Mitochondrial Genetics of Alloplasmic Male-Sterile *Brassica napus* Lines

Matti Leino
*Faculty of Natural Resources and Agricultural Sciences*
*Department of Plant Biology and Forest Genetics*
*Uppsala*

Doctoral thesis
Swedish University of Agricultural Sciences
Uppsala 2005
Reliable and economical production of hybrid (F₁) varieties requires efficient means to control pollination. The predominant system for pollination control of most field crops today is cytoplasmic male sterility (CMS). Since CMS is a consequence of disturbed nuclear-mitochondrial interactions, the CMS trait is maternally inherited. A common way to produce CMS plants is to combine the nuclear genome from one species with the mitochondria from another; plants of such origin are described as alloplasmic CMS. This thesis describes the production and molecular characterization of a novel alloplasmic CMS system produced from rapeseed (Brassica napus) (+) Arabidopsis thaliana somatic hybrids.

A population of 170 B. napus (+) A. thaliana somatic hybrid lines was backcrossed to B. napus and 22 lines with male sterility and/or aberrant flower morphology were found. Nine of these were analysed for RFLP and found to contain nuclear and plastid DNA from B. napus, whilst the mitochondrial genomes were frequently recombined containing both B. napus and A. thaliana DNA. Besides the male-sterile trait, the majority of the lines displayed homeotic conversions of anthers to carpeloid organs. Vegetative growth was reduced to some extent in some of the lines. One CMS line segregated both fertile and sterile plants. Fertility co-segregated with molecular markers from A. thaliana chr III. By a dihaploidisation strategy we produced plants with a disomic addition of A. thaliana chr III, which had a stable inheritance of the fertile phenotype.

Global mitochondrial mRNA expression was analysed in the fertility-restored line, the corresponding CMS line, B. napus and A. thaliana. Run-on experiments showed that transcriptional activities were highly variable between B. napus and A. thaliana and that in the CMS line transcriptional activity was reduced for several ribosomal protein genes and increased for orf139. Steady-state levels were more homogenous in comparison to transcriptional activities showing that RNA turnover is an important regulatory mechanism. Increased transcript abundance of several genes was observed in the CMS line, often correlated with presence of longer transcripts.

Transcripts of three A. thaliana loci, orf139, orf240a and orf294 accumulated in the CMS line, but with reduced levels in the restored line. The orf139 and orf294 transcripts accumulated differentially in a tissue and genotype-specific manner, while orf240a was constitutively expressed throughout the plant. Both orf240a and orf294 transcripts can be polyadenylated, thus providing an explanation for their post-transcriptional regulation. Segregation analysis of sterile and fertile alloplasmic lines indicates that orf139 and orf240a are less likely candidates to be responsible for the male-sterile phenotype, whereas the orf294 can be CMS-associated. However, it is likely that more than one locus in the A. thaliana mt-DNA could encode CMS in the nuclear background of B. napus.

Keywords: Arabidopsis thaliana, rapeseed, cytoplasmic male sterility (CMS), restorer of fertility, cybrids, mitochondrial gene expression, RNA stability, mitochondrial open reading frames (orf), species-specific factors.

Author’s address: Matti Leino, Department of Plant Biology and Forest Genetics, Swedish University of Agricultural Sciences. Box 7080, S-750 07 UPPSALA, Sweden.
### Abbreviations

List of selected abbreviations used in the text:

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>bp</td>
<td>basepairs</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>CHA</td>
<td>chemical hybridizing agents</td>
</tr>
<tr>
<td>chr</td>
<td>chromosome</td>
</tr>
<tr>
<td>CMS</td>
<td>cytoplasmic male sterility</td>
</tr>
<tr>
<td>CNM</td>
<td>conserved nonanucleotide motif</td>
</tr>
<tr>
<td>cRT-PCR</td>
<td>circular RT-PCR</td>
</tr>
<tr>
<td>Fr</td>
<td>fertility-restorer</td>
</tr>
<tr>
<td>GISH</td>
<td>genomic in situ hybridisation</td>
</tr>
<tr>
<td>GMS</td>
<td>genetic/genic male sterility</td>
</tr>
<tr>
<td>kb</td>
<td>kilo basepairs</td>
</tr>
<tr>
<td>kDa</td>
<td>kilo Dalton</td>
</tr>
<tr>
<td>ms</td>
<td>male-sterility</td>
</tr>
<tr>
<td>mt</td>
<td>mitochondrial</td>
</tr>
<tr>
<td>NMS</td>
<td>nuclear male sterility</td>
</tr>
<tr>
<td>nt</td>
<td>nucleotides</td>
</tr>
<tr>
<td>orf</td>
<td>open reading frame</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PPR</td>
<td>pentatricopeptide repeat</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>reverse transcriptase PCR</td>
</tr>
<tr>
<td>Rf</td>
<td>restorer of fertility</td>
</tr>
<tr>
<td>RFLP</td>
<td>restriction fragment length polymorphism</td>
</tr>
<tr>
<td>rRNA</td>
<td>ribosomal RNA</td>
</tr>
<tr>
<td>SSLP</td>
<td>simple sequence length polymorphism</td>
</tr>
<tr>
<td>tRNA</td>
<td>transfer RNA</td>
</tr>
<tr>
<td>UTR</td>
<td>untranslated region</td>
</tr>
</tbody>
</table>
Contents

Introduction, 7
Fertility regulating mechanisms in plant breeding, 7
  Production of hybrid seed, 7
  Manual, mechanical and chemical pollination control, 8
  Genetical pollination control, 8
Plant mitochondrial genetics, 11
  Genome structure and content, 11
  Regulation of gene expression, 15
Cytoplasmic male sterility, 16
  CMS-associated genes, 16
  Nuclear fertility-restorer genes, 19
  Mechanisms of CMS, 20

Aims of the study, 21

Results and discussion, 22
Establishment of a CMS-system (I, II), 22
Mitochondrial DNA recombination (I, III), 24
Species-specific control of gene expression (III), 26
  Transcriptional activity is genome-specific and nuclearly regulated, 26
  RNA turnover regulates transcript steady-state levels, 27
Characterization of putative CMS-associated genes (III, IV), 27

Conclusions, 30

Future perspectives, 31

References, 33

Acknowledgements, 42
Appendix

Papers I-IV

The present thesis is based on the following papers, which will be referred to by their Roman numerals:


IV. Leino, M., Sundström, J., Landgren, M. & Glimelius, K. 2005. Expression analysis of the Arabidopsis mitochondrial *orf139, orf240a* and *orf294* in different tissues of male-sterile and fertile alloplasmic *Brassica napus* lines. (manuscript)

Paper I, II and III are reproduced by permission of the journals concerned.
Introduction

The relationship between pollen and the fertilization of flowers for formation of grains and fruit has probably been known since the early establishment of agriculture. In parallel with the development of more sophisticated agricultural techniques, ways to control pollination have been sought, in order to produce more desirable agricultural products. By 2000 BC the Assyrians practiced artificial pollination of date palm trees. The first report of plants with inhibited pollen production is probably Köhreuter’s (1763), in which bisexual flowers were described with aborted anthers. In 1877 Darwin proposed male sterility as a strategy to force outbreeding. That male sterility could be the consequence of cytoplasmic genes – “plasmon” was suggested in the beginning of the 20th century when the first examples of non-Mendelian inheritance were described. By studying populations of Satureja, Silene and Plantago, Correns (1904, 1906) demonstrated that gynodioecy could be maternally inherited. The phenomenon of cytoplasmic male sterility (CMS) has been widely exploited for pollination control in the production of hybrid seed since the first report of its use for production of hybrid onion (Jones and Clarke, 1943).

With the development of tools for molecular genetics that arose in the 1970’s and 1980’s plant development could be examined on a molecular level. Not the least, the development and differentiation of flowers and flower organs in higher plants have been surveyed thoroughly lately. By the identification of meristem identity-genes, ABC-genes and ABC-regulators (reviewed by Coen and Meyerowitz, 1991; Theissen, 2001; Theissen and Saedler, 2001; Lohman & Weigel, 2002) a basic understanding of the genetic control of floral patterning has now been obtained. However, several important pieces of the puzzle are still missing, there among the impact of cytoplasmic genes.

Fertility regulating mechanisms in plant breeding

Production of hybrid seed

Heterosis, or hybrid vigour, was discovered and exploited by corn breeders over a century ago (Shull, 1908). The phenomenon is defined as the increased productivity, speed of development and fertility resulting from differences in parental gametes, although the molecular basis of heterosis is still a matter of debate (reviewed by Birchler et al., 2003). By controlled crosses of inbred parental lines, hybrid (F₁) varieties are produced. The hybrid varieties combine the two desired traits of heterosis and heterogeneity, that otherwise are conflicting breeding objectives when breeding open-pollinating crops. Hybrid varieties are nowadays produced in a large number of crop species, e.g., corn, sorghum, pearl millet, wheat, sunflower, cotton, oil palm, rice, sugar beet, rapeseed and numerous vegetable crops (Maunder, 1999). Reliable and economical hybrid seed production depends on three biological requirements: 1) presence of hybrid vigour, 2) prevention of self-pollination of the female parent, 3) adequate pollination by the male parent (Wright, 1980). The latter two requirements can be referred to as pollination control, for which several different practices have been developed.
Manual, mechanical and chemical pollination control

The simplest way to prevent self-pollination of the female parent is to remove the pollen bearing organs of the plant. This procedure is feasible in monoecious species, like corn and cucumbers, where male and female flowers are positioned on separate parts of the plant. In corn the procedure, called detasseling, can even be mechanised (Wright, 1980). For crop plants with bisexual flowers manual emasculation, in this case removal of anthers, is extremely labourous. Nevertheless, manual emasculation is practised for hybrid seed production of several horticultural crops, e.g. tomato (Kalloo, 1993). Induction of male sterility by chemicals, referred to as gametocides, pollen suppressants or chemical hybridizing agents (CHA) has been known for over 50 years (Moore, 1950). In spite of some use in rice and wheat hybrid production, CHAs are only marginally used, mainly due to incomplete male sterility or to severe damage to the rest of the plant (Tu & Banga, 1998).

Genetical pollination control

In light of the above mentioned obstacles to prevention of self-pollination by manual, mechanical or chemical measures a better alternative is provided by genetical pollination control. This can be classified into two major categories, self-incompatibility and male sterility. Self-incompatibility acts through the inhibition of pollen germination or pollen growth in the style of carpels of the same plant from which pollen is derived. The incompatibility can either be gametophytic where the pollen ability to grow through the stigma is determined by the haploid genotype of the pollen grain, or sporophytic where the ability of pollen to penetrate the stigma is determined by the genotype of the pollen parent. Self-incompatibility is regulated by multiple alleles of single loci. Lately, several of the genes responsible for the reaction have been identified (reviewed by Charlesworth, 2000; Stone and Goring, 2001). Surprisingly, a common mechanism between plant families seems not to be the case. In theory, self-incompatibility provide the breeder with an ideal hybrid system since each plant can serve both as a pollen donor and as a producer of hybrid seed resulting from pollination by another plant. However, the system is not widely used, mainly due to significant difficulties in stabilising and maintaining the parental lines. Exceptions include some vegetable species belonging to Brassica oleracea (Gray, 1993; Ockendon & Smith, 1993; Chiang et al., 1993; Crisp & Tapsell, 1993).

Male sterility is defined as the inability to produce functional anthers and/or pollen. The trait can be regulated either by nuclear or cytoplasmic genes. Nuclear male sterility (NMS), also termed genic or genetic male sterility (GMS) is usually caused by recessive alleles (ms). Mutations causing NMS are commonly found in nature and a large number of male sterility genes have been found in economically important genera, such as Zea, Lycopersicon, Hordeum, Pisum, Capsicum, Gossypium, Glycine and Oryza (reviewed by Horner and Palmer, 1995). The drawback with NMS is the impossibility to create 100% male sterile populations. The male sterile plants must be maintained by crossings with an isogenic line heterozygous for the male sterility gene resulting in offspring populations of 50%
fertile and 50% sterile plants. Thus, in seed production plots the fertile plants have to be rogued out before anthesis (fig 1a). Unless the male sterile gene is linked to some kind of marker gene, these plants are difficult to identify. Even though several types of phenotypic markers, conditional lethal markers and partial male sterilities sensitive to environmental conditions have been produced (reviewed by Krishna Rao et al., 1990) NMS is seldomly used in field scales. An important exception is NMS rice where the male sterility trait is linked to photoperiod-sensitivity (Yuan et al., 1993).

A special form of NMS has been accomplished by genetic engineering. In 1990 Mariani et al. produced male sterile Brassica napus plants by fusing a tapetum specific promoter to an RNase gene – Barnase from Bacillus amyloliquefaciens or RNase T1 from Aspergillus oryzae. By linking the male sterility gene to a selectable marker, in this case the bar gene that confers resistance to ammonium-glufoxinate, plants not expressing the transgene can be eliminated. Thus, by treating the segregating 1:1 population of sterile and fertile plants that occur after crossing with the fertile maintainer line with the herbicide a pure population of sterile plants can be obtained. The system was further developed by the introduction of fertility restorer lines (Mariani et al., 1992). These lines carry the Barstar gene under control of the same tapetum specific promoter. The Barstar gene, that is an RNase inhibitor, has the ability to counteract the effect of the Barnase gene so that fertility is restored. The utilisation of the barnase-barstar system in hybrid seed production is illustrated in figure 1b. Male-sterile transgenes developed in other systems promise that the techniques can be applied in more crop species (reviewed by Williams, 1995).

Male sterility can also be caused by cytoplasmic genes and is then termed cytoplasmic male sterility (CMS). CMS can result from spontaneous mutations (autoplasmic CMS) or, more commonly, by the combination of the nucleus from one species with the cytoplasm from another (alloplasmic CMS). CMS has been demonstrated in over 150 plant species (Kaul, 1988). As the cytoplasm in most species is inherited via the eggcell and not from pollen the trait is maternally inherited. This makes CMS extra valuable for production of hybrid seed since crossings with the maintainer line result in 100% sterile plants. CMS was first exploited in the production of hybrid onion (Jones and Clarke, 1943) and has since been extensively used in a number of crop species, including corn, sorghum, pearl millet, sugar beet, sunflower, rice and carrot.

Fertility can be restored by the introgression of specific nuclear genes to plants with CMS-inducing cytoplasm. This is sometimes referred to as genetic-cytoplasmic male sterility. The nuclear fertility-restoring genes are alternatively symbolised by Fr (for fertility-restorer) or Rf (for restorer of fertility). For production of hybrid seed in species where the hybrid plants must be fertile, e.g. cereals, the male parent must carry an Rf allele. The system to use CMS and Rf genes for hybrid seed production is illustrated in figure 1c. In all CMS-inducing cytoplasms investigated so far, the CMS trait has been associated with the mitochondrial genome.
Figure 1. Use of different types of male sterility in hybrid seed production. A and B represents the gametes of two different parental lines. A) Use of NMS governed by the recessive male sterility gene $ms$. B) Use of the transgenic Barnase/Barstar system. C) Use of CMS. The genotype of the cytoplasm, $N=$ normal and $S=$ sterile, and the presence of dominant restorer of fertility, $Rf$, genes are indicated.
Plant mitochondrial genetics

Genome structure and content

The genomes of higher plant mitochondria differ significantly from those of other organisms due to their large size, from 222 kb of the *B. napus* mt-genome (Handa, 2003) to some species within the Cucurbitaceae where mt-genomes appear to extend to over 2000 kb (Ward *et al.*, 1981). With the complete mitochondrial genome sequence now available from *A. thaliana* (Unseld *et al*., 1997), *B. napus* (Handa, 2003), sugar beet – *Beta vulgaris* (Kubo *et al*., 2000), rice – *Oryza sativa* (Notsu *et al*., 2002) and corn – *Zea mays* (Clifton *et al*., 2004) some general features of plant mitochondrial genomes can be listed.

First, gene density must be considered very low. For example, the identified genes of *A. thaliana* account for only 10% of the genome. Introns, duplications, integrations of DNA of nuclear and plastid origin and large unidentified open reading frames account for another 30%. Consequently, 60% of the genome has no known function (Unseld *et al*., 1997). Second, the gene content (table 1) is highly conserved, although some individual genes have been transferred to the nucleus in some species. Third, the identified genes are often ordered in clusters enabling co-transcription from the same promoter. This clustering has also been experimentally shown for some clusters in *B. napus* (Itani & Handa, 1998) and *A. thaliana* (Brandt *et al*., 1992). However, the gene clusters are only occasionally conserved between species and the order among clusters usually shifts completely (fig 2). Fourth, each genome contains internal repeats. These repeated sequences could be responsible for recombination and the formation of subgenomic molecules. For example, a tricircular structure is suggested for *B. napus* (Palmer & Shields, 1984) and five circular molecules can be formed in *A. thaliana* (Klein *et al*., 1994).

All five species have an identical set of 18 genes coding for subunits of complex I, III, IV and V. A pseudogene for the complexII gene *sdh4* is found in *A. thaliana* and *B. napus*, but not in the other species. In the liverwort *Marchantia polymorpha* this gene is intact (Oda *et al*., 1992). The conserved *orf25* and *orfB* have recently been identified as the ATPsynthase subunits *atp4* (Heazlewood *et al*., 2003) and *atp8* (Heazlewood *et al*., 2003; Sabar *et al*., 2003). Furthermore, six genes for cytochrome c biogenesis are present, some of which have been transferred to the nucleus in some species. Likewise, of the total of 14 genes encoding ribosomal subunits among these mt-genomes, each species has a different set. Noteworthy is the *rps14* gene that is functional in *B. napus*, but only a pseudogene in *A. thaliana*. Several tRNA genes are present, but not always the same ones. In none of the species does the set of tRNA genes account for all 20 amino acids found in proteins. The remaining tRNAs must be imported from the nucleus or in rare cases from the chloroplast (see e.g. Dietrich *et al*., 1996). The three ribosomal RNA genes *rrn26*, *rrn18* and *rrn5* are found in all genomes, with the last two always in close physical association. Two additional genes, the maturase gene *mat-R* and *orfX* (*mttB*), proposed to encode a transporter protein (Bonnard and Grienenberger, 1995) are also found in all species. In conclusion, plant mitochondrial genomes vary significantly in structure, size and sequence, but share a set of very well conserved genes.
Table 1. Comparison of gene content in higher plant mitochondrial genomes, *Arabidopsis thaliana* (A. t.), rapeseed – *Brassica napus* (B. n.), sugar beet – *Beta vulgaris* (B. v.), rice – *Oryza sativa* (O. s.) and corn NB – *Zea mays* (Z. m.)

<table>
<thead>
<tr>
<th>Gene</th>
<th>A. t.</th>
<th>B. n.</th>
<th>B. v.</th>
<th>O. s.</th>
<th>Z. m.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NADH dehydrogenases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Complex I)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>nad1</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>nad2</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>nad3</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>nad4</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>nad4L</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>nad5</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>nad6</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>nad7</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>nad9</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><strong>Succinate dehydrogenase</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Complex II)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>sdhA</em></td>
<td>Ψ</td>
<td>Ψ</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Apocytochrome b</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Complex III)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>cob</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><strong>Cytochrome oxidase</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Complex IV)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>cox1</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>cox2</em></td>
<td>+</td>
<td>+ (2)</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>cox3</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><strong>ATP synthase</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Complex V)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>atp1</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+ (2)</td>
</tr>
<tr>
<td><em>atp4</em> (orf25)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>atp6</em></td>
<td>+</td>
<td>+ (2)</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>atp8</em> (orfB)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>atp9</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><strong>Cytochrome c biogenesis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>ccmB</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>ccmC</em></td>
<td>+</td>
<td>+</td>
<td>Ψ</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>ccmFN</em></td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>ccmFN1</em></td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>ccmFN2</em></td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>ccmFC</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><strong>Ribosomal proteins</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(small subunit)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>rps1</em></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>rps2A</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>rps2B</em></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>rps3</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>rps4</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>rps7</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>rps11</em></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Ψ</td>
<td>-</td>
</tr>
<tr>
<td><em>rps12</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>rps13</em></td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><em>rps14</em></td>
<td>Ψ</td>
<td>+</td>
<td>-</td>
<td>Ψ</td>
<td>-</td>
</tr>
<tr>
<td><em>rps19</em></td>
<td>Ψ</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gene</td>
<td>A. t.</td>
<td>B. n.</td>
<td>B. v.</td>
<td>O. s.</td>
<td>Z. m.</td>
</tr>
<tr>
<td>----------------------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td><strong>Ribosomal proteins</strong> (large subunit)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rpl2</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>rpl5</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>rpl16</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><strong>Other proteins</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mat-R</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>orfX (mmtB)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><strong>Ribosomal RNAs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rrs5</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>rrs18</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>rrs26</td>
<td>+</td>
<td>+</td>
<td>+ (3)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><strong>Transfer RNAs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>trnA</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>- (Ψ)</td>
</tr>
<tr>
<td>trnR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>- (Ψ)</td>
<td>(Ψ)</td>
</tr>
<tr>
<td>trnN</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+ (2)</td>
</tr>
<tr>
<td>trnD</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+ (2)</td>
</tr>
<tr>
<td>trnC</td>
<td>+</td>
<td>+</td>
<td>+ (2)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>trnE</td>
<td>+</td>
<td>+</td>
<td>+ (2)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>trnQ</td>
<td>+</td>
<td>+</td>
<td>+ (2)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>trnG</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>trnH</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+ (2)</td>
<td>+</td>
</tr>
<tr>
<td>trnI</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+ (3)</td>
</tr>
<tr>
<td>trnK</td>
<td>+ (2)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>trnL</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>- (Ψ)</td>
</tr>
<tr>
<td>trnM</td>
<td>+ (2)</td>
<td>+</td>
<td>+</td>
<td>+ (3)</td>
<td>+</td>
</tr>
<tr>
<td>trnM</td>
<td>+</td>
<td>+</td>
<td>+ (4)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>trnP</td>
<td>+</td>
<td>+</td>
<td>+ (2)</td>
<td>+ (2)</td>
<td>+ (2)</td>
</tr>
<tr>
<td>trnS</td>
<td>+ (5)</td>
<td>+ (2)</td>
<td>+ (3)</td>
<td>+ (3)</td>
<td>+ (2)</td>
</tr>
<tr>
<td>trnW</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+ (2)</td>
<td>+</td>
</tr>
<tr>
<td>trnY</td>
<td>+ (2)</td>
<td>+</td>
<td>+</td>
<td>+ (2)</td>
<td>+</td>
</tr>
<tr>
<td>trnV</td>
<td>-</td>
<td>-</td>
<td>- (Ψ)</td>
<td>- (Ψ)</td>
<td>- (Ψ)</td>
</tr>
</tbody>
</table>

* trans-spliced. (+) present, (-) absent, (Ψ) pseudogene. Numbers in parenthesis indicate if more than one gene copy exists.
Figure 2. Circular maps of higher plant mt-genomes. For simplicity only protein coding genes and rRNA genes are indicated and trans-spliced exons omitted.
Regulation of gene expression

In addition to, or perhaps as a result of the complex mt-genomes in plants, regulation of gene expression has unique features. The multiple steps of transcriptional, post-transcriptional and translational control have been reviewed by Binder et al. (1996), Giegé and Brennicke (2001), Hoffmann et al. (2001) and Binder and Brennicke (2003).

The large physical distances between genes suggest that several promoters are required for transcription. From run-on experiments (Finnegan & Brown, 1990; Mulligan et al., 1991; Muise and Hauswirth, 1992; Giegé et al., 2000) it is also clear that promoters have different strength, since differences in transcriptional activity are found for individual genes. By in vitro capping experiments with the enzyme guanylyltransferase, transcription initiation sites and corresponding promoter sequences can be identified (see e.g. Covello & Gray, 1991; Brown et al., 1991; Binder & Brennicke, 1993). A motif of four nucleotides, CRTA, where R represents either a guanine or an adenine, is common to promoters identified this way in both monocots and dicots. By in vitro transcription studies the importance of nucleotides surrounding this motif has been further investigated (Binder et al., 1995; Dombrowski et al., 1999; Hoffmann & Binder, 2002). The region from at least 14 bases upstream and 4 bases downstream of the transcription initiation site seems to be involved. The combined in vitro capping and in vitro transcription studies have defined a dicot consensus nonanucleotide motif (CNM) promoter sequence, CRTAAGAGA. Searches for this sequence (including some variations) in the mt-genomic sequences of A. thaliana (Dombrowski et al., 1998) and B. napus (Handa, 2003) resulted in two major conclusions: About half of the genes lacked this promoter motif and the locations of the CNM for specific genes are not always conserved between the two species. Consequently, there must be other, yet unidentified, promoter motifs. Recently, Kühn et al. (2005) identified a large number of A. thaliana promoters which differ from the CNM. Additionally, they clearly showed that multiple promoters for a single gene are often present. Although RNApolymerases in plant mitochondria have been identified (Hedke et al., 1997, 2000) cofactors mediating transcription initiation are still unknown. Cofactors have been shown to be species-specific in corn/teosinte (Newton et al., 1995) and Nicotiana (Edqvist & Bergman, 2002). In both these cases particular promoters require a specific nuclear background for transcript initiation.

Transcription is only the first step in determining the steady-state population of translatable RNA. Equally important are post-transcriptional events. For example, transcript mapping shows that 5´ and 3´ processing most often takes place in plant mitochondrial transcripts. The enzymes responsible for processing are currently unknown, although conserved sequence elements marking processing sites have been suggested by Schuster & Brennicke (1989). In transcript 3´ends stem-loop structures are regularly found. In the pea atp9 transcripts, such structures were found to act as processing signals by impeding exonucleolytic degradation from 3´ends (Dombrowski et al., 1997). Stem-loop structures in Ogura CMS orf138 transcript 3´regions were found to play roles both as processing and stabilizing signals (Bellaoui et al., 1997). Recently it has been discovered that
polyadenylation of 3’ends results triggers degradation by a 3’ to 5’-exoribonuclease activity. (Gagliardi and Leaver, 1999; Lupold et al., 1999; Gagliardi et al., 2001; Kuhn et al., 2001; Perrin et al., 2004a; Perrin et al., 2004b). In A. thaliana two exoribonucleases, AtmTPNPase and AtRNaseII, have been cloned (Perrin et al., 2004a). Mutants of AtmTPNPase result in accumulation of pretranscripts and polyadenylated transcripts (Perrin et al., 2004a; 2004b).

Group II introns are found in many protein coding genes but vary from species to species (Binder et al., 1996). The splicosome and potential cofactors have not been identified. Trans-splicing of the exons of nad1, nad2 and nad5, spread all over the genome, occur in every flowering plant hitherto examined (Malek et al., 1997). A unique feature of plant mitochondrial gene expression is RNA editing, first detected by Covello & Gray (1989), Hiesel et al. (1989) and Gualberto et al. (1989). The editing events change Cs to Us in the RNA. Giegé & Brennicke (1999) undertook a global investigation of editing sites in A. thaliana mitochondria. A total of 441 editing sites was found, mostly located in coding regions, some in introns and leading or trailer sequences, but extremely rarely found elsewhere. In a similar investigation performed in B. napus Handa (2003) discovered 427 sites, 81% of which were common for both species. Compared to the average DNA sequence similarity for protein coding genes of 99.2%, diversification of editing seem higher than coding information. In neither of these surveys of editing sites, could a common sequence element for editing be found. Electroporation experiments indicate that cis-elements, 20 nt upstream of the editing site could have a function (Farré et al., 2001). The role of editing for gene regulation has been extensively examined. The only clear role of editing to date is for a tRNA, encoding phenylalanine in Oenothera, which requires editing to excise from its precursor molecule (Marchfelder et al., 1996). Mitochondrial gene expression is probably also regulated on a translational level, however the mechanisms regulating translation of transcripts are virtually unknown.

**Cytoplasmic male sterility**

**CMS-associated genes**

Molecular studies of CMS plants have correlated the trait with expression of putative genes in the mitochondrial genome (reviewed by Schnable & Wise 1998; Budar & Pelletier, 2001; Budar et al., 2003; Hanson & Bentolila, 2004). Although mitochondrial CMS-associated genes have been suggested in numerous systems clear evidence for a link to the male-sterile phenotype is often lacking. However, in some cases, shown in figure 3, strong correlations between the gene and the CMS-phenotype have been demonstrated.

Of the CMS-associated loci hitherto found, some shared similarities are worth noting. First, open reading frames (orfs) are found, comprised of novel sequences of unknown origin combined with sequences of known mitochondrial genes. Examples include the Brassica pol and nap orf224 and orf222, which contain parts of the atp8 gene (Singh & Brown, 1991, Brown, 1999), as does the sunflower PET1 orf522 (Laver et al., 1991), the Sorghum A3 orf 107, which contains a piece of atp9 (Tang et al., 1996), the rice Boro orf79 (Akagi et al., 1994) and wheat
thimophevii orf256 (Rathburn & Hedcoth, 1991), which contain small fragments of cox1 and the orf13-T in corn CMS-T, which contains a small fragment from rrn26 (Dewey et al., 1986). Other loci are even more complex, such as the Brassica tour orf263 which contains parts of nad5a and atp6 (Landgren et al., 1996), the Petunia pcf (orf402) which contains parts of atp9 and cox2 (Young and Hanson, 1987), and the orf77 of corn CMS-S, which contains several small pieces from the atp9 gene (Zabala et al., 1997).

Figure 3. Genes associated with CMS and corresponding references. CMS-genes are denoted by the current convention of number of codons, with the exception of the historical convention of sunflower orf522 that rather encodes 522 nucleotides. The alternative names for common bean orf239, corn T orf115 and petunia orf402 are also indicated. Boxes indicate orfs. Blue indicates standard mt-genes. Yellow indicates CMS-associated genes with fragments of standard genes indicated by green.
A second common feature is close physical association to, and in some cases co-transcription with, standard mitochondrial genes. In fact, a common strategy for identifying CMS-associated loci is to look for aberrant transcription patterns in Northern analysis using standard mt-gene probes. Surprisingly often, CMS-associated loci are found close to ATP synthase subunit genes. This is the case for corn CMS-T, in which the *urf13-T* is positioned upstream of *atp4* (Dewey et al., 1986, 1987), the *Brassica pol orf224* positioned upstream of and co-transcribed with the *atp6* gene (Singh and Brown, 1991), the *Brassica Ogura orf138* (Bonhomme et al., 1991, 1992) and the radish Kosena *orf125* (Iwabuchi et al., 1999) which are positioned upstream of *atp8* and co-transcribed with this gene, the *Brassica tour orf263* (Landgren et al., 1996) and rice Boro *orf79* (Akagi et al., 1994) which are positioned downstream of and co-transcribed with *atp6*, and the sunflower PET1 *orf522* which is positioned downstream of and co-transcribed with *atpi* (Köhler et al., 1991). Apart from these examples, the wheat *thimophevii orf256* is located upstream of and co-transcribed with *cox1*. The *orf77* of corn CMS-S is also closely positioned to *cox1* but do not appear co-transcribed with it (Zabala et al., 1997). The *pvs* locus in *Petunia* is positioned upstream of and co-transcribed with both *nad3* and *rps12* (Pruitt and Hanson, 1991).

Besides these two common, but not compulsory, requisites CMS loci have few other features in common. The parts of the *orfs* not containing fragments from standard mt-genes have limited sequence similarity with the exception of some pairwise similarities of *Brassica pol orf224* and *nap orf 222*, the *Brassica Ogura orf138* and radish Kosena *orf125*, and the rice Boro *orf79* and sorghum A3 *orf107*.

Since efficient methods for transforming mitochondria are still lacking the ultimate proof for a role of CMS-associated genes remains to be demonstrated. Attempts to provide proof by transforming plants with CMS-associated *orfs* linked to a mitochondrial targeting sequence have been made. Chaumont et al. (1995) transformed tobacco with the *urf13* from CMS-T corn, but did not obtain male-sterile plants. Wintz et al. (1995) transformed *Petunia* with the CMS-associated *pcf* gene, but did neither obtain male-sterile plants. He et al. (1996) managed to obtain sterile tobacco plants upon transformation with the common bean *pvs* CMS-associated gene. However, male-sterility was found independent whether or not the gene was mitochondrially targeted. All in all, experiments with nuclear transformation of CMS-associated genes must be considered inconclusive.

An interesting observation from the fully sequenced mt-genomes of *A. thaliana* (Unseld et al., 1997) and corn (Clifton et al., 2004) is that mitochondrial genomes of fertile plants contain many more *orfs* with properties typical for CMS than would occur by random. In the *A. thaliana* mt-genome Marienfeld et al. (1997) found 8 *orfs* with a mosaic structure involving fragments of known genes. An interesting hypothesis (see e.g. Budar & Pelletier, 2001) is that these *orfs* can cause CMS, but are suppressed by nuclear regulators (*Rf* genes). In alloplasmic plants the putative CMS-genes are exposed to a nuclear background, which has not evolved or maintained the corresponding restorer genes. Consequently, this would allow the CMS gene and phenotype to be expressed.
Nuclear fertility-restorer genes

That CMS is the result of disturbed nuclear-mitochondrial interactions is clearly demonstrated by the restoration of flower morphology and fertility through introgression of certain nuclear restorer of fertility – Rf genes. In the case of alloplasmic CMS Rf genes are often found in the nuclear genome of the cytoplasmic donor species. Restored lines can thus be produced by introgression of parts of nuclear DNA from this species to the CMS line. Restoration has been accomplished in, for example, Nicotiana alloplasmic CMS lines (Burns et al., 1978; Gerstel et al., 1978), wheat with T. thimophevii cytoplasm (Livers, 1964) and numerous Brassica juncea interspecific hybrids (Prakash et al., 2001; Pathania et al., 2003; Banga et al., 2003).

Rf genes have been demonstrated to affect the CMS-associated genes by several different modes of action. In common bean the Fr locus has the ability to shift the stoichiometric relationship in the tripartiate mt-genome so that the molecule containing the CMS-associated pvs gene is suppressed to substoichiometric levels (Mackenzie and Chase, 1990; Janska et al., 1998; Sarria et al., 1999). Therefore, transcription of the pvs gene is indirectly reduced as there are fewer DNA molecules encoding it. This mechanism has not been reported from any other CMS-system, although a nuclear influence on mitochondrial genome structure has been reported in, for example, Nicotiana (Håkansson and Glimelius, 1991) and A. thaliana (Abdelnoor et al., 2003). An influence of Rf genes on transcriptional level is more common. In corn the non-allelic restorer genes Rf1, Rf8 and Rf* all have the ability to process the CMS-associated T-urf13 transcript (Dill et al., 1997; Wise et al., 1999). Likewise, in Brassica with pol and nap cytoplasms the allelic genes Rfp and Rfn have the ability to process the orf224-atp6 and orf222-nad5c-orf139 co-transcripts, respectively (Li et al., 1998; Menassa et al., 1999; Brown et al., 1999). The Rf3 gene in sorghum (Tang et al., 1996) and Rf3 gene of corn CMS-S (Wen & Chase, 1999) also appear to process CMS-associated transcripts. Interestingly, these Rf loci often also process other mitochondrial transcripts.

Other modes of restorer action have been reported. In a restored CMS-line of Nicotiana tabaccum with N. repanda cytoplasm the transcript initiation site of the CMS-associated locus is affected (Edqvist and Bergman, 2002). Restorer effects on editing have been reported for atp6 transcripts of rice Boro CMS (Iwabuchi et al., 1993) and sorghum A3 (Howad & Kempken, 1997) but the relationship of editing to fertility restoration is unclear. In sterile and fertility-restored lines of sunflower PET1, transcripts are identical but preferably degraded in the restored line due to polyadenylation of 3’ ends (Monéger et al., 1994; Gagliardi & Leaver, 1999). Finally, restorer genes can also act on the translational or post-translational level and inhibit accumulation of the CMS-associated protein. This action of restorer genes has been demonstrated in the similar Brassica Ogura (Bellaoui et al., 1999) and Radish/Brassica Kosena (Koizuka et al., 2000) CMS systems.

The molecular identity of Rf genes was elusive until 1996, when the Rf2 gene restoring CMS-T corn was identified as an aldehyde dehydrogenase that assembles in the mitochondrial matrix (Cui et al., 1996; Liu et al., 2001). If this gene is a true
restorer gene has been debated (Touzet, 2002; Schnable, 2002). First, \( R_f2 \) does not affect the chimeric gene or its protein. Second, \( R_f2 \) has an important physiological role also in plants with normal fertile cytoplasm, as heterozygous \( rf2 \) plants have impaired anther formation. Lately, four more restorer genes, that do affect levels of the CMS-associated protein, have been cloned in a Petunia CMS system (Bentolila \textit{et al.}, 2002), in rice Boro (Kazama and Toriyama, 2003; Komori \textit{et al.}, 2004), in Brassica Ogura (Brown \textit{et al.}, 2003; Desloire \textit{et al.}, 2003), and Raphanus/Brassica Kosena (Koizuka \textit{et al.}, 2003). The latter two are identical. All these genes bear a PPR motif (pentatricopeptide repeat), a gene family recently characterized by Lurin \textit{et al.} (2004) of organelar targeted, and probably RNA binding, proteins. In \textit{A. thaliana} 441 PPR genes have been found to date and the gene family is probably well represented also in other plant species. Interestingly, the restorer genes cloned so far, all belong to the same subgroup of PPR genes and have similar motifs.

\textit{Mechanisms of CMS}

In spite of the progress made in identifying mitochondrial CMS-loci and nuclear restorer genes, the interactions leading to male sterility remains puzzling. The translation of \textit{Brassica} Ogura \textit{orf}138 yields a 19kD protein (Grelon \textit{et al.}, 1994), the Radish Kosena \textit{orf}125 is translated into a 17 kDa protein (Iwabuchhi \textit{et al.}, 1999), the wheat \textit{thimophevii orf} 256 to a 7 kDa protein (Song & Hedgcoth, 1994), the sunflower \textit{PET1 orf}522 to a 15-16 kDa protein (Horn \textit{et al.}, 1991; Monéger \textit{et al.}, 1994), the common bean \textit{pvs} to a 27 kDa protein (Abad \textit{et al.}, 1995), the corn CMS-T \textit{orf}13 to a 13 kDa protein (Wise \textit{et al.}, 1987) and the Petunia \textit{pcf} to a 19.5 kDa protein (Nivison & Hanson, 1989). These proteins are generally hydrophobic, suggesting association with mitochondrial membranes. The membrane association has indeed been shown for the \textit{Brassica} Ogura ORF138 (Grelon \textit{et al.}, 1994), the sunflower ORF522 (Horn \textit{et al.}, 1996), and the corn CMS-T \textit{orf}13 (Dewey \textit{et al.}, 1987). Expression of URF13 is also associated with susceptibility to toxin from the corn fungal pathogen \textit{Cochliobolus heterostrophus} (reviewed by Levings, 1993). In presence of the toxin the protein forms a pore in the mitochondrial membrane. Whether this pore, is formed also in sterile anthers is not known. Interestingly, expression of a novel mitochondrial gene due to altered processing in \textit{Citrus jambhiri} also leads to membrane pore formation and susceptibility to the fungal pathogen \textit{Alternaria alternata} (Ohtani \textit{et al.}, 2002). The membrane association and/or the co-transcription with subunits of ATP synthase (Complex V) could also suggest disturbed ATP production. In an alloplasmic \textit{Nicotiana} CMS system reduced ATP/ADP ratios have been observed in flower buds in the sterile line (Bergman \textit{et al.} 2000). In the alloplasmic CMS-lines of \textit{Brassica napus} described in this thesis, Teixiera \textit{et al.} (2005a) found lower ATP levels in comparison to the fertile line. Lowered ATP synthase activity has also been reported in sunflower PET1 (Sabar \textit{et al.}, 2003). Thus, impaired energy production could be one explanation for the developmental aberrations. Balk & Leaver (2001) showed that programmed cell death (PCD) occurs in premature tapetal cells of sunflower PET1 developing anthers. The authors propose that insufficient energy access may activate the PCD-associated proteins. However, why lowered ATP levels should specifically disturb pollen development and not other energy demanding processes, like germination, remains mysterious.
CMS-associated polypeptides are generally constitutively expressed, with a noteworthy exception: In common bean, the PVS-ORF239 proteins accumulate in microspore cell walls and callose layers and seem to toxify microspore development (Abad et al., 1995; Sarria et al., 1998). Restorer genes have also been shown to act in a tissue-specific manner. For example, the Rfp restorer of Brassica pol CMS processes the orf224 more efficiently in the inner three whorls of the flower (Menassa et al., 1999), the Rfn of Brassica nap CMS reduces the orf222 in developing anthers (Geddy et al., 2005) and the orf522 of PET1 sunflower is most efficiently degraded in floral tissue (Monéger et al., 1994). Although the general feature of CMS is inhibited pollen production, a wide range of different morphological aberrations is found among different CMS systems. In some systems, such as Petunia, sunflower PET1 and corn CMS-T, only the pollen production in anthers is affected, whereas in other systems stamens undergo homeotic conversions of floral organs. In Brassica Ogura, especially grown in cold temperatures, conversions of stamens to carpeloid structures were found (Polowick & Sahwney, 1987), in Brassica juncea with Enarthrocarpus lyratus cytoplasm petaloid anthers are formed (Banga et al., 2003), in Nicotiana interspecific hybrids or cybrids a range of flower modifications occur (Gerstel et al., 1978; Kofer et al., 1991, Farbos et al., 2001), in CMS carrot both petaloid and carpeloid stamens are formed (Linke et al., 1999) and in CMS wheat pistilloid stamens are found (Murai et al., 2002). Interestingly, the homologues of the floral homeotic gene APETALA3 in A. thaliana conferring anther identity (Coen and Meyerowitz, 1991) are found affected in several CMS-systems (Murai et al., 2002; Linke et al., 2003; Geddy et al., 2005; Teixeira et al., 2005b). The link between CMS-associated genes, their products, mitochondrial function and pollen/flower formation, however, remains unclear.

Aims of the study

The overall aim of the project was to study the genetic mechanisms involved in the nuclear-mitochondrial interaction leading to CMS in Brassica napus (+) Arabidopsis thaliana somatic hybrids. Specific aims were to:

- produce stable B. napus CMS lines from a population of B. napus (+) A. thaliana somatic hybrids and to characterize genotypes and phenotypes of these lines;
- produce stable restorer of fertility lines through introgression of nuclear A. thaliana DNA and to characterize the nuclear DNA involved in fertility restoration;
- develop methods for high-throughput expression profiling of mitochondrial genes;
- analyze species-specific effects on mitochondrial transcriptional and post-transcriptional processes;
- analyze developmental and molecular regulatory mechanisms of putative CMS-associated genes.
Results and discussion

Establishment of a CMS-system (I, II)

Forsberg et al. (1998) performed protoplast fusions between *Brassica napus* and *Arabidopsis thaliana*. From these experiments approximately 200 lines derived from individual calli were obtained. After a backcross to *B. napus* a BC\(_1\) population of 170 lines was screened for plants with male sterility and/or aberrant flower morphologies. Twenty-two lines were identified. Nine lines with good vigour and seed set were selected and backcrossed further to obtain a BC\(_3\) generation. These lines were characterized with RFLP and specific nuclear and organellar probes and found to have nuclear and plastidic DNA from *B. napus*, whilst mitochondrial DNA was recombined with varying portions of *A. thaliana* and *B. napus* mtDNA.

Male-sterile plants could be classified into two major categories, one group with stamens of reduced size and limited pollen production and one group with stamens converted into feminized organs, including ovules and stigmas, that were completely unable to produce pollen. Both groups displayed reduced petal size, which was especially true for the second group. The first group shared clear morphological similarities with the *nap, pol* (Fan & Stefansson, 1986) and *tour* (Liu et al., 1996) CMS-inducing cytoplasms in *B. napus* whereas the second group resembles cold-grown plants of the original type of Ogura CMS in *B. napus* (Polowick and Sahwney, 1987). The second group was more common among the screened lines. Detailed investigations of two lines in group two (4:19 and 41:17) showed that morphological differences of flower meristems can be noticed from stage 4 and onwards during flower development (Teixeira et al., 2005b). Interestingly, the floral homeotic genes APETALA3 and PISTILLATA, responsible for proper stamen and petal development are downregulated in these lines (Teixeira et al., 2005b, Carlsson et al., manuscript in preparation). Some reduction in vegetative development and later flowering time were also observed in these lines. The phenotypes were stably maternally inherited between generations and have to date been backcrossed for 12 generations without any noticeable shift in phenotype.

In the first backcrossed generation of CMS-line 4:19, one plant with partially restored male-fertility was observed. This line was subjected to recurrent backcrosses with the maintainer line *B. napus* cv. Hanna and analysed for inheritance of the restoration trait. The restorer trait appeared not to be inherited in a Mendelian ratio, nor was it possible to stabilise the trait via selfings. In a segregating population mapping with RFLP markers was performed. These results showed that restoration co-segregated with markers covering the whole *A. thaliana* chromosome III. From these results we concluded that the restored plants possessed a monosomic addition of the foreign chromosome. Thus, we utilized a dihaploidisation strategy to stabilize the line. Haploid plants were produced from microspore cultures, screened with SSLP-markers for presence of the *A. thaliana* chromosome and doubled with colchicine. One dihaploid line had a restored phenotype that was stably inherited after selfing and by GISH we showed that this
The restored line, the CMS line, and the two parental species have served as research material (fig. 3) in this thesis. As the full mitochondrial genome sequence from both *B. napus* (Handa, 2003) and *A. thaliana* (Unseld et al., 1997) as well as the full nuclear genome sequence of *A. thaliana* (The Arabidopsis Genome Initiative, 2000) are available the material is especially suitable for mitochondrial genetic studies. In addition, the finding of several lines with different male sterile or aberrant flower morphologies and heterogenous mitochondrial genomes offers the possibility to correlate certain flower abnormalities with a certain mitochondrial DNA composition.

![Schematic presentation of the nuclear-mitochondrial composition in the plant material used in this thesis.](image)

Figure 4. Schematic presentation of the nuclear-mitochondrial composition in the plant material used in this thesis. Note that *B. napus* (maintainer line) and the CMS line are isogenic in respect of the nuclear genome, whilst the CMS and restored lines are isogenic in respect of mitochondrial genomes.
Mitochondrial DNA composition (I, III)

Nine of the CMS-lines were characterized for mitochondrial genome composition using RFLP analysis. Almost any gene probe - restriction enzyme combination resulted in polymorphism between *A. thaliana* and *B. napus*, a finding reflecting the large differences in non-coding sequence and structure between the two species. The RFLP pattern for each CMS line was compared to the parental lines to determine from which parental species a specific gene region was derived (fig 5a). This methodology is not foolproof as fragments visually estimated to be of the same size could be wrongly interpreted to be similar. However, two general conclusions could be drawn from this investigation. The CMS lines contain both *A. thaliana* and *B. napus* mtDNA and frequent recombinations have occurred. In fact none of the nine lines showed an identical RFLP patterning. Of special interest is the pattern obtained with the *cox2* gene probe, for which a very high degree of polymorphism between lines was found (fig 5b). The *cox2* gene is located in the repeat regions in *B. napus* (Handa, 2003) and close to repeat I in *A. thaliana* (Unseld et al., 1997). The repeat regions could be involved in intragenomic recombination and the formation of subcircular molecules. Whether novel subcircular molecules also formed when the two species were combined is not known, but all the same quite possible. In the presence of the *A. thaliana* chr III in the restored line a shift in genome structure in this region was observed. Nuclear influence on genome structure has also been reported in alloplasmic *Nicotiana* lines (Håkansson and Glimelius, 1991) and in CMS lines of common bean (Janska et al., 1998). In *A. thaliana* a gene influencing genome structure, CHM, has been cloned (Abdelnoor et al., 2003). This gene is indeed positioned on chr III, but it is not known whether this gene is specifically responsible for the structural shifts observed in the *cox2* region.

Novel recombined hybrid-specific fragments were found occasionally. In the investigation of the nine CMS lines with ten gene probes recombinations were found, in one or more lines, with the probes *atp1*, *atp9*, *cob*, *cox2* and *nad5a*. Three of the lines, 4:19, 41:17 and 14:103, were analysed for RFLP and 34 gene probes (III and unpublished results). Additional recombined fragments were found, in one or more lines with the probes *nad1*, *cox3*, *atp6*, *ccmB*, *rpl16*, *rps7*, *rps3*, *rps4*, *rps12* and *orfX*. In the extensive analysis of mitochondrial recombinations after somatic hybridisation between different Brassicaceae species performed by Landgren and Glimelius (1990, 1994) and Landgren et al. (1994) hotspots for recombinations were found, for example, with *atp1* and *atp9*. As suggested by Clark et al. (1986) the recombination events could preferably take place at sites of short repeated sequences. In fact, both the *B. napus* (Handa, 2003) and *A. thaliana* mt-genomes (Unseld et al., 1997) are filled with short repeats. We have not sequenced the rearranged fragments, but with the full genome sequences available this is easily feasible and could determine if a certain particular sequence motif facilitates recombination.
Figure 5. Examples of RFLP hybridisation patterns of the parental and different CMS lines obtained with different gene probes. (a) The three different CMS lines 4:19, 41:17 and 14:103 and the two parental species, *B. napus* (B) and *A. thaliana* (A). Notice presence of gene copies from either or both parental species in the CMS lines as well as recombined fragments. (b) RFLP hybridization pattern obtained with the gene probe *cox2* from eight different CMS lines. DNA was analysed from two individual plants of each line.
Species-specific control of gene expression (III)

To investigate mitochondrial gene expression and RNA turnover in the different lines a macroarray-system was set up. The system is based on the methods described by Giegé et al. (2000), but refined and supplemented by additional gene probes. Using sequence specific primers, sequences were PCR amplified from all identified protein coding and rRNA coding genes as well as several orfs in the A. thaliana mitochondrial genome. The PCR products were dotted onto nylon membranes using a pipetting robot in order to produce a high number of identical filters. The filters could then be used for DNA/RNA profiling studies by hybridization of differentially labeled nucleic acids from the set of fertile and sterile lines.

Transcriptional activity is genome-specific and nuclearly regulated

By disrupting mitochondrial preparations in the presence of [$\alpha$-$^{32}$P]G/UTP and extending previously initiated transcripts, labelled run-on transcripts were produced. These transcripts were used as probes on the filters and by quantifying the hybridisation signal from each gene the relative transcriptional activities could be estimated. In all lines the rRNA genes were the most actively transcribed. This finding is in contrast with the investigation performed by Giegé et al. (2000) in A. thaliana, who found rRNA genes no more transcriptionally active than other mitochondrial genes. However, it is in accordance with the investigations performed in corn (Finnegan and Brown, 1990; Mulligan et al., 1991; Muise and Hauswirth, 1992) in which the rRNA genes were more active than the protein coding mitochondrial genes. An important difference in the methodology used by us and by Giegé et al. (2000) is our use of RNase to remove adjacent single stranded RNA before the quantification. As mitochondrial transcripts regularly extend extensively outside the reading frames this procedure reduces the influence of differential transcript lengths, which cannot be normalized for. Additionally, we used flower bud mitochondria, whereas Giegé et al. (2000) used cell culture mitochondria. These very different plant materials may account for some of the differences observed.

Interestingly, large differences were found in transcriptional activity for individual genes between the lines, a finding which correlates well with the comparison of B. napus and A. thaliana mitochondrial genomes made by Handa (2003). Handa’s study demonstrated large structural shifts in genome structure between the two lines, although the gene content was basically identical. Consequently, promoters for a large number of the investigated genes seem to have shifted. The CMS and restored line had mainly similar transcriptional activities as expected due to their identical mitochondrial genomes. However, in comparison to the parental species, some interesting alterations in transcriptional activity were observed. Several of the ribosomal protein subunit genes had very reduced transcriptional activities and the orf139 had highly increased activity. RFLP analysis demonstrated that these genes were exclusively inherited from A. thaliana and as a consequence they would be under the control of A. thaliana-specific promoters that work differently, and in most cases less efficient, in the B. napus
nuclear background. In addition to promoter strength transcriptional activity can also be regulated by selective DNA amplification in the mitochondrial genome (Muise & Hauswirth, 1995). We could, however, only detect marginal alterations of gene copy numbers by probing the filters with end-labelled mtDNA.

**RNA turnover regulates transcript steady-state levels**

Steady-state levels of transcripts from individual genes were also investigated in order to compare with the transcriptional activities. Transcript abundance was estimated by end-labelling RNA isolated from the same mitochondrial preparations used in the run-on assays and hybridising this RNA to identical filters. From these results it was obvious that steady-state levels of RNA were not directly proportional to transcriptional activity i.e. promoter strength. Thus, RNA steady-state levels seem to be the result of transcriptional activity as well as differential degradation rates. This finding is in accordance with the studies performed by Mulligan *et al.* (1991) and Giegé *et al.* (2000) in corn and *A. thaliana* respectively, who also demonstrated the influence of post-transcriptional regulation. The steady-state levels of individual genes shifted less than transcriptional activities between *B. napus* and *A. thaliana*. In contrast, the CMS line had higher transcript abundance than the parental lines for several genes, including *nad4, cox3, atp1, atp9, ccmC, ccmFC, rpl2, rpl16, Ψsdh4, orf294 and orf240a*. Northern analysis revealed that the elevated steady-state levels often were correlated with presence of transcripts longer than those found in the parental lines. RFLP analysis revealed that the DNA for encoding these genes in the CMS line was inherited from *A. thaliana*. Although we cannot exclude the possibility that the novel transcripts are a consequence of genome rearrangements they are more likely due to alloplasmic incompatibilities of transcriptional regulation. Similar observations were reported in alloplasmic tobacco (Håkansson & Glimelius, 1991), maize (Wen & Chase, 1999) and *Brassica* (Li *et al.*, 1998). Alloplasmic-induced transcripts can result both from novel transcript initiation sites, such as demonstrated with the *cox2* gene of maize (Newton *et al.*, 1995) and the *atp1* of tobacco (Edqvist and Bergman, 2002) and from differences in processing ability, as shown for restorer effects of *T-urf13* in maize (Dill *et al.*, 1997), *nap orf222* and *pol orf224* in *B. napus* (Li *et al.*, 1998; Menassa *et al.*, 1999; Brown *et al.*, 1999) and the CMS-associated loci in sorghum (Tang *et al.*, 1996) and corn CMS-S (Wen and Chase, 1999). Furthermore, Northern analysis revealed more clearly quantitative differences between lines for low abundant transcripts. The *A. thaliana orf139, orf240a* and *orf294* were more abundant in the CMS line than in *A. thaliana* and the fertility-restored line. Thus, we focused on control of expression of these putative CMS-associated genes.

**Characterization of putative CMS-associated genes (III, IV)**

In order to characterize the post-transcriptional and developmental regulation of *orf139, orf240a* and *orf294* different analyses were performed. Mapping of the transcript ends was performed with circular RT-PCR. This method, first utilised in plant mitochondria by Kuhn & Binder (2002) utilises the ability of T4 RNA ligase to form junctions of transcript 5´and 3´ends that can be detected by PCR and
sequencing. RT-PCR on cDNA primed with dT-oligos was utilized for the mapping of putative polyadenylation sites (Gagliardi & Leaver, 1999). Additionally we performed RT-PCR and quantitative real-time RT-PCR experiments on RNA isolated from different plant organs as well as in situ hybridisation to examine the possible developmental regulation of the transcripts. Finally, to investigate whether expression of the three orfs correlated with the male-sterile phenotype we investigated a population of both sterile, semi-sterile and fertile alloplasmic lines. The results of these experiments are summarized in Table 3.

Table 3. Summary of transcript characteristics for orf139, orf240a and orf294. (A) A. thaliana, (C) CMS, (R) restored

<table>
<thead>
<tr>
<th>feature</th>
<th>orf139</th>
<th>orf240a</th>
<th>orf294</th>
</tr>
</thead>
<tbody>
<tr>
<td>transcriptional activity</td>
<td>very high (C, R)</td>
<td>high (A)</td>
<td>moderate (A, C, R)</td>
</tr>
<tr>
<td></td>
<td>high (A)</td>
<td>moderate (C, R)</td>
<td></td>
</tr>
<tr>
<td>steady-state levels</td>
<td>extremely low (A)</td>
<td>very low (A)</td>
<td>very low (A, R)</td>
</tr>
<tr>
<td></td>
<td>very low (R)</td>
<td>low (R)</td>
<td>moderate (C)</td>
</tr>
<tr>
<td></td>
<td>low (C)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5’ ends</td>
<td>identical in C, R</td>
<td>identical</td>
<td>similar in C, R</td>
</tr>
<tr>
<td></td>
<td>unknown in A</td>
<td></td>
<td>shorter in A</td>
</tr>
<tr>
<td>3’ ends</td>
<td>heterogenous (C, R)</td>
<td>identical</td>
<td>heterogenous, extend into atp1</td>
</tr>
<tr>
<td></td>
<td>unknown in A*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNM near 5’ends</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>editing</td>
<td>-</td>
<td>yes (A, C, R)</td>
<td>-</td>
</tr>
<tr>
<td>polyadenylation</td>
<td>- *</td>
<td>yes (A)</td>
<td>yes (A, C, R)</td>
</tr>
<tr>
<td>mosaic structure</td>
<td>no</td>
<td>yes, part of rps3</td>
<td>no</td>
</tr>
<tr>
<td>similarities to</td>
<td>no</td>
<td>yes, to orf222 and orf224 in nap and pol B. napus CMS</td>
<td></td>
</tr>
<tr>
<td>identified genes</td>
<td>no</td>
<td></td>
<td>no</td>
</tr>
<tr>
<td>tissue-specific</td>
<td>yes</td>
<td>-</td>
<td>yes</td>
</tr>
<tr>
<td>expression levels</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>co-segregation with</td>
<td>no</td>
<td>no</td>
<td>possibly</td>
</tr>
<tr>
<td>male-sterility</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a polyadenylated transcript end 596 nt upstream of the start codon was found in A. thaliana.

From the run-on and kinase labelling experiments it was determined that orf139 is very actively transcribed but degraded to very low steady state levels. In the CMS and restored line a common type of transcript 5’end mapped 1093 nt upstream of the start codon, close to a consensus promoter sequence. The 3’ends were more heterogenous. We thus propose that the orf139 transcripts possess relatively stable 5’ ends, but are actively degraded from 3’ ends. Unfortunately, we were unable to map the transcript ends in A. thaliana by the cRT-PCR experiment,
probably due to the low transcript abundance. However, by RT-PCR experiments with a forward primer 956 nt upstream of the start codon a polyadenylated 3’ end at a position 596 nt upstream of the start codon was found. This processing/polyA site is absent in the CMS and restored line (unpublished results) and thus mediated by species-specific factors. The orf139 does not have typical CMS features such as fragments of standard genes or co-transcription with standard genes. The finding of a fertile line among the alloplasmic lines, which expresses the orf makes it an unlikely candidate to be responsible for the male-sterile phenotype. On the other hand, other facts that pinpoint a function for this gene region are the presence of a CNM promoter, the presence of a small non-messenger RNA (Marker et al., 2002) in the 5’UTR of the longest transcripts and a tissue-specific transcript accumulation, i.e., in carpels of A. thaliana, in flower buds of the CMS line and in roots, buds and carpels of the restored line.

Orf240a has several features consistent with its being a true gene, including distinct transcripts, which appear to be constitutively transcribed, editing of the transcripts and a location far from other identified genes. Editing of transcripts is very rare in regions of non-coding mtDNA (Giegé & Brennicke, 1999). The comparison of transcriptional activity and steady-state levels indicated that the gene is less stable in A. thaliana than in the other materials. Since the cRT-PCR analysis showed that orf240a transcripts are identical in all three lines differential processing could not be the cause of stability differences. A difference in polyadenylation-mediated transcript degradation probably provides the best explanation for the differences in transcript stability between lines. Indeed, we could demonstrate the presence of polyA-tails in A. thaliana transcripts. Even though we lack quantitative data of polyadenylation in the investigated material, a hypothesis is that in the CMS line, polyadenylation is less efficient. Alternatively, polyadenylated transcripts are less efficiently degraded. Orf240a has a high predicted amino acid similarity with orf222 and orf224 in the B. napus CMS-inducing cytoplasts nap and pol. Furthermore orf240a contains a small part of the standard gene, rps3, a characteristic typical for CMS-associated genes. However, among the alloplasmic lines, one fertile line expressing orf240a was found. Although it cannot be excluded that this fertile line have retained nuclear restorer genes from A. thaliana acting on a translational level, orf240a is therefore less likely CMS-associated.

Run-on experiments and Northern analysis revealed that orf294 transcripts are both differentially processed and degraded among the lines. In the CMS and restored line transcripts are generally about 400 nt longer than in A. thaliana. The cRT-PCR analysis revealed that this difference is due to longer transcript 5’ends. One cRT-PCR clone from A. thaliana extends several hundred nt farther upstream close to a CNM promoter. This cRT-PCR clone could represent a de novo transcript with the remaining transcript ends resulting from differential processing. Several polyadenylation sites were found in transcript 3’ends. Thus, polyadenylation-mediated degradation is a possible regulatory mechanism. 3’ ends were found that extended into atp1. The possible co-transcription with atp1 is a typical CMS-associated characteristic. One interesting speculation is that orf294-atp1 co-transcripts interfere with production of ATP1 and complex V of the
respiratory chain. Reduced ATP production has been reported in this CMS system (Teixeira et al., 2005a) as well as in Nicotiana CMS (Bergman et al., 2000) and could interfere with cell proliferation in flower meristems (Farbos et al., 2001, Teixeira et al., 2005b). The developmental patterns suggest a genotype and tissue-specific regulation since accumulation seems to occur preferentially in carpels of A. thaliana, roots of the restored line and less in petals of all three lines. Additionally, in situ hybridisation showed that orf294 is expressed in early flower meristems of A. thaliana and to some extent in the CMS-line, but not in the restored line. Among the alloplasmic lines transcription of orf294 was exclusively found in sterile lines. However, it was also absent in one sterile and one semi-sterile line. This does not exclude orf294 from being CMS-associated, but implying presence of other CMS-associated genes in these sterile lines.

To summarize, orf139 and orf240a seem not to be CMS-associated, although the possibility of retained nuclear restorer genes acting on a translational level in the fertile line expressing orf139 and orf240a transcripts cannot be ignored. Nevertheless, orf139 and orf240a may well have other functions in plant mitochondria. The orf294 may be causing the CMS-phenotype. However, the exception with the sterile and semi-sterile lines not expressing orf294, implies that other CMS-inducing genes are present in these lines. Taken together these results reveal the complexity of alloplasmic induced mitochondrial gene expression aberrations and the difficulty to associate particular transcripts with the CMS-trait. The diverse range of flower phenotypes among the B. napus (+) A. thaliana alloplasmic lines indicate that several CMS-associated, or flower morphology affecting loci, in the A. thaliana mt-genome are present and differently retained in the different lines. The A. thaliana mt-genome contains 460 orfs longer than 60 codons (Unseld et al., 1997). The expression pattern, or alloplasmic effects on expression of the vast majority of these orfs remain unknown. The alloplasmic influenced accumulation of orf139, orf240a and orf294 transcripts provide examples of how putative genes, without known function, can be activated when removed from their nuclear suppressors.

Conclusions

The main conclusions from the results presented in this thesis follow.

- Introgression of A. thaliana mitochondrial DNA to the nuclear background of B. napus results in cytoplasmic male-sterile plants.
- The cytoplasmic male-sterile lines often display homeotic conversions of anthers to carpeloid structures, reduced petal size and some disturbances in vegetative growth.
- Introgression of A. thaliana chr III into CMS line 4:19 restores anther development and pollen production.
- A. thaliana and B. napus have different promoter strength for a large number of individual genes.
As demonstrated in a CMS-line, the transcriptional activity of *A. thaliana* mt-DNA is often altered in the nuclear background of *B. napus*. In particular, several ribosomal protein subunit genes have reduced transcriptional activities and *orf139* has increased transcriptional activity.

Copy-number alteration of individual mt-genes is not a major mechanism for gene expression regulation.

Post-transcriptional events largely counter-balance differences in transcriptional activity.

Novel mitochondrial transcripts were formed in the CMS line, probably due to processing aberrations.

Transcripts of several *A. thaliana* mitochondrial *orfs* accumulate in the nuclear background of *B. napus*.

In *A. thaliana* and the fertility restored line *orf240a* and *orf294* transcripts are more efficiently degraded in comparison to the CMS-line. This degradation could be polyadenylation-mediated as polyA-tails were found in *A. thaliana* *orf240a* transcripts and in *orf294* transcripts from all three lines.

Transcripts of *orf139* and *orf294* accumulate differently in mitochondria dependent on genotype and tissue.

The *orf139* and *orf240a* are less likely candidates to be responsible for the male-sterile phenotype, whereas the *orf294* can be CMS-associated. However, it is likely that more than one locus in the *A. thaliana* mt-DNA could encode CMS in the nuclear background of *B. napus*.

**Future perspectives**

This thesis describes the establishment and characterization of a *B. napus (A. thaliana)* CMS system. The system was utilized to generate information about the genetics of CMS. Additionally, the unique combination of two fully sequenced mitochondrial genomes provides possibilities to study species-specific factors regulating mitochondrial gene expression. I suggest the following avenues of research as examples of how to exploit the material for further investigations.

As visualised in the CMS-line, a large number of novel transcripts are produced as a result of alloplasmic incompatibilities, e.g. for *atp9*, *cox3* and *ccmC*. Mapping of these transcripts would indicate if insufficient processing of precursor transcripts is a general phenomenon and if this can be related to specific sequence motifs.

Alloplasmic effects on transcriptional activity were found, most obvious for several of the ribosomal protein genes. Mapping of the *de novo* transcript 5’ ends and promoter elements would indicate if these genes are regulated by a special kind of promoter motif that would be driven by species-specific transcription factors.

Even though neither *orf139*, *orf240a* or *orf294* might be causing CMS their function is still intriguing. Polysome analysis to indicate if the transcripts are translationally active would be an initial step. As a second
step production of antibodies against the putative polypeptides and
determination of the cellular localization and developmental regulation of
the proteins could be performed.

- The results indicate that more than one locus in the *A. thaliana* mt-
genome might result in CMS when transferred to the *B. napus* nuclear
background. In paper IV we used 16 sterile and 8 fertile lines. More
fertile cybrid lines are available. Mapping of the genomes with more *A.
thaliana* specific probes and a larger population of plants could reveal
other putative CMS-loci.

- Although the restorer element consists of the whole *A. thaliana* chr III,
the full sequence of it is available. Some particularly interesting genes
have already been mentioned. Transformation of the CMS-line with these
putative restorer genes and corresponding analyses of mitochondrial gene
expression and flower development would provide means to identify
genes involved in the nuclear-mitochondrial interaction resulting in CMS.
To limit the number of putative restorer genes narrowing the amount of
*A. thaliana* would be useful. However, translocations of *A. thaliana* DNA
to *B. napus* chromosomes in this material are extremely rare, and would
probably require X-ray treatment.

- The combination of *A. thaliana* mt-DNA with the *B. napus* nuclear
background obviously resulted in CMS. It would be interesting to see if
the reversed combination – *B. napus* mtDNA in *A. thaliana* nuclear
background also would result in CMS. An alternative strategy would be
to create alloplasmic *A. thaliana* CMS using a cytoplasmic donor species
more closely related to *A. thaliana*, for instance *A. lyrata*. With the
wealth of information and tools for *A. thaliana* available, an *A. thaliana*
CMS system would provide extraordinary good opportunities to study
especially the nuclear genes involved in the expression of male sterility.
References


Bonhomme, S., Budar, F., Lancelin, D., Small, I., Defrance, M.C. & Pelletier, G. 1992. Sequence and transcript analysis of the NcoI 2.5 Ogura-specific fragment correlated with


Dombrowski, S., Brennicke, A. & Binder, S. 1997. 3´-Inverted repeats in plant mitochondrial mRNAs are processing signals rather than transcription terminators. EMBO Journal 16, 5069-5076.
Farbos, I., Mours, A., Bereterbide, A. & Glimeilus, K. 2001. Defective cell proliferation in the floral meristem of alloplasmic plants of Nicotiana tabacum leads to abnormal floral organ development and male sterility. The Plant Journal 26, 131-142.
Gagliardi, D. & Leaver, C.J. 1999. Polyadenylation accelerates the degradation of the mitochondrial mRNA associated with cytoplasmic male sterility in sunflower. EMBO Journal 18, 3757-3766.
Hanson, M. R. & Bentolila, S. 2004 Interactions of mitochondrial and nuclear genes that affect male gametophyte development Plant Cell, 16, 154-169.


Landgren, M. & Glimelius, K. 1994. A high frequency of intergenomic mitochondrial recombination and an overall biased segregation of B. campestris or recombined B. campestris mitochondria were found in somatic hybrids within Brassicaceae. Theoretical and Applied Genetics 87, 854-862.

Landgren, M., Sundberg, E. & Glimelius, K. 1994. Biased mitochondrial segregation, independent of cell type used for fusion and of hybrid nuclear content, was found in Brassica napus (+) Brassica oleracea somatic hybrids. Plant Science 103, 51-57.


Sabar, M., Gagliardi, D., Balk, J. & Leaver, C.J. 2003. ORFB is a subunit of F1F0ATPsynthase: Insight into the basis of cytoplasmic male sterility in sunflower. EMBO Reports 4, 1-6.


Song, J. & Hedgcoth, C. 1994. A chimeric gene (orf256) is expressed as protein only in cytoplasmic male-sterile lines of wheat. Plant Molecular Biology 26, 535–539.


Teixeira, R.T., Farbos, I. & Glimelius, K. 2005b. Expression levels of meristem identity and homeotic genes are modified by nuclear-mitochondrial interactions in alloplasmic male-sterile lines of Brassica napus. The Plant Journal (in press).


Acknowledgements

As previously stated by Tsarouhas (2002) and Wretblad (2002) acknowledgements is a very important section. This thesis would never have been accomplished without the help from great colleagues and friends. Thus, I would like to express my sincere gratitude to all of you who have contributed to and supported this work one way or another:

First, Kristina Glimelius, my supervisor, for your advice, support and encouragement. For always finding the time for helping me. For allowing me to test new ideas, however, never letting me loose focus.

My co-supervisor Maria Landgren for your excellent advice in all kinds of technical and tactical matters. For introducing me into the field and remaining so helpful throughout the work.

My co-workers: Susanna Thyselius - what a team we made in the lab, without you not even half of this thesis would have been realised. Ingrid Eriksson, for help and advice in small and big technical matters, you turned my green fingers gradually white. Rita Teixeira and Jenny Carlsson, fellows and collaborators, it has been such a comfortable situation for me to have you ahead and behind me in the PhD succession. Jenny, I am convinced that your thesis will be as excellent as your chocolate muffins. Jens Sundström for all good advice and your enthusiasm for novel smashing ideas. Gun Rönnqvist for caring about both tissue cultures as well as group members. My summer-students Linda Carlsson and Johanna Johansson for great help in the lab.

Kristin-Sophie Mellsjö, Urban Pettersson and Ewa Winkler for taking such good care of my greenhouse plants. Special thanks to Urban for occasional agricultural chats. Ingrid Schenning and Yvonne Tillman in sequencing unit. Lars-Olof Hansson in the library for finding me the most impossible and long forgotten references. Lena Johansson and Björn Nicander for nursing my computer (of the better brand). Birgitta Eriksson for always fixing all kinds of practicalities so smoothly.

The mitomics group: Carina Knorpp, Jenni Hammargren, Monika Johansson and Per Bergman. For inspiring Journal Clubs, scientific advice and lots of fun on mitochondrial conference travels.

Past and present people at the department for all kinds of scientific and social events. Especially the members of the plant breeding group: Anki, Anna, Berit (thanks for lending me the potter’s wheel!), Christina (for introducing me to plant breeding), Delal, Guillermo, Gunilla, Harald, Janne (leading molecular biology expert), Jens, Johan, Maj-Britt, Maria, Marie (life IS circling around the short-film festival), Patrick, Oksana, Svante, Sofia, Ulf and Urban. All members in the in the floor-hockey team and the pub-crew. My former room-mates Karin and
Lisa. And all others making Plant Biology (and Forest Genetics) a really enjoyable place.

Bo Gertsson, my industrial mentor, for showing me Brassica breeding in practise and for providing the cover illustration (a Chinese hybrid seed production field of Qinyou#2, the first large-scale commercial rapeseed hybrid variety).

Dominique Gagliardi for my very productive time in Strasbourg, Howard Bonett for improving both the science and English of my manuscripts and Axel Brennicke for your valuable comments on presentation of data.

Additionally, warm thanks to Runar and Marianne Andersson at Nenninge gård where my true interest in agriculture started thirteen years ago, and to Lars Andersson, Nils-Ove Bertholdsson and Bengt Lundegårdh who introduced me into the world of plant research.

This work was funded by the Swedish University for Agricultural Sciences, the Swedish Research Council (VR), the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (FORMAS), the Swedish Foundation for International Cooperation in Research and Higher Education (STINT), SvalövWeibull AB, Carl Tryggers Foundation, the Nilsson-Ehle Foundation and Martha and Fredrik Nilssons Rememberance Foundation.

Finally, I would like to thank my friends and family for support and all the fun times we have together: Agneta for your enjoyable gardening mails whenever work is to boring, Peter for our regular beer drinking and vegetable chatting events, Per Rune for your friendship, I hope we can continue our expedition tradition to more places like Gästrikland. Tack Jenny! Johan, we still have lots of peaks to climb. Kusse Karin, the second half of Kusindesign, for fruitful Raku kiln construction and firings. My parents, Ti and Olle (Pelle), who encouraged my interest in plants from very early years. And last but not least, super-extra-special (massamycket) thanks to Linda (np!).