2022

Targeting Transgenic Proteins to the Mitochondria in Plants

Susan Qudus
University of Nebraska - Lincoln

Jeffrey P. Mower
University of Nebraska - Lincoln

TARGETING TRANSGENIC PROTEINS TO THE MITOCHONDRIA IN PLANTS

An Undergraduate Honors Thesis
Submitted in Partial fulfillment of
University Honors Program Requirements
University of Nebraska-Lincoln

by
Susan Qudus, BS
Biological Systems Engineering
College of Engineering

April 23, 2022

Faculty Mentor:
Jeffrey Mower, PhD, Center for Plant Innovation

Abstract

Cytoplasmic male sterility (CMS) in plants can convert an individual into female through sterilization of the male reproductive parts, creating an inability to produce functional pollen. In nature, this is advantageous for genetic diversity and producing progeny with unique genes that may be beneficial in evolution. For agriculture, this method is preferential for selection of desirable traits in crops when cross breeding. CMS is typically caused by genes in the mitochondrial DNA (mtDNA), such as novel open reading frames (ORFs) that form after imperfect repair of double-stranded breaks in the mtDNA. Specific CMS genes vary among various plant species. It is currently unclear if the CMS gene of one species is effective in inducing sterility to another. These genes are easily predictable from the gene sequence but are rarely experimentally tested to determine their effect to cause male sterility. In this paper, I worked to characterize four different CMS genes identified from *Brassica*. Using a mitochondrial targeting system, I assessed the ability to transgenically target one of these genes to the mitochondria. For future considerations if these genes can be targeted to the mitochondria, the analysis into the ability to cause CMS can be studied by transgenic expression in *Arabidopsis*.

Key Words: cytoplasmic male sterility, chimeric genes, mitochondrial targeting, *orf288, NtATPb*

Targeting Transgenic Proteins to the Mitochondria in Plants

Introduction

Most plants are hermaphrodites, with male and female parts present on each flower, or more rarely, separate male and female flowers. In nature and in agriculture, a process known as cytoplasmic male sterility (CMS) can convert some plants to female by the partial or complete sterility of male floral parts, resulting in an inability to produce viable pollen (Hanson and Bentolila 2004). Wild plant populations experiencing CMS can be naturally advantageous as this encourages cross pollination, which promotes increased genetic variation in progeny. CMS is widely utilized in agriculture to force cross-pollination of plant cultivars, which can produce hybrid progeny with enhanced features, such as increased height, growth rate, and yield; The enhanced characters of hybrids relative to the parental lines is a phenomenon known as heterosis, or hybrid vigor (Hochholdinger 2018). Controlling pollen production in GMO crops can also provide benefits in deliberately deterring the cross-pollination with non-GMO plants, thereby preventing the unwanted spread of transgenic material into wild plants or non-GMO crops. The reduction in pollen production could also offer advantages in reducing seasonal allergies caused by excessive pollen distribution.

The process of CMS is usually caused by novel factors within plant mitochondria. Mitochondria have their own DNA (mtDNA), containing genes involved in crucial metabolic processes, such as the tricarboxylic acid cycle and oxidative phosphorylation pathway (Millar *et al.* 2011). Some mtDNA carries novel open reading frames (ORFs) that can cause CMS. These CMS ORFs are typically created after mtDNA experiences double-stranded breaks, which can be imperfectly repaired to create novel combinations of genetic material. The CMS ORF can disrupt mitochondrial function during floral development or pollen production, resulting in male sterility (Hanson and Bentolila 2004). However, the sterility effect can be reversed, particularly when the nuclear genome carries a fertility restorer (Rf) gene. These Rf genes are imported into mitochondria and disrupt expression of CMS ORF, eliminating the activity of the CMS ORF and restoring the plant to male fertility (Hanson and Bentolila 2004).

These CMS genes are typically chimeric genes (Hanson and Bentolila 2004). Chimerism refers to when sections of two or more coding regions combine to create a new open reading frame (ORF). This occurs through various methods, such as faults in DNA replication or repair where two or more genes are mistakenly integrated. Additionally, a transcript of a gene in one location of the genome can be copied and inserted in a novel location through retrotransposition to form chimeric genes. This process may change the recognition of exons to create a new combination of nucleotides. The effect or novel function of most chimeric genes are largely unknown (Bonde and Bulow 2013).

To take advantage of heterosis, natural CMS systems are commonly used to produce hybrid breeding systems in many crops, such as rice, beans, and canola. However, not all crops take advantage of CMS and heterosis. In some crops, identified CMS ORFs may not produce 100% sterility, while Rfs may not be 100% efficient at fertility restoration. Furthermore, even if the CMS ORFs and Rfs are highly efficient, it is often very time consuming to move these genes into a different cultivar, requiring multiple rounds of back-crossing. Importantly, CMS factors in mtDNA of one species can be used to induce sterility when transgenically inserted and expressed in other plants (Landgren *et al.* 1996; Yang *et al.* 2010). The long-term goals of this project are to test the ability and efficiency of CMS ORFs identified in various crops to generate CMS in the model plant *Arabidopsis thaliana*, which is otherwise a predominantly self-pollinated plant.

As CMS proteins function within the mitochondria, but transgenes introduced by transformation are located in the nuclear genome, to transgenically induce CMS in another plant species a mitochondrial targeting sequence (MTS) is necessary. An MTS is a short peptide that contains positively charged residues alternating with hydrophobic residues to create an amphipathic helix at the N-terminal of a protein sequence (Chatre *et al.* 2009). The MTS will ensure the transgenically expressed CMS protein is transported to the mitochondria to function. Natural CMS genes do not include this precursor sequence to target the mitochondria as they are encoded in the mtDNA. Two MTSs are commonly in use to target transgenic proteins into mitochondria, and they were evaluated for the purposes of this experiment: *ATP2 (NtATPb*) from *Nicotiana* and *AOX1a* from *Arabidopsis* (Yang *et al.* 2010, Carrie *et al.* 2008).

In this study, I will characterize four CMS genes from Brassica species, use bioinformatics to evaluate MTS from two nuclear genes to predict their ability to target a protein to the mitochondria, and experimentally assess the ability of one MTS to target a transgenic CMS protein to the mitochondria

Methods

Characterizing the CMS genes

The program Blastx is utilized to identify genes surrounding the CMS genes in the genome and to characterize the chimeric elements of each CMS gene.

Prediction of Mitochondrial Targeting

As we are focusing on targeting specific proteins to the mitochondria, we used the program TargetP-2.0 to predict the location that two different MTS sequences will target. The protein sequence was input into the neural network predictor program to investigate whether the sequence will be targeting the mitochondria or other locations in the cell, such as chloroplast or secreted from the cell.

Production of CMS expression construct

Genes were synthesized from Genscript with modifications made to the gene sequence for optimal translation in the model *Arabidopsis*. Restriction enzyme cut sites were added around the gene to aid in recombinant cloning into the expression plasmid, *pMDC83*. This plasmid contains the *35S* promoter used for expression, in-frame GFP marker for transgenic visualization, Kanamycin resistance for bacterial selection, and Hygromycin for plant selection. Internal restriction enzyme cut sites were removed from the gene sequences for isolation of the selected enzyme cut sites. Using recombinant cloning, the MTS+CMS construct was transformed into *pMDC83.* The recombinant construct was confirmed using restriction digestion patterns.

Cloning into E. coli

The stock plasmid DNA was transformed into *E. coli* through heat shock transformation using DH5α competent cells. The experiment followed ThermoFisher "Transforming Chemically Competent Cells" protocol. Competent cells were thawed on ice, mixed with the ligation reaction, and incubated on ice for 30 minutes. The mixture was then heat shocked at 42℃ for 30 seconds and mixed with SOC

medium to shake for one hour at 37°C. The transformation was spread on a selection plate and bacterial colonies were selected and grown the following day.

Extracting Purified Plasmid

To extract and obtain high concentrations of plasmid, the QIAGEN Plasmid Midi Kit was utilized. The bacterial cultures were grown in 50 mL of media overnight. Buffer P1 was added to resuspend the bacterial pellet after centrifugation. Buffer P2 was used to lyse the mixture. The addition of Buffer P3 precipitates cellular material. The DNA was bound to a column and then washed to purify the plasmid. The plasmid was eluted and maintained in 30 mL of water.

Agrobacterium Electroporation

The PMDC83 expression plasmid was transformed into agrobacterium C58C1 cells by electroporation following a lab protocol available in the Mower lab. Bio-Rad electroporation cuvettes were cooled to 4°C. The *Agrobacterium* C58C1 stock cells were thawed on ice and ~20-30 ng of purified plasmid was added. The mixture was transferred to the cuvette, and a voltage of 2,000 V was administered. The cells were transferred to 1 mL of SOC medium and grown at 28°C for one hour to increase the transformation efficiency. The mixture was spun down in a centrifuge, ~750 µL of supernatant was removed, and the remaining transformation was spread on pre-warmed Kanamycin selection plates.

Transient Transformation into Tobacco

Bacteria colonies were collected from the electroporation plates and grown overnight in 5 mL of LB media with antibiotic selection then moved to a volume of 25 mL of LB with antibiotics. The culture was centrifuged to create a pellet and resuspended in MMA media with OD_{600}= 1.0. MMA media was created by mixing 10 mM MES, 10 mM $MgCl_2$, and 200 µM acetosyringone. The cultures were then incubated for 2-4 hours at 28°C. The bacteria were transiently infected into tobacco plants with a blunt tip syringe to the underside of fresh and healthy leaves. The solution was injected and spread

through labeled areas on the tobacco plant. The plant was left for 48 hours to successfully infect the plants.

Microscopy to Assess Subcellular Localization

After the transient transformation into tobacco, microscopy was used to visualize the effectiveness of the plasmid to infect and target the mitochondria. The infected leaves were cut into small disks and analyzed under the microscope. With GFP as the marker, green fluorescence was observed to visually assess the effectiveness of mitochondrial targeting.

Results & Discussion

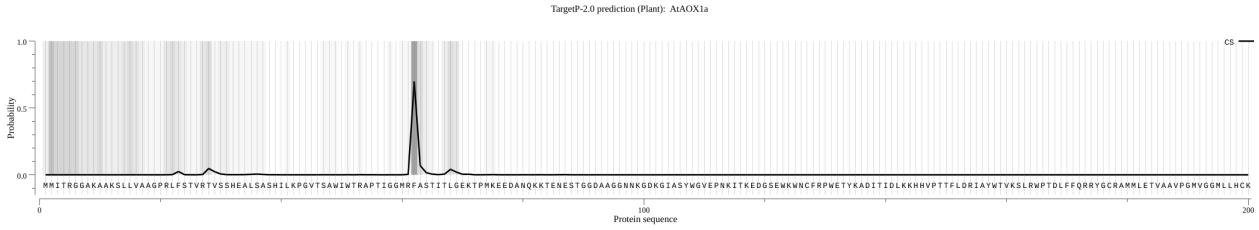
Prediction of Targeting

Results from the intracellular prediction program TargetP-2.0 are displayed in Table 1 and Figure 1. The MTS gene *NtATPb* was predicted to have a 99.84% probability to target the mitochondria in plants. MTS gene *AtAOX1a* had an 85.71% probability to localize to the mitochondria in plants. As *NtATPb* shows a stronger probability towards targeting the mitochondria, this MTS was selected to synthesize with the CMS genes and used for cloning into the destination plasmid- *pMDC83.*

Table 1. MTS Targeting Predictions

TargetP-2.0 Targeting Predictions		
Location	***NtATPb***	***AtAOX1a***
Mitochondria	0.9984	0.8571
Chloroplast	0.007	0.1183
Other	0.009	0.0344

A.



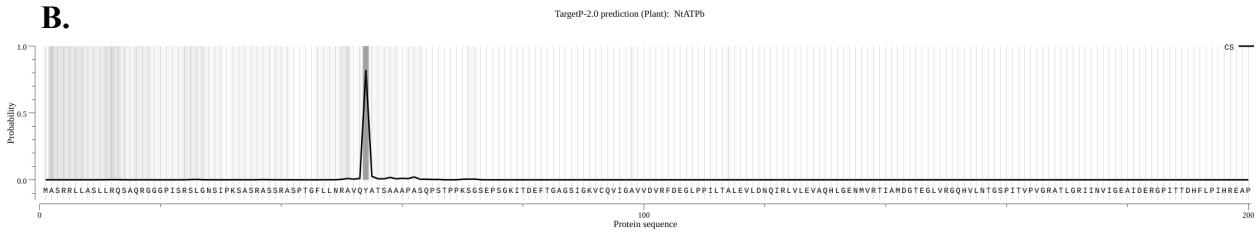


Figure 1. Targeting Prediction for **A.** NtATPb and **B.** AtAOX1a

Synthesis of CMS Construct

Four different CMS genes were selected from literature that suggests an associated relationship for these genes to cause male sterility in plants. The gene *BjORF288* has been experimentally correlated in multiple papers to cause CMS (Heng *et al.* 2012, Heng *et al.* 2018). It was shown to cause defective anther and pollen development when infected into *Arabidopsis thaliana* (Jing *et al.* 2012). We utilized this gene as our positive control for association to cause male sterility in plants. *BtORF263* is also shown in previous literature to be associated with cytoplasmic male sterility in *Brassica tournefortii* (Landgren 1996). The remaining genes were predicted to cause male sterility based on genomic analysis (Grewe *et al*. 2014).

Prior to synthesis, edits were made to each gene sequence for the benefit of experimental usage. For *BtORF263,* a partial XbaI enzyme cut site at the end of the sequence was removed. In *BjORF288*, the EcoRI enzyme cut site near the end and XbaI cut site at the end of the sequence were removed. For *BoORF287,* the XbaI enzyme cut site at the end was removed. A silent edit was made to a single amino acid in the sequence. In the sequence for *BoORF266*, a BamHI cut site was removed from the middle. Three edited enzyme sites were also added to the sequence. Stop codons at the end of each gene sequence were removed to promote transcription and translation of the gene with the GFP marker at the end of the construct.

Table 2. CMS genes selected, origins, and GenBank accession numbers

Gene	Origin	GenBank Expression Number
BtORF263	*Brassica tournefortii*	X83692
BjORF288	*Brassica juncea*	KF736092
BoORF287	*Brassica oleracea*	KJ820683
BoORF266	*Brassica oleracea*	KJ820683

After selecting distinct CMS genes to analyze and an effective MTS gene, one CMS gene was chosen to clone into the designation plasmid. the plasmid was constructed and collected by cloning into *E. coli* and using the QIAGEN Midi Prep Kit. The plasmid created (as seen in Figure 2) contains the MTS component *NtATPb* to target the mitochondria, *BjORF288* to induce CMS, and GFP to visually mark the targeting in the cells. Before moving to electroporation, the plasmid was confirmed through restriction enzyme patterns.

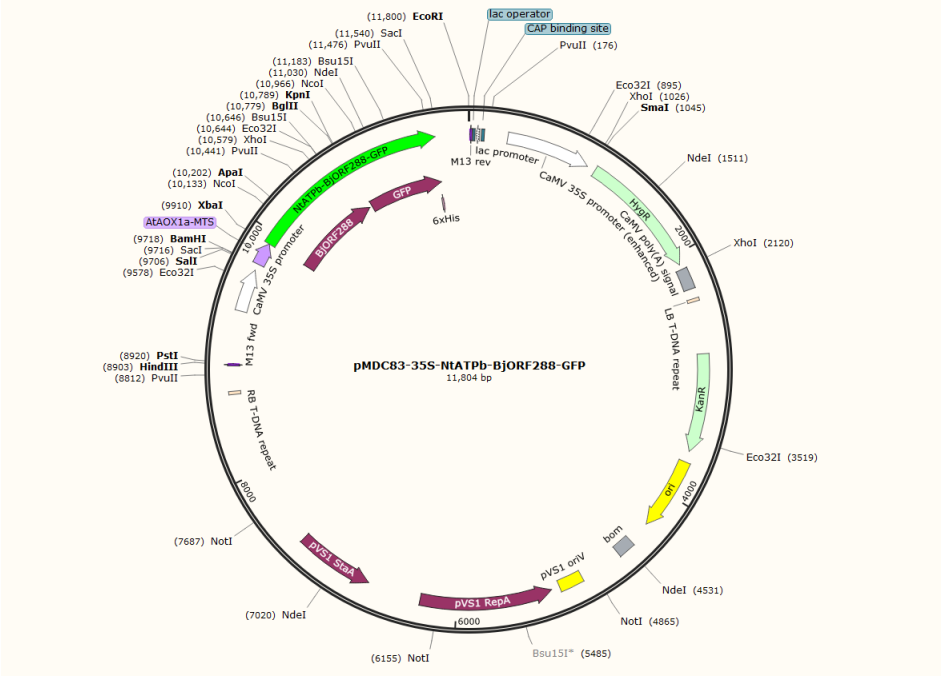


Figure 2. Construct map of synthesized plasmid *pMDC83-35S-NtATPb-BjORF288-GFP*

Gene Characterization & Analysis

As CMS genes are typically chimeric in nature and expressed with other mitochondrial genes, we assessed the chimeric content of each chosen CMS gene and their closest gene neighbors in the mitochondria. The structures and neighboring genes of the four selected different CMS genes were analyzed and contracted in a visual model in Figure 3.

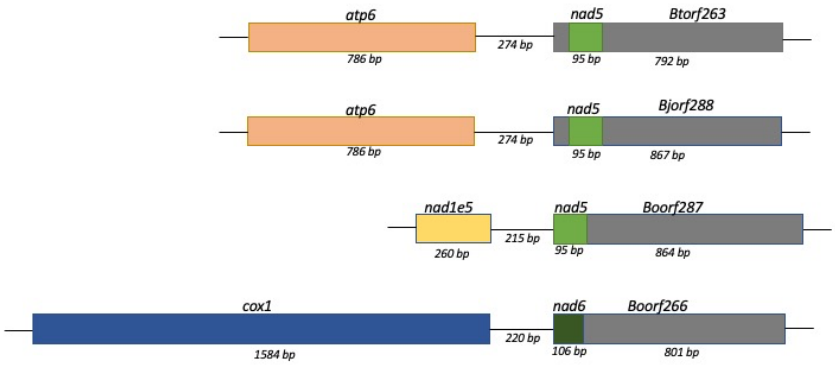


Figure 3. Gene mapping of upstream genes and the chimeric nature of the CMS genes of interest

As seen from first glance at the *BtORF263* and *BjORF288* structures, these genes are nearly identical in composition. They have an *atp6* gene upstream and are chimeric with a *nad5* gene fragment that begins towards the beginning of the CMS sequence. The main difference lies in the size of the CMS gene itself, with a difference of 75 base pairs in the larger *BjORF288* gene.

The *BoORF287* gene is similar in structure to *BtORF263* and *BjORF288*, though it contains a different upstream gene, *nad1e5*. This gene is distinctly shorter than *atp6* and lies closer to the CMS gene with a 59 base pair difference in length. The internal gene fragment *nad5* is the same though it is present at the beginning of the CMS gene rather than being near the start.

The *BoORF266* gene is the most distinctly different than the other three CMS genes. The upstream gene, *cox1*, is largely different in length compared to the previous two with a 798 base pair increase from *atp6* and 1,324 base pair increase from *nad1e5*. The internal gene fragment, *nad6*, differs

from the previous CMS and lies in the same region as *BoORF287* at the start of the CMS gene. The length of the CMS gene is similar to *BtORF263* and shorter than *BjORF288* and *BoORF287*.

Based on this analysis, we expect *BtORF263* and *BjORF288* to behave the most similarly when expressed in the host plant. As *BjORF288* has been shown to cause male sterility in literature, we expect *BtORF263* to also cause CMS based on the genetic similarities. The behavior of *BoORF287* is predicted to be alike the previous two. Though with genomic differences, we may expect a different expression of male sterility, such as differing efficacy to cause degenerative pollen production or anther development. *BoORF266* is expected to behave the most dissimilarly than the remaining three genes and may induce CMS by a different mechanism.

Subcellular Localization

After transforming the plasmid into *Agrobacterium* C58C1 cells and infecting into tobacco by transient transformation, the infected leaf disks were observed under the microscope. Images of the microscopy can be observed in Figure 4.

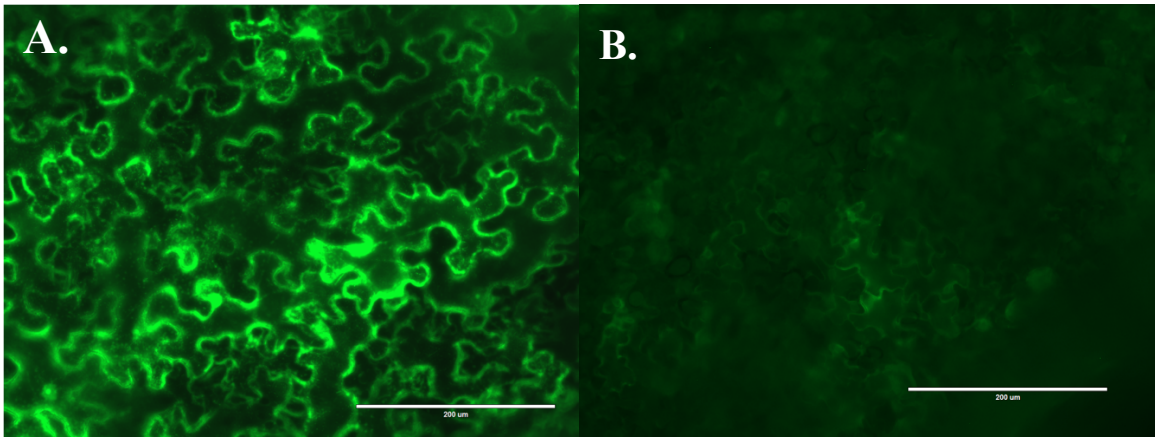


Figure 4. A. Positive Control- *35S-NtATPb-GFP* infected into tobacco**. B.** Experimental construct, *pMDC83-35S-NtATPb-BjORF288-GFP* infected into tobacco, failed to target mitochondria

Figure 4A shows a mitochondrial signal for a positive control plasmid *35S-NtATPb-GFP*. This plasmid lacks the CMS gene to observe and compare the behavior of the same construct with the transgene present. The GFP florescence is strong and mitochondrial targeting can be seen based on the spattering of dots inside of the cell walls. This indicates the positive control plasmid successfully

localized to the mitochondria and acts as a positive control to compare results with the constructed plasmid in this experiment.

Figure 4B shows the microscopy results for the experimental construct *pMDC83-35S-NtATPb-BjORF288-GFP* infected into tobacco. From this image, the GFP florescence can be seen majorly along the cell walls of the tobacco, though the signal is relatively week compared to Figure 4A. Analyzing inside of the cells in Figure 4B, there is no apparent GFP fluoresce of mitochondria observed. This indicates the construct did not successfully localize to the mitochondria.

The experimental vector failing to target the mitochondria indicates the plasmid construct was not working properly. The error may lie in issues with the destination plasmid, as there were issues in verifying certain enzyme cut sites to the known sequence and enzyme map in the lab. This means during the transformation experiments, that the plasmid construct could rearrange in undesired ways. The issue could also be the chosen MTS gene, *NtATPb*. This conclusion is unlikely as seen in Figure 4A that the MTS component worked to localize the mitochondria and produce a strong signal. This MTS gene is commonly used in literature and the field as it strongly targets the mitochondria seen also in the prediction from the results of the program TargetP2.0 in Table 1. The GFP indicator utilized may also have unfavorable interactions with the plasmid construct, resulting in weak signals when analyzed. The CMS gene could be the source of error, as it is the factor changed in the positive control and experiential results. The transgene may not have favorable interactions with the MTS sequence to fail to target to the mitochondria. Ultimately, it could be some interaction of these various components that fails to target the MTS+CMS+GFP protein to the mitochondria. To recognize the true error, future efforts are needed to experimentally determine the source.

Conclusion

In conclusion, a strong mitochondrial targeting sequence was identified, and four CMS genes of interest were recognized from previous literature and genomic analysis. The plasmid construct *pMDC83-35S-NtATPb-BjORF288-GFP* was observably unable to successfully localize to the mitochondria. This step is necessary to express the male sterility transgene in the mitochondria and observe if the selected

four CMS genes cause male sterility. Future consideration and experimentation are necessary to resolve the cause of this error and to improve the results to further the project. The ultimate goal of this project is to use transgenic methods to determine if these genes can cause cytoplasmic male sterility in a distinctly different species than sourced, like *Arabidopsis.* Each CMS and MTS gene combination can be created and stably infected into a large quantity of plant samples to observe the phenotypic changes and possible reduction viability or capacity of pollen production.

References

Bonde, J.S., Bulow, L. (2013) Chimeric Genes, Proteins, Brenner's Encyclopedia of Genetics, Second Edition, Elsevier.

Carrie, C., Murcha, M.W., Kuehn, K., Duncan, O., Barthet, M., Smith, P.M., Eubel, H., Meyer, E., Day, D.A., Millar, A.H. and Whelan, J. (2008) Type II NAD(P)H Dehydrogenases Are Targeted to Mitochondria and Chloroplasts or Peroxisomes in *Arabidopsis thaliana*. *Federation of European Biochemical Societies Letters*, 582(20), 3073-3079.

Chatre, L., Matheson, L.A., Jack, A.S., Hanton, S.L. and Brandizzi, F. (2009) Efficient Mitochondrial Targeting Relies on Co-Operation of Multiple Protein Signals in Plants. *Journal of Experimental Botany*, 60(3), 741-749.

Colas des Francs-Small, C. and Small, I. (2014) Surrogate mutants for studying mitochondrially encoded functions. *Biochimie,* 100, 234-242.

Grewe, F., Edger, P.P., Keren, I., Sultan, L., Pires, J.C., Ostersetzer-Biran, O. and Mower, J.P. (2014) Comparative analysis of 11 Brassicales mitochondrial genomes and the mitochondrial transcriptome of *Brassica oleracea. Mitochondrion*, 19, 135-143.

Hanson, M.R. and Bentolila, S. (2004) Interactions of mitochondrial and nuclear genes that affect male gametophyte development. *The Plant Cell,* 16, 154-169.

Heng, S., Wei, C., Jing, B., Wan, Z., Wen, J., Yi, B., Ma, C., Tu, J., Fu, T. and Shen, J. (2014) Comparative analysis of mitochondrial genomes between the *hau* cytoplasmic male sterility (CMS) line and its iso-nuclear maintainer line in *Brassica juncea* to reveal the origin of the CMS-associated gene *orf288. BMC Genomics,* 15(1), 322.

Heng, S., Gao, J., Wei, C., Chen, F., Li, X., Wen, J., Yi, B., Ma, C., Tu, J., Fu, T. and Shen, J. (2018) Transcript levels of *orf288* are associated with the *hau* cytoplasmic male sterility system and altered nuclear gene expression in *Brassica juncea*. *Journal of Experimental Botany*, 69, 455-466.

Hochholdinger, F., Baldauf J.A. (2018) Heterosis in plants. *Current Biology*, 28(18), R1089-R1092.

Jing, B., Heng, S., Tong, D., Wan, Z., Fu, T., Tu, J., Ma, C., Yi, B., Wen, J. and Shen, J. (2012) A male sterility-associated cytotoxic protein ORF288 in *Brassica juncea* causes aborted pollen development. *Journal of Experimental Botany*, 63, 1285-1295.

Kazama, T., Okuno, M., Watari, Y., Yanase, S., Koizuka, C., Tsuruta, Y., Sugaya, H., Toyoda, A., Itoh, T., Tsutsumi, N., Toriyama, K., Koizuka, N. and Arimura, S.-I. (2019) Curing cytoplasmic male sterility via TALEN-mediated mitochondrial genome editing. *Nature Plants*, 5, 722-730.

Landgren, M., Zetterstrand, M., Sundeberg, E. and Glimelius, K. (1996) Alloplasmic male-sterile *Brassica* lines containing *B. tournefortii* mitochondria express an ORF 3' of the *atp6* gene and a 32 kDa protein. *Plant Molecular Biology*, 32, 879-90.

One Shot® TOP10 Competent Cells. ThermoFisher Scientific., Waltham, MA., USA, 2013.

QIAGEN Plasmid Midi Kit. QIAGEN., Hilden, Germany. 2016.

Yang, J., Liu, X., Yang, X. and Zhang, M. (2010) Mitochondrially-Targeted Expression of a Cytoplasmic Male Sterility-Associated *orf220* Gene Causes Male Sterility in *Brassica juncea*. *BMC Plant Biology*, 10, 231.