 base pair fragment from sterile plants (Figure 3A). Plants heterozygous at this locus produce both

fragments. This polymorphism can be visualized using agarose gel electrophoresis and ethidium bromide staining.

Table 1. Sequence and position of primers used for PCR based markers

Name	Position	Forward Primer	Reverse Primer	Enzyme
64K20-9	NA	AACAAGAAGTAGAGTTGATGAAACC	AATGAATGTATGATCCTAAGTTTTGAG	NA
10I4.T73F2R	52,250	GAAGAGTGAGCGTGAATCCAAT	AGCGAGTAAAACGAAACTACACG	*SphI*
12M3.T73F1R	117,742	AGTGCGCTAAAGAGTGATGC	TGCTCCTACTCGGACACTATTTG	NA
PEP3-1	158,588	CATGATGATCGAGAGGAAT	GAAGAGCTGCAGCTTATGCT	NA
PEP6 4F3R	178,245	CGCAGCTTCAAACTCTCTC	TTAGCTTGACGGGACCGAC	NA
52P9 1.5	198,732	GATTCAGTGCCCCAGTTCG	CTCAAGCAGCCAAGCTCTTAG	NA
12H9-16	NA	TCTGTTTCTTGGCTGTGGTTA	ATTAGTGCTGCGGAGTTGTTT	*BLPI*

The 125H9-16 marker was also generated by comparing sequences generated from sterile and fertile plants using primers derived from the sequence of a BAC known to map near Rfo (also provided by DNA LandMarks). In this case, although the sizes of the products from sterile and fertile plants were identical (~1080 bp), it was found that t the sequence from fertile plants contained two *Blp I* restriction sites, while the sterile plant sequence had none. Thus, after digestion with the enzyme *Blp I,* sterile plants retain the 1080 fragment, while fertile plants show two bands of 400 and 600 base pairs, as well as a 60bp fragment that is usually not detected (Figure 3B).

The more markers there are in a given region the more precisely the gene can be mapped. Thus, the region of interest was further screened with PCR and RFLP markers (Tables 1 and 2), to more precisely identify the location of the recombination event. Five additional PCR based markers were developed to aid in the quick delimiting of the region that could contain the restorer gene. The PEP3-1, PEP6- 4F3R, and the 52p9 1.5 are SCAR type markers that amplify sections of sequenced genes in the *Rfo* region. The 10I4.T73F2R, and 12M3.T73F1R markers were created using DNA sequence from the ends of Bacterial Artificial Chromosomes (BACs). Nucleotide sequence analysis revealed differences between the fertile and sterile BACs that resulted in the addition or removal

of a restriction enzyme site. These sequence polymorphisms between fertile and sterile plants were used in the creation of CAPS type markers (provide a reference for CAPS markers). DNA sequences from the radish BACs were used as RFLPs probes (designed at DNA Landmarks, Inc.) to further define the genetic region containing *Rfo*. A list of the RFLP probes and their location within the Rfo region sequence is provided in Table 3.

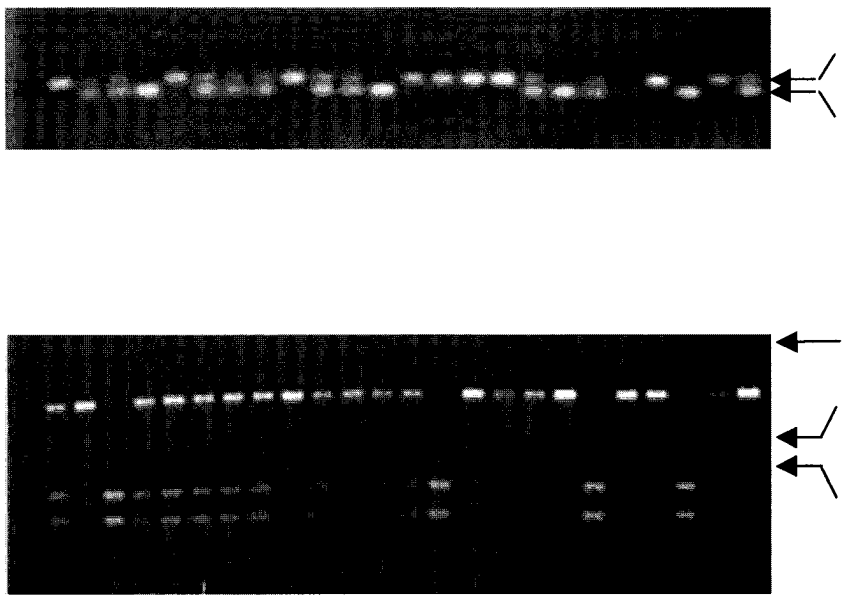

Figure 3. Examples of PCR based markers used to screen a large population segregating for *Rfo* in order to identify individuals informative for fine mapping of the gene. A. The STS marker 64K20-9; a 212 bp band is amplified from plants homozygous for the sterile allele and a 181 bp band is amplified from plants homozygous for the fertile allele. B. The CAPS marker 125H9-16/*Blp1*; Following amplification and cleavage with *Blp1*, a 1080 bp band is obtained from plants homozygous for the sterile allele and three bands of 600, 400 and 60 bp are obtained from plants homozygous for the fertile allele. The 60 bp band does not appear in photo.

Table 2. Positions and types of PCR based markers used. * The region containing the restorer gene was sequenced while this project was underway. The position is given relative to an arbitrary position selected during the sequencing process.

Name	Size	Position	Enzyme	Fragment size
D8p5sh45	1.8	49,635	BamHI	4.2+ 2.5
10I4p2sh64	41	58,574	BamHI	42.5
10I4p8sh8	1.3	68,914	EcoRV	9.6
12M3p4sh18	1.6	91,894	BamHI	42.5
12M3p9sh53	1.5	97,522	HindIII	2.2 +13.1
7A7	-	97,940	-	-
12M3p3sh30	1.6	100,797	BamHI	34.3
12M3p4sh77	1.4	102,646	EcoRI	8.8
8A8	0.399	102,980	EcoRV	-
9A9	0.272	107,710	-	-
12M3p23sh39	2.1	106,166	HindIII	13.1
12M3p6sh45	1	111,310	EcoRv	2.2
C15p1sh60	2.1	120,486	HindIII	4.4
Aip3sh53	1.2	139,253	EcoRV	8.6
101G11p2sh82	1.2	145,329	BamHI	28.9
101G11p1sh65	0.64	154,012	HindIII	8.7
101G11LE	-	160,473	BamHI	28.9
101G11pep 3-2	-	160,473	-	-
50E19p2sh85	2.9	188,432	EcoRV	1.7+8.9
50E19LE	-	192,414	-	-

Data Analysis

The genotype of each plant at a particular marker was obtained by comparison of the banding pattern obtained to the banding patterns obtained from fertile and sterile bulk DNA. The raw PCR and RFLP data was entered into an Excel spread sheet, and subsequently transferred to MapMaker. The MapMaker software (Lander *et al.*, 1987) was used to calculate map distances.

Sequence homology to extensively studied model organism *Arabidopsis* was determined using the BLAST program at www.ncbi.nih.gov. Analysis of the protein domains found in the candidate genes was determined using pfam (Protein families

database of alignments) at www.sanger.ac.uk/Software/Pfam/index.shtml. The targeting of candidate proteins to the mitochondrial was determined using the programs MITOPROTII (Claros and Vincens, 1996) and PREDOTAR (HTTP://WWW.INRA.FR/PREDOTAR/)

Gene Expression

The expression of *Rfo* in different plant tissues was determined by standard reverse transcriptase polymerase chain reaction (RT-PCR) analyses using SuperScript One-Step RT-PCR with Platinum *Taq* from Invitrogen Life Technologies according to the protocol provided by the supplier. The expression of the two flanking PPR proteins P24 and P27 was also examined using RT-PCR methods, using gene specific primers (Table 3). All RT-PCR products were cloned and sequenced to confirm their identity TOPO TA Cloning® Kit for Sequencing, using protocols provided by the supplier.

Table 3
Sequence and position of primers used for RT-PCR analysis of PPR genes at the restorer locus.

Gene Name	Primer Name	Position	Sequence (5' to 3')	Size of PCR product
P24	P24-F1	161991	AGTGACATGCTTCGATCTCGTCCTTTA	347
	P24-R1	161644	ACCCTCGCGGCAAAGACCATT	
P25	P25-F1	164047	AACCCTAGCGCTACGGAATGCT	301
	P25-R1	164348	GGTGATGCATGTGAAGCGTTCTC	
P26 (*Rfo*)	P26-F3		CTCGTCCTTTACCTTCTGTGGTTGATT	
	P26-R3		GACCATCTTCCATCATCCGATCAAG	
P27	P27-F1	178188	GATGCGATTGATTTGTTCGGTGATATGG	359
	P27-R1	177829	AAGACCGTTCATCAGCGTGGTGA	

DNA Sequencing

DNA sequencing, sequence assembly and sequence annotation for the radish restorer gene region was performed at DNA LandMarks Inc. (St-Jean-sur- Richelieu, Quebec). BAC and cosmid clones mapping to the *Rfo* interval were sheared to an average size of 5 to 10 kb, and these fragments were subcloned into a standard plasmid vector and sequenced on an ABI 3700 sequencer. The sequence was assembled using the STADEN software package, and annotated for the presence of potential protein coding sequences with GENSCAN software,

RESULTS

Screen for recombinants in the vicinity of Rfo

In order to reduce the number of plants that would be subjected to genetic analysis an initial screen was performed. A previous mapping study had established that the restorer gene lay within a region contained by the two markers, 64K20-9 and 125H9-16 located at position P1 and P2 on the map (Figure 4). The exact size of the region was unknown since the region had yet to be sequenced in radish.

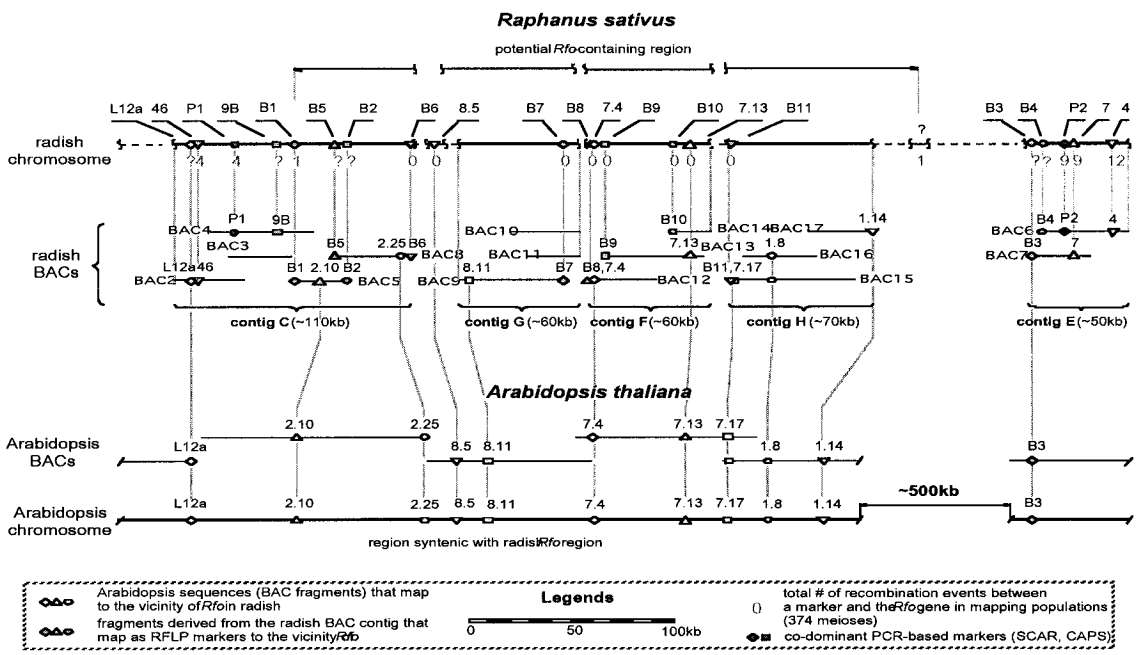

Figure 4. A previous mapping study had established that the restorer gene Rfo lay within a region contained by the two markers two markers, 64K20-9 and 125H9-16 located at position P1 and P2 on the map

A population of 1082 radish plants was screened with two markers, 64K20-9 and 125H9-16 (Figure 3). Each plant was co-dominantly scored for the two markers which flanked the region containing the restorer gene. The genotype was assessed at both the 64k20-9 and 125H9-16 loci. The plants were scored a being homozygous for the restorer line allele (A), heterozygous (H), or homozygous for the sterile line allele (B). The PCR reaction used for the 64k20-9 and 125H9-16 markers failed to yield a PCR product in 1.57% and 4.44% of samples respectively. A recombination event was deemed to have occurred in plants that did not show the same genotype at both loci examined. Plants that showed a recombination between the two markers were selected for further analysis.

A total of 66 (6.09%) the radishes showed a recombination between the 64k20-9 and 125H9-16 markers, these plants were selected for further analysis. The map distance between the two markers was calculated to be 3.7 cM. This initial screen did not narrow the region where *Rfo* could be located on the chromosome. However this screen limited the number of plants, which needed to be scrutinized for the construction of the fine scale map.

Preliminary data analysis revealed that the markers segregate in a Mendelian fashion (Table 4). The alleles segregated in a 1:2:1 ratio as expected from a cross of two heterozygous individuals. This showed that the there was no selection for either the presence or absence of the *Rfo* gene.

Fine Scale Mapping

A fine scale map of the region was created to limit the region that could genetically contain the restorer gene *Rfo*. Seven PCR based markers, 64K20-9, 10I4.T73F2R, 12M3.T73F1R,PEP3-1, PEP6 4F3R, 52P9 1.5,12H9-16, were used to

cover the region were used to further define the region containing *Rfo*. The genotypes were scored: A (*Rfo/Rfo),* H (*Rfo/rfo*), B (*rfo/rfo*), D (*Rfo/---*) (Table5). The closest flanking PCR based markers not co-segregating with the gene were 10I4 T7 and 52P9-1.5 (Table 8). These markers allowed the potential *Rfo* coding region to be narrowed to approximately 146 kb.

Table 4. The frequencies of the alleles used in the initial screen. A (*Rfo/Rfo),* H (*Rfo/rfo*), B (*rfo/rfo*), D (*Rfo/---*)

Population	64K20-9				125H9-16			
	H (*Rfo/rfo)*	A (*Rfo/Rfo)*	B (*rfo/rfo)*	no score	H (*Rfo/rfo)*	A (*Rfo/Rfo)*	B (*rfo/rfo)*	no score
1	40	20	24	-	38	22	24	-
A	32	38	20	-	31	39	20	-
B	73	47	45	-	77	39	20	29
C	236	107	108	11	232	110	113	7
D	88	38	43	1	87	39	41	3
E	60	25	21	5	57	21	23	10
TOTAL	529	275	261	17	522	270	241	48
% of total scorable	49.67%	25.82%	24.51%	1.60%	50.48%	26.11%	23.31%	4.64%

Table 5. The genotypes, 64K20-9, 10I4.T73F2R, 12M3.T73F1R,PEP3, PEP6, 52P9 1.5,12H9-16, were used to cover the region were used to further define the region containing *Rfo*. The genotypes were scored: A (*Rfo/Rfo),* H (*Rfo/rfo*), B (*rfo/rfo*), D (*Rfo/--*)

Plant	64k20-9	10I4 T7	12M3	Pep 3	PEP 6	52P9-1.5	125H9-16	Rfo
9	H	A	A	D	A	A	A	D
10	B	B	B	B	B	B	H	B
17	B	H	H	D	H	H	H	D
62	H			D	A	A	A	D
72	H	H	A	D	A	A	A	D
73	H			D	A	A	B	D
74	A	A	A	D	A	-	H	D
A 3	H	H	H	D	H	H	B	D
A 5	A	H	H	D	H	H	H	D
A 10	B	H	H	D	H	H	H	D
A 70	H	H	A	D	A	A	A	D
A 108	H	H	A	D	A	A	A	D
B 33	H	A	A	D	A	A	A	D
B 57	H	-	H	D		H	A	D
B 68	B	B	B	B	B	B	H	B
B 73	A	-	A	D	A	A	H	D
B 78	B	B	B	B	B	H	H	B
B 79	B	H	H	-	H	-	H	D
B 91	A	A	A	-	A	H	H	D
B 96	A	H	H	D	H	H	H	D
B 110	H	H	H	D	H	H	B	D
B 111	A	A	A	D	A	-	H	D
B 128	A	A	A	D	A	-	H	D
B 139	H	-	-	-	H	-	B	D
B 140	H	-	-	-	H	-	B	D
B 149	H	-	-	-	H	-	A	D
B 167	H	-	-	-	H	-	A	D
B 169	B	B	H	D	H	-	H	D
B 178	A	-	-	-	H	-	H	D
B 189	A	-	-	-	A	-	H	D
C 6	H	H	B	B	B	-	B	B
C 22	B	B	B	B	B	-	H	B

C 31	A	H	H	D	H	-	H	D
C 34	H	A	A	D	A	-	A	D
C 60	H	H	H	D	H	-	A	D
C 74	B	H	H	D	-	-	A	D
C 77	H	H	H	D	H	H	A	D
C 94	H	-	B	B	B	-	B	
C 97	A	A	A	D	H	-	H	B
C 174	B	B	B	B	B	H	H	D
C 182	A	A	A	D	H	H	H	B
C 190	B	H	H	D	H	H	H	D
C 239	A	A	A	D	-	A	H	D
C 252	H	H	H	D	-	H	B	D
C 269	H	H	H	B	-	H	A	D
C 271	A	H	H	D	H	H	B	D
C 273	A	H	H	D	-	-	B	D
C 280	H	A	A	D	A	A	A	D
C 281	H	H	H	B	H	H	A	D
C 296	H	H	H	D	H	H	B	D
C 316	H	A	A	D	A	A	A	D
C 326	H	H	H	B	H	H	A	D
C 367	H	H	H	D	H	H	B	D
C 369	H	H	H	D	H	H	B	D
C 383	B	B	B	B	B	-	H	-
C 398	H	A	A	D	A	A	A	B
C 456	H	H	H	D	H	H	B	D
D 35	H	-	H	D	H	A	A	D
D 75	H	-	B	B	B	B	B	D
D 80	B	-	-	B	H	-	H	B
D 139	B	-	H	D	H	-	H	D
D 145	B	-	H	D	H	H	H	D
E 32	H	-	H	D	H	H	B	D
E 37	A	-	A	D	A	A	H	D
E 55	H	-	H	D	H	H	B	D
E 68	H	-	A	D	A	A	A	D

The plants showing genetic recombination in the vicinity of the *Rfo* gene were then screened with an additional 20 RFLP markers (Table 2). The results of this analysis are presented in Table 5. The map distance between the RFLP markers was computed using MapMaker software. The closest flanking RFLP markers not co-segregating with the gene are 12M3p3sh30 and 50E19LE. A sterile plant, B78, which scored as heterozygous for marker to the right of 50E19LE and homozygous for the non-restoring allele for markers to the left (Table 7) defined the right boundary of the region that could genetically contain *Rfo*. Plant B78 established that the restorer gene must lie to the right of the marker 50E19LE. Another plant, C6, defined the left boundary of the region. Plant the non-restoring allele for all markers tested to the right (Table 7). These two plants narrowed the region that could contain the restorer gene to approximately 92 kb (Figure 5). A sequence analysis of the region that could contain the restorer gene revealed 13 putative candidate genes, G16 to G31 (Figure 5).

Table 6. The plants showing genetic recombination in the vicinity of the *Rfo* gene were then screened with an additional 20 RFLP markers. The genotypes were scored: A (*Rfo/Rfo)*, H (*Rfo/rfo*), B (*rfo/rfo*), D (*Rfo/---*)

plant	D8p5sh45	10I4p2sh64	10I4p8sh8	12M3p4sh18	12M3p9sh53	7A7	12M3p3sh30	12M3p4sh77	8A8	9A9	12M3p23sh39	12M3p6sh45	C15p1sh60	AIp3sh53	101G11p2sh82	101G11p1sh65	101G11LE	101G11pep 3-2	50E19p2sh85	50E19LE
9	A	A	A	D	D	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
17	H	H	H	D	D	H	H	H	H	H	A	H	H	H	H	H	H	H	H	H
A3	H	H	H	D	D	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
A5	H	H	H	D	D	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
A10	H	H	H	D	D	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
A70	A	A	A	D	D	A	A	A	A	A	-	A	A	A	A	A	A	A	A	A
A108	H	A	A	D	D	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
B33	A	A	A	D	D	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
B68	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
B73	A	A	A	D	D	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
B78	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	H
B79	H	H	H	D	D	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
B91	A	A	-	D	D	A	A	-	A	A	-	A	-	-	A	-	A	A	A	A
B96	H	H	H	D	-	H	H	H	H	H	-	H	-	H	H	-	H	H	H	H
B110	H	H	-	D	D	H	H	H	H	H	H	H	H	-	H	H	H	H	H	H
B111	A	A	-	D	D	A	A	A	A	A	-	A	-	A	A	-	A	A	A	A
B128	A	A	-	D	D	A	A	A	A	A	-	A	A	A	A	A	A	A	H	A
B169	H	H	-	D	-	H	H	-	-	-	-	-	-	-	H	-	H	H	-	-
C6	H	H	-	D	D	H	H	B	B	B	B	B	B	B	B	B	B	B	B	B
C31	H	H	H	D	D	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
C34	A	A	-	D	D	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
C60	H	H	-	D	D	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
C74	H	H	H	D	D	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
C77	H	H	-	D	D	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
C80	H	H	-	D	D	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
C97	A	A	-	D	D	-	A	A	A	A	A	A	A	A	-	A	A	A	A	A
C174	B	B	-	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
C182	A	A	A	D	D	A	A	A	A	A	A	A	A	A	A	A	H	H	A	A
C190	H	H	-	D	D	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
C239	A	A	-	D	D	A	A	H	H	H	-	H	A	H	A	A	A	A	H	H
C252	H	H	-	D	D	H	H	A	A	A	H	A	H	A	H	H	H	H	A	A

plant	D8p5sh45	10I4p2sh64	10I4p8sh8	12M3p4sh18	12M3p9sh53	7A7	12M3p3sh30	12M3p4sh77	8A8	9A9	12M3p23sh39	12M3p6sh45	C15p1sh60	AIp3sh53	101G11p2sh82	101G11p1sh65	101G11LE	101G11pep 3-2	50E19p2sh85	50E19LE
C271	-	-	D	D	D	H	H	H	H	H	H	H	-	H	-	H	H	H	H	H
C280	-	-	D	D	D	A	A	A	A	A	A	A	-	A	-	A	A	A	A	A
C281	-	-	D	D	D	H	H	H	H	H	H	H	-	H	-	D	H	H	H	H
C296	-	-	D	D	D	H	H	H	H	H	-	H	-	H	-	D	H	H	H	H
C367	-	-	-	D	D	H	H	H	H	H	-	H	-	H	-	D	H	A	H	-
C456	-	-	D	D	D	H	H	H	H	H	H	H	-	H	-	H	H	H	H	H
72	-	-	D	D	D	A	A	A	A	A	A	A	-	A	-	A	A	A	A	A
C398	-	-	D	D	D	A	A	A	A	A	A	A	-	A	-	A	A	A	A	A
10	-	-	B	B	B	B	B	B	B	B	B	B	-	B	-	B	B	B	B	B
C316	-	-	D	D	D	A	A	A	A	A	A	A	-	A	-	A	A	A	A	A
C369	-	-	D	D	D	H	H	H	H	H	H	H	-	H	-	H	H	H	H	H
C22	-	-	B	-	B	B	B	B	B	B	-	B	-	-	-	-	B	B	B	B
C326	-	-	-	-	D	-	H	H	H	H	-	H	-	H	-	A?	H	H	H	H
C383	-	-	B	-	B	B	B	B	A	B	-	B	-	B	-	B	B	B	B	B
F bulk	-	-	B	B	B	B	B	-	B	B	B	B	-	B	-	B	B	-	B	B
S bulk	-	-	D	D	A	-	-	-	A	A	A	A	-	A	-	A	A	-	A	A

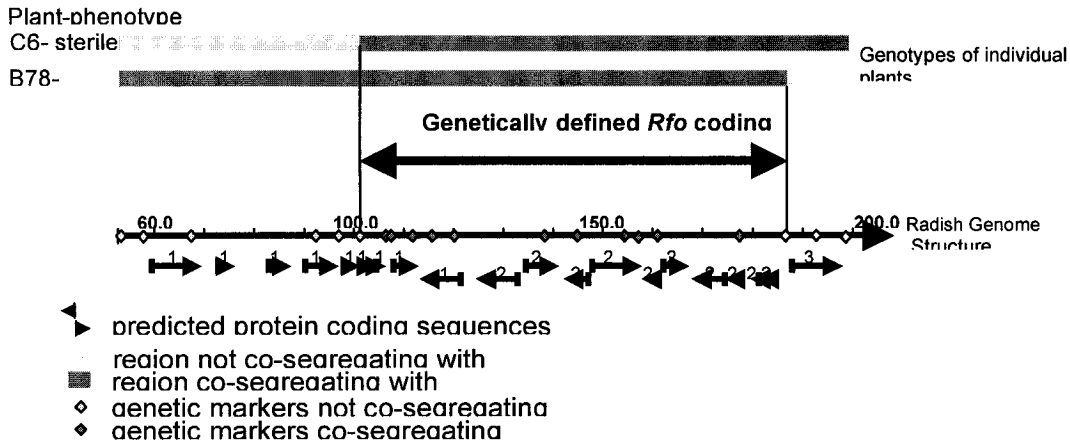

Figure 5. The genetically defined boundaries for the location of *Rfo* The black indicates sequence co-ordinates of 270-kb segment of radish DNA surrounding *Rfo*. The genotypes and location of the recombination is indicated by the colored bars above. The genetically defined boundaries for the location of *Rfo* based on the analysis of F5 individuals using different sequences from within the sequenced region as RFLP probes. The location and direction of transcription of GENSCAN predicted protein coding sequences (11, 12...) are indicated by the arrows beneath the bar.

Table 7. Genotypes and phenotypes of the recombinant plants near to the *Rfo* locus. The genotypes (H- *Rfo/rfo*, B- *rfo/rfo*, D-*Rfo/ Rfo* or *rfo*)were determined using a combination of RFLP and PCR based markers.

Name	B78	C6
64k20-9	B	H
10I4 T7	B	H
D8p5sh45	B	H
10I4p2sh64	B	H
10I4p8sh8	B	-
12M3p4sh18	B	D
12M3p9sh53	B	D
7A7	B	H
12M3p3sh30	B	H
12M3p4sh77	B	B
8A8	B	B
9A9	B	B
12M3p23sh39	B	B
12M3p6sh45	B	B
12M3	B	B
C15p1sh60	B	B
Aip3sh53	B	B
101G11p2sh82	B	B
101G11p1sh65	B	B
Pep 3	B	B
101G11LE	B	B
101G11pep 3-2	B	B
50E19p2sh85	B	B
PEP 6	B	B
50E19LE	H	B
52P9-1.5	H	B
125H9-16	H	B
Phenotype	Sterile	Sterile

Identification of Rfo *by complementation*

Plant transformation experiments were conducted to determine the capacity of the various predicted genes to act as restorers of Ogura CMS in *Brassica napus*. The candidate genes were subcloned from the genomic BAC or cosmid clones, individually or in combination with flanking genes, into binary transformation vectors, and introduced into Ogura CMS *B. napus* plants by *Agrobacterium*-mediated transformation. Only one of the constructs, Bgl-5, led to the recovery of plants displaying male fertility (Figure 7) that could be stably co-transmitted to T1 progeny plants along with the introduced DNA. These male fertile plants showed no phenotypic alterations apart from the change in floral phenotype from Ogura CMS to the completely male fertile restored phenotype. Only a single open reading frame was found to be present on the Bgl-5 construct, G26 (*Rfo*).

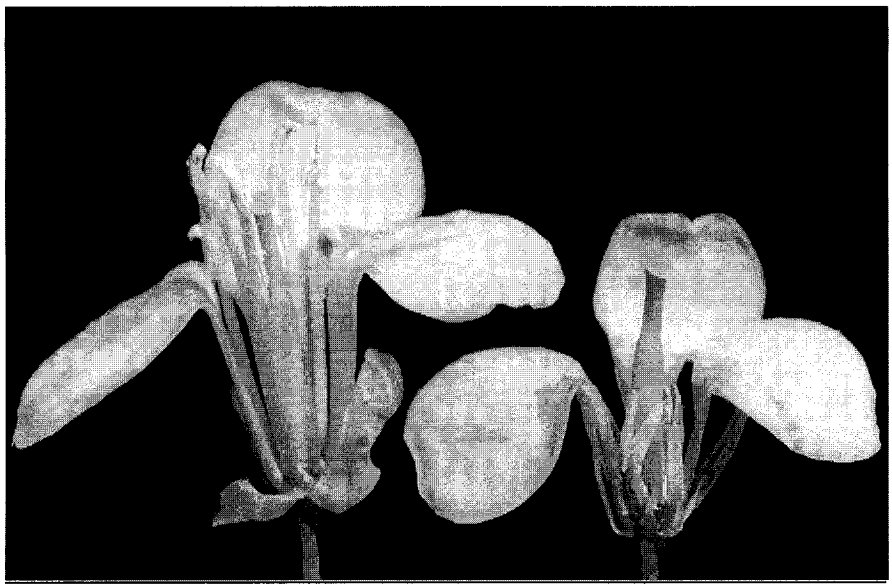

Figure 7. Flowers of a fertile T1 g26 (Rfo) transgenic *Brassica napus* plant with *ogu* CMS cytoplasm (left) and a sterile, non-transgenic sib (right). A portion of the corolla of each flower has been removed to clearly display stamen morphology

Structure of the Rfo gene

The restorer gene *Rfo* consists of a radish gene encoding a 687 amino acid protein comprising 16 PPR domains (Figures.8 and 9) and a mitochondrial targeting pre-sequence. This gene like other members of PPR family of proteins is composed nearly entirely of a 35 amino acid pentatricopeptide repeat (PPR) (Small and Peeters, 2000). The N-terminus of G26 (*Rfo)* is predicted with a likelihood of over 98% to function as a mitochondrial targeting presequence by the programs MitoProtII and Predotar.

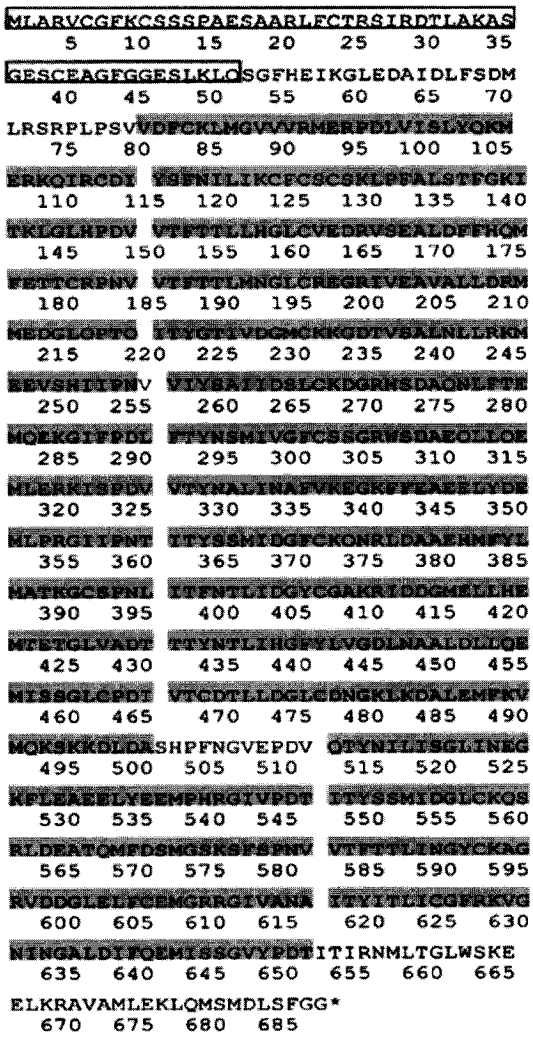

Figure 8. Predicted sequence of the protein encoded by the *Rfo* gene (g26). The predicted mitochondrial targeting presequence is indicated by the open box. Shaded residues indicate PPR domains.

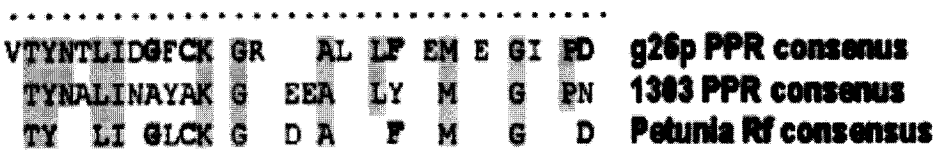

Figure 9. Consensus sequences of PPR domains of: the Rfo protein (g26p) (residues present in 6 of 16 domains); the MIME derived consensus of 1303 PPR domains; and *Petunia* Rf protein (residues present in 6 of 14 domains). Shading indicates residues conserved between the 1303 PPR consensus and one or both of the Rf protein PPR consensus sequences. Residues found in both restorer protein consensus sequences but not the 1303 consensus are indicated in bold

Expression of of G26 (Rfo) in Ogura radish

The expression of G26 (*Rfo)* was examined using RT-PCR techniques. RT-PCR was preformed using primers P26F3and P26R3 (Table 4) internal to the open reading frame of G26 (*Rfo)* on RNA extracted roots, cotyledons, leaves, and flowers of Ogura radish. All RT-PCR products were cloned and sequenced, to confirm their identity. Transcripts corresponding to G26 (*Rfo*) were detected in cotyledons, leaves, and flowers but no transcript was detected in RNA extracted from roots (Figure 10). This pattern of expression is consistent with previous work that shows that *Rfo* acts to decrease the level of expression of the CMS-associated protein ORF138 in leaves and flowers, but not in roots (Krishnasamy and Makaroff, 1994). An identical pattern of expression was observed in transgenic *Brassica napus* bearing the restorer gene on the *Bgl-5* construct.

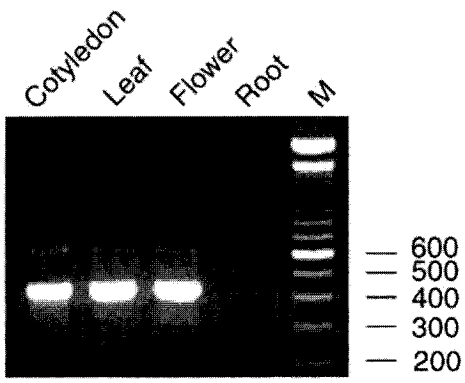

Figure 10. Rfo is not transcribed in roots, but is transcribed in most other tissues
RT-PCR products obtained with *Rfo*-specific primers from different organs of homozygous restored radish plants. *Rfo* transcripts as well as transcripts of the other PPR protein encoding genes of the Rfo region, g24 and g27, are expressed in flowers, leaves and cotyledons, but not in roots. The expression pattern of *Rfo* is consistent with its organ-specific capacity to reduce levels of the *ogu*-CMS associated protein ORF138

Non-Restoring allele

A transcript unique to sterile plants was identified by comparing the RT-PCR products obtained from RNA pooled from fertile and sterile plants derived from the segregating population (Figure 11). This transcript suggests the presence of a transcribed but non-functional allele of *Rfo*. The RT-PCR product was sequenced and aligned to the sequence of the *Rfo* gene. A twelve base pair deletion corresponding to the third PPR domain of the gene was observed. This result suggests the presence of a transcribed but non-functional allele of the *Rfo* gene.

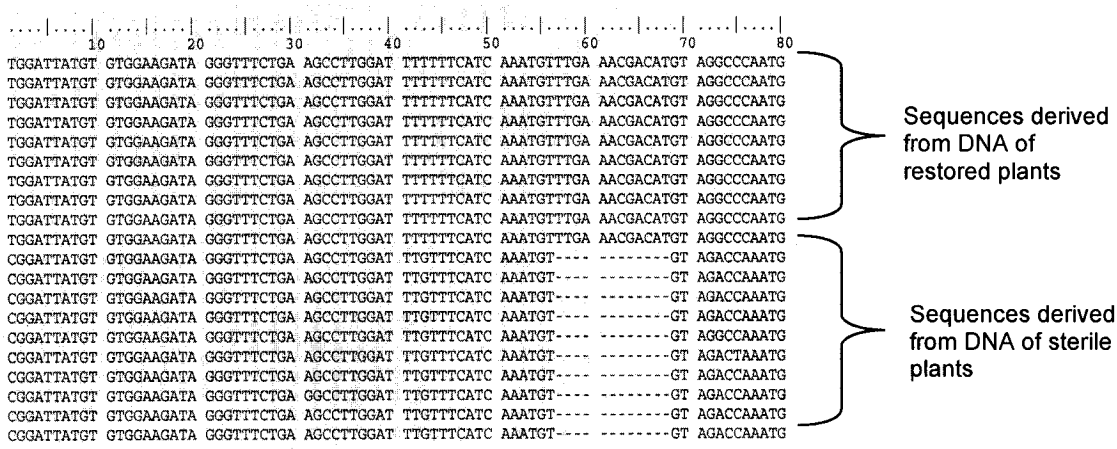

Figure 11. Structure of an *rfo*-like transcript unique to sterile plants. PCR on sterile and restored genomic Bulk DNA reveals a size polymorphism of 12 base pair corresponding to the third PPR domain of the Rfo. This suggests that the sterile allele of Rfo (rfo), like the non-restoring PPR protein-encoding genes in the Rfo locus, g24 and g27, lacks this third domain. Some non-specific amplification of closely related PPR encoding genes was observed from both fertile and sterile bulks. However, the 12 base pair segment appears does not appear in any of the the sterile bulk sequencesPossible modes of action of Rfo

Development of a SCAR marker inside the Rfo open reading frame

The discovery of sequence polymorphisms between transcripts from fertile and sterile plants lead to the creation of a SCAR marker that is located inside of the coding region of the *Rfo* gene. This genetic marker located inside the coding region of the gene allows for the tracking of the gene at any stage of development.

Flanking PPR Proteins

Two additional genes G24 and G27 fall within a 20 kb segment of the radish genome encompassing *Rfo* region contain multiple PPR domains. The three genes are 72 to 86% identical and 80 to 90% similar at the amino acid sequence level to *Rfo*. However, they do not appear to function as restorers. The number of PPR domains differs between the three genes with G24 containing 15 PPR domains, G27 containing 14 PPR domains compared with 16 found in G26 (*Rfo)*. A depiction of the alignment of these 3 related PPR proteins is presented in Figure 12. Intriguingly, the third PPR domain is disrupted in both of the flanking (and non-restoring) PPR containing genes as it is in *rfo*. This could indicate that the third PPR domain is essential for functioning of the protein.

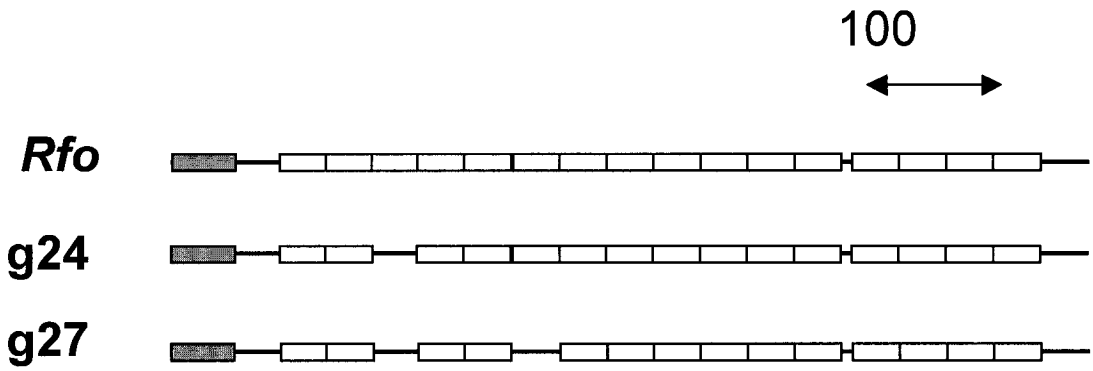

Figure 12. PPR domain-encoding proteins located in the radish Rfo region
Structure of three PPR domain-encoding proteins located in the radish Rfo region. Note the absence of the third PPR domain in the producst of the non-restoring genes g24 and g27.

DISCUSSION

Fine mapping analysis and its relationship to the identification of the restorer gene by complementation studies

As stated in the Introduction, map-based cloning of a gene proceeds through several phases: rough genetic mapping of the target gene with DNA markers, fine genetic mapping, cloning and sequencing of the region that is defined to contain the gene on the basis of the mapping studies, complementation of the appropriate recessive gene through transformation studies and finally characterization of expression of the identified gene and of alternative/mutant alleles of the gene. This thesis was directed towards fine genetic mapping of the Ogura restorer gene *Rfo* in its initial phases and towards characterization of expression of this gene and identification of the recessive *rfo* allele in its latter stages.

In order to identify the Ogura *Rfo* restore gene a large mapping population of Japanese radishes bearing the Ogura cytoplasm was grown. Plants that showed a genetic recombination in the vicinity of the restorer gene were subjected to further genetic analysis. The plants in which the phenotype of the plant did not co-segregate with the genotype of the genetic marker used in the screen were identified. This indicates that the restorer gene is not in the region covered by that particular marker. Other markers were found to co-segregate with the *Rfo* in all 1082 examined individuals of the population. These markers therefore define the region of the radish genome that could potentially encode *Rfo*. On this basis a fine scale linkage map was created that allowed the gene to be located within a 92 kb region. DNA sequence analysis of this region predicted the presence of 13 candidate genes. This mapping effort therefore limited the size of the

region and number of candidate genes that needed to be tested in the in complementation screen for the restorer gene.

One draw back to this approach is that it does not allow for the detection of plants in which a double recombination event occurred within the region being studied. However, the size of the area being studied is relatively small and the incidence of double recombination events could therefore be considered to be relatively insignificant.

The PCR reaction used for the 64k20-9 and 125H9-16 markers failed to yield a PCR product in 1.57% and 4.44% of samples respectively. This could be due to one of several factors from the extraction of DNA to problems with the PCR reaction. In an effort to screen the maximum number of plants in a short period of time these plants were not used in the generation of the creation of the linkage map.

Plant transformation experiments were conducted to determine the capacity of the various predicted genes to act as restorers of Ogura CMS in *Brassica napus*. The candidate genes were sub-cloned from the genomic BAC or cosmid clones, individually or in combination with flanking genes, into binary transformation vectors, and introduced into Ogura CMS *B. napus* plants by *Agrobacterium*-mediated transformation. Only one of the constructs, Bgl-5, led to the recovery of plants displaying male fertility that could be stably co-transmitted to T1 progeny plants along with the introduced DNA. These male fertile plants showed no phenotypic alterations apart from the change in floral phenotype from Ogura CMS to the completely male fertile restored phenotype. Only a single open reading frame was found to be present on the Bgl-5 construct, G26 (*Rfo*).

Characteristics of the Rfo gene product

The P26 (*Rfo)* open reading frame was found to encode a protein of 687 amino acids in length. The *Rfo* protein was predicted to code for a pentatricopeptide repeat (PPR)-containing protein with16, PPR domains. This protein was predicted to with a likelihood of over 98% to contain an N-terminal mitochondrial targeting pre-sequence by the programs MitoprotII and Predotar. The targeting of G26 to the mitochondria places the P26 gene product in the organelle where the effect of *Rfo* is observed gives further support to the idea the G26 is *Rfo.*

The Ogura restorer gene is very similar in structure to other recently characterized restorer genes such as the *Rf-PPR592* restorer of fertility gene in petunia (Bentolila et al., 2002), and *Rf-1* a nuclear fertility restorer gene from rice (Komori *et al.,* 2004). *Rf-PPR592* is predicted to encode a protein with 14 tandem copies of the PPR motif as well as a mitochondrial targeting sequence. *Rf-1* also encodes mitochondrially targeted protein containing 16 repeats of the PPR motif. It is expected that regardless of the CMS system that the restorer gene product should be targeted to the mitochondria.

To determine if the expression pattern of the candidate gene G26 matches the tissues where the restorer gene is known to be functionally active, the expression pattern of the P26 was examined using RT-PCR. The candidate gene P26 was expressed in the tissues where *Rfo* is known to have an effect of mitochondrial levels of *orf138.* These experiments strongly indicated that the restorer gene in the genetically defined *Rfo* region corresponded to the G26 ORF.

Koizuka et al. (2003) recently reported the genetic characterization of a pentatricopeptide repeat protein gene, orf687, which restores fertility in the cytoplasmic male–sterile Kosena radish. This gene contains 16 copies of the PPR domain. This gene differs slightly at the nucleotide sequence from Rfo, but the predicted protein it encodes is identical to g26p. This result is not completely unexpected, since the chimeric open reading frame ORF125 associated with Kosena CMS system is very similar to ORF138 that is responsible for male sterility Ogura radish. It has been previously shown that pollen containing the Ogura restorer gene Rfo is capable of restoring fertility to CMS Kosena plants and vice versa.

Model for how Rfo could restore fertility

In plants bearing the Ogura cytoplasm the levels of the CMS associated protein ORF138, are noticeably reduced in flowers and leaves, but not in roots, of restored versus *ogu* CMS radish plants (Krishnasamy and Makaroff, 1994). *Rfo* appears to act at either the translational or post-translational level, since the transcript levels of the CMS associated ORF138 do not appear to be altered in both restored and sterile plants. ((Krishnasamy & Makaroff, 1995).

The consensus PPR domain structure that predominates in the *Rfo* protein shows homology to the TPR (tetratricopeptide) motif, which is found in wide array of proteins that interact with other proteins (Small & Peeters, 2001). Bellaoui et al. (1999) showed using total polysome analyses of buds and anthers show that the *orf138* transcripts are translated with the same efficiency in sterile and restored plants. From these results we infer that the *Rfo* gene product acts on the post-translational stability of the ORF138

protein, leading to a decrease in the accumulation of the protein and a restoration of fertility. On the basis of these results, it has been suggested that *Rfo* acts post-translationally in *B. napus* by selectively targeting ORF138 for proteolysis (Bellaoui *et al.*, 1999). It is possible that one or more of the PPR domains of *Rfo* mediate its specific binding, directly or indirectly; to ORF138, and that this binding might then target ORF138 for specific proteolysis.

The restorer gene may act directly to alter translation of the *orf138* protein. Other PPR containing proteins have been found to serve as regulators of organelle gene expression in plants and fungi (Small and Peeters, 2000), and most act to modify the translation and/or processing of specific organelle-encoded mRNAs. It has been suggested that tandem PPRs may fold into a superhelix that could encircle and possibly bind an RNA strand (Small and Peeters, 2000).

The mechanism of restoration of fertility is different between the restorer genes. In the Ogura system the level of the CMS associated transcript *orf138* are not affected by the presence of the restorer gene. However, the petunia restorer Rf-PPR592 acts to reduce both the levels of the CMS associated transcript *pcf* as well as the PCF protein (Nivison and Hanson, 1989). The PPR restorer genes may act in different manners in different CMS systems. It is possible that one or multiple PPR domains are responsible for recognizing the substrate. The PPR protein may act alone in binding or in concert with other proteins. It is possible that one or more of the catalytic proteins may not be present in the Ogura CMS system resulting in the down regulation of translation, in the absence of transcript degradation.

Agronomic use of Rfo

The presence of a restorer of fertility gene in the nucleus is essential if the seed is the crop being harvested. In the absence of a restorer gene, CMS cannot be used to create hybrid F1 seeds. If a CMS plant is crossed with a fertile plant, all of the seeds collected from the sterile plant will be the result of cross pollination and will, therefore, be hybrid. However, all of the plants arising from the hybrid seeds will be male sterile, since that trait is maternally inherited. The presence of pollen is essential for seed production. The presence of a restorer of fertility gene in the nucleus of the male parental plant would restore fertility to the F1 generation hybrid plants, resulting in fully fertile plants. The availability of an isolated *Rfo* gene provides an opportunity to develop new restorer lines by introducing *Rfo* into *Brassica napus* by transformation There have been limits to to use of the *ogu* CMS system due to the difficulties in the development of restorer lines free of negative agronomic characters that were introduced from radish into *B. napus* along with *Rfo*. In particular, most existing restorer lines have a dominant radish gene that conditions elevated levels of seed glucosinolates such that hybrids produced with these lines cannot be approved to be grown as canola (Delourme et al., 1998).

Future Work

Now that the identity of the Ogura restorer gene has been ascertained a mechanism for how it controls needs to be established. A construct containing the open reading frame of *Rfo* fused to coding sequence of GFP (green fluorescent protein) would be used to establish that the *Rfo*(G26) is translated into a protein. This construct would

also be valuable in the confirming the putative sub-cellular targeting of the gene to the mitochondria. An *Rfo* GFP fusion could also aid in the isolation and purification of the protein.

The *Rfo* protein is composed of multiple tandem repeats of a 35 amino acid PPR motif. It is possible that one or more of these repeats are responsible for the functionality of the gene. A series of deletion of one or several of the se domains could lead to a greater understanding of which domains are essential. There are two very similar proteins that consist of multiple PPR domains. Swapping domains between these three genes and observing the affect on male fertility could reveal which components of the restorer are essential for proper function.

Protein-RNA

The restorer gene consists of tandem PPRs that may fold into a superhelix that could encircle and possibly bind the *orf138* mRNA strand directly. A gel shift experiment using the *orf138* transcript could show that *Rfo* interacts directly with the *orf138* mRNA. This experiment could be repeated using fragments of the *orf138* transcript, this would aid in the determination of segments of *orf138* interact with *Rfo*. This data could also be applied to the closely related *orf125* CMS associated transcript from Kosena radish.

Protein-Protein

The theory that *Rfo* effects fertility by way of direct interaction with the CMS associated protein ORF138, could be tested by looking for protein-protein interactions. If the proteins interact directly a yeast two-hybrid could be done to see if *Rfo* interacts

directly with either of these two proteins. A yeast two-hybrid assay in which only ORF138 could show that the two proteins interact.

One single gene restores fertility in both the Ogura and Kosena CMS systems. These two systems have very similar CMS associated proteins ORF138 and ORF125 respectively. If the proteins interact directly the next step would be to determine which domains are involved in the interaction. One way to do this would be to employ a yeast two-hybrid assay using sections of the ORF138 protein. Another approach would be to use a GST-pull down assay to see if any proteins co-purify with the restorer gene. This approach could be made more sensitive by using proteins obtained from isolated mitochondria of restored plants.

REFRENCES

Aubourg, S, Boudet, N, Kreis, M, Lecharny, A, (2000) In Arabidopsis thaliana, 1% of the genome codes for a novel protein family unique to plants. Plant Mol. Biol. **42**: 603–613.

Bannerot, H., Boulidard, L., Cauderon, Y., Tempé, J. (1974) Transfer of cytoplasmic male sterility from Raphanus sativus to Brassica oleracea. Proceedings EUCARPIA Crop Section Cruciferae **25**: 52–54.

Bellaoui, M., Gerlon, M., Pelletier, G., Budar, F. (1999) The restorer *Rfo* gene acts post-translstionally on the stability of the ORF138 Ogura CMS-associated protein in reproductive tissues of rapeseed cybrids. Plant Mol Biol. **40**:893-902

Bentolila S, Alfonso AA, Hanson MR. (2002) A pentatricopeptide repeat-containing gene restores fertility to cytoplasmic male-sterile plants. Proc Natl Acad Sci U S A. **99**:10887-92.

Bonhomme, S., Budar, F., Lancelin, D., Small, I., Defrance, MC., Peltier, G. (1992) Sequence and transcript analysis of the *Nco2.5* Ogura-specific fragment correlated with cytoplasmic male sterility in *Brassica* cybrids. Mol Gen Genet **235**:340-348.

Budar, F., Pelletier, G. (2001) Male sterility in plants: occurrence, determinism, significance and use. C R Acad Sci III **324**:543-50.

Brown, G.G., Formanová, 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:**262-72.

Cheung, W.Y., Champagne, G., Hubert, N., Tulsieram, L., Charne, D., Patel, J. and Landry, B.S. (1997) Conservation of S-locus for self-incompatibility in Brassica napus (L.) and Brassica oleracea (L.) Theor. Appl. Genet. **95:** 73-82.

Claros MG, Vincens P (1996) Computational method to predict mitochondrially imported proteins and their targeting sequences. Eur J Biochem **241:** 779-786

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.

Delourme, R., Foisset, N., Horvais, R., Barret, P., Champagne, G., Cheung, W.Y., Landry, B.S. and Renard, M. (1998) Characterisation of the radish introgression carrying the Rfo restorer gene for the Ogu-INRA cytoplasmic male sterility in rapeseed (Brassica napus L.) Theor. Appl. Genet. **97**: 129-134.

Fu T.D. (1981) Production and research on rapeseed in the Republic in China. Crucifeae Newsl. **6**: 6-7.

Grelon, M., Budar, F., Bonhomme, S., Pelletie, G. (1994) Ogura cytoplasmic male-sterility(CMS)-associated *orf138* is translated into a mitochondrial membrane polypeptide in male-sterile *Brassica* cybrids. Mol Gen Genet **243**: 540-547.

Heyn, F.W. (1976) Transfer of restorer genes from *Raphanus* to cytoplasmic male sterile *Brassica napus*. Cruciferae Newsl **1**: 15-16.

Iwabuchi M., Kyozuka J., Shimamoto K. (1993) Processing followed by complete editing of an altered mitochondrial atp6 RNA restores fertility of cytoplasmic male sterile rice. EMBO J. **12**:1437–1446

Iwabuchi, M., Koizuka, N., Fujimoto, H. Sakai, T., Imamura, J. (1999) Identification and expresion of the kosena radish (*Raphanus sativus* cv.Kosena) homologue of the ogura radish CMS-associated gene, orf138. Plant molecular Biology **39**: 183-188.

Jean, M., Brown, G. G., 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

Kadowaki, K., Suzuki, T., 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.

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

Koizuka N, Imai R, Fujimoto H, Hayakawa T, Kimura Y, Kohno-Murase J, Sakai T, Kawasaki S, Imamura J. (2003) Genetic characterization of a pentatricopeptide repeat protein gene, orf687, that restores fertility in the cytoplasmic male-sterile Kosena radish. Plant J. 34:407-15.

Konieczny A, Ausubel FM. (1993) A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers Plant J. **4**:403-410.

Koizuka, N., Imai, R., Iwabuchi, M., Sakai, T. and Imamura, J. (2000) Genetic analysis of fertility restoration and accumulation of ORF125 mitochondrial protein in the Kosena radish (Raphanus sativus. L. cv. Kosena) and a Brassica napus restorer line. Theor. Appl. Genet. **100**: 949-955.

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

Krishnasamy, S. and Makaroff, C.A. (1994) Organ-specific reduction in the abundance of a mitochondrial protein accompanies fertility restoration in cytoplasmic male-sterile radish. Plant Mol. Biol. **32**: 879–890.

Lander,E., Botstein, D. (1989) Mapping Mendelian Factors Underlying Quantitative Traits Using RFLP Linkage Maps. Genetics **121**: 185-199.

L'Homme Y, Stahl RJ, Li XQ, Hameed A, Brown GG. (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. **3**: 325-335.

Li XQ, Jean M, Landry BS, Brown GG. (1998) Restorer genes for different forms of brassica cytoplasmic male sterility map to a single nuclear locus that modifies transcripts of several mitochondrial genes. Proc Natl Acad Sci U S A. **95**:10032-7.

Liu, F., Cui, X., Horner, H.T., Weiner, H. and Schnable, P.S. (2001) Mitochondrial aldehyde dehydrogenase activity is required for male fertility in maize. Plant Cell, **13**: 1063-1078.

Nivison, HT, Hanson MR. (1989) Identification of a mitochondrial protein associated with cytoplasmic male sterility in petunia. Plant Cell **1**: 1121–1130

Ogura H (1968) Studies on the new male sterility in Japanese radish with special reference to the utilization of this sterility towards the practical raising of hybrids seeds Mem Fac Agric Kagosima Univ **6**: 39-78.

Paran I, Michelmore RW (1993) Development of reliable PCR-based markers linked to downey mildew resistance genes in lettuce. Theor Appl Genet **85**:985–993

Pelletier, G., Primard, C., Vedel, F., Chétrit, P., Rémy, R., Rousselle, P. and Renard, M (1983) Intergeneric cytoplasmic hybridization in Cruciferae by protoplast fusion. Mol. Gen. Genet. **191**: 244–250

Pruitt KD, Hanson MR. (1991) Transcription of the Petunia mitochondrial CMS-associated Pcf locus in male sterile and fertility-restored lines. Mol Gen Genet. **227**:348-355.

Sakai T., Imamura J., (1992) Alteration of mitochondrial genomes containing atpA genes in the sexual progeny of cybrids between Raphanus sativus cms line and Brassica napus cv. Westar. Theor Appl Genet **84**:923–929.

Sakai T, Liu HJ, Iwabuchi M, Kohno-Murase J, Imamura J (1996) Introduction of a gene from fertility restored radish (Raphanus sativus) into Brassica napus by fusion of X-irradiated protoplasts from a radish restorer line and iodoacetamide-treated protoplasts from a cytoplasmic male-sterile cybrid of B. napus. Theor Appl Genet **93**:373–379

Schnable, P., Wise, R.P. (1998) The molecular basis of cytoplasmic male sterility and fertility restoration. Trends Plant Sci. **3**:175-180.

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

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.

Young EG, Hanson MR. (1987) A fused mitochondrial gene associated with cytoplasmic male sterility is developmentally regulated. Cell **50**:41-49

Zhang, J. and Stewart, J.McD. (2000) Economical and rapid method for extracting cotton genomic DNA. J. Cotton Sci. **4**: 193-201.