

Article

The Mitochondrial Genome Assembly of Fennel (*Foeniculum vulgare*) Reveals Two Different *atp6* Gene Sequences in Cytoplasmic Male Sterile Accessions

Fabio Palumbo ¹,*[®], Nicola Vitulo ²[®], Alessandro Vannozzi ¹[®], Gabriele Magon ¹ and Gianni Barcaccia ¹[®]

- ¹ Department of Agronomy Food Natural Resources Animals and Environment (DAFNAE), University of Padova, Campus of Agripolis, Viale dell'Università 16, 35020 Legnaro, PD, Italy;
- alessandro.vannozzi@unipd.it (A.V.); gabriele.magon@phd.unipd.it (G.M.); gianni.barcaccia@unipd.it (G.B.)
 ² Department of Biotechnology, University of Verona, Strada Le Grazie 15, 37134 Verona, Italy; nicola.vitulo@univr.it
- * Correspondence: fabio.palumbo@unipd.it

Received: 6 June 2020; Accepted: 29 June 2020; Published: 30 June 2020



Abstract: Cytoplasmic male sterility (CMS) has always aroused interest among researchers and breeders, being a valuable resource widely exploited not only to breed F1 hybrid varieties but also to investigate genes that control stamen and pollen development. With the aim of identifying candidate genes for CMS in fennel, we adopted an effective strategy relying on the comparison between mitochondrial genomes (mtDNA) of both fertile and sterile genotypes. mtDNA raw reads derived from a CMS genotype were assembled in a single molecule (296,483 bp), while a draft mtDNA assembly (166,124 nucleotides, 94 contigs) was performed using male fertile sample (MF) sequences. From their annotation and alignment, two *atp6*-like sequences were identified. *atp6*⁻, the putative mutant copy with a 300 bp truncation at the 5'-end, was found only in the mtDNA of CMS samples, while the wild type copy $(atp6^+)$ was detected only in the MF mtDNA. Further analyses (i.e., reads mapping and Sanger sequencing), revealed an $atp6^+$ copy also in CMS samples, probably in the nuclear DNA. However, qPCRs performed on different tissues proved that, despite its availability, *atp6*⁺ is expressed only in MF samples, while *apt6*⁻ mRNA was always detected in CMS individuals. In the light of these findings, the energy deficiency model could explain the pollen deficiency observed in male sterile flower. *atp6⁻* could represent a gene whose mRNA is translated into a not-fully functional protein leading to suboptimal ATP production that guarantees essential cellular processes but not a high energy demand process such as pollen development. Our study provides novel insights into the fennel mtDNA genome and its *atp6* genes, and paves the way for further studies aimed at understanding their functional roles in the determination of male sterility.

Keywords: male sterility; ATP synthase; mtDNA; organelle assembly; energy deficiency model; F1 hybrid breeding; pollen; marker-assisted selection

1. Introduction

Plant male sterility (MS), is generally defined as the inability of plants to produce functional pollen, dehiscent anthers, or viable male gametes [1]. According to their mode of inheritance, two types of MS are recognized: Cytoplasmic male sterility (CMS), which is caused by mitochondrial genes coupled with fertility restoring nuclear genes, and nuclear or genic male sterility (NMS or GMS), determined exclusively by nuclear genes [2].

Both types of MS are currently exploited for breeding purposes and hybrid constitution and the worldwide crop production has tremendously benefited from their utilization. It is estimated that



hybrid vigor can offer, on average, from 20% to over 50% increases [3] and this would explain why hybrid seeds market, valued at \$52 billion in 2016, is expected to reach at \$99 billion by 2023, with a Compound Annual Growth Rate (CAGR) of 9.7% [4].

CMS is more widely used than GMS because this maternally inherited trait can be applied, preserved and restored more efficiently [5,6]. In fact, a CMS-based hybrid seed technology only requires a CMS line, a sterility maintainer line, and a fertility restorer line [7,8]. Moreover, in addition to its wide application in hybrid production, CMS provides important resources to investigate stamen and pollen development and to study the crosstalk between nuclear and mitochondrial genomes [9]. Most of the CMS genes reviewed by Chen and Liu [1] belonged to the mitochondrial electron transfer chain (mtETC) or to the ATP synthase complex and, among them, it is worth mentioning cox1 [9,10], *atp8* [11], and *atp6* [5,12,13]. In other cases, species-specific mitochondrial open reading frames (ORFs) encoding for transmembrane proteins were found to be responsible for CMS [14,15]. According to the National Center for Biotechnology Information (NCBI) [16] CMS-related genes have been identified and deposited for at least 76 plant species, including cereals such as bread wheat (*Triticum aestivum*) [17], rice (Oryza sativa) [18], and maize (Zea mays) [19], and crops like stem mustard (Brassica juncea) [20,21], onion (Allium cepa) [9], pepper (Capsicum annuum) [5], radish (Raphanus sativus) [22], sunflower (Helianthus annuus) [23], and carrot (Daucus carota) [11]. In fennel (Foeniculum vulgare), the presence of a naturally occurring NMS system has never been reported while several spontaneous CMS mutants have been discovered and exploited by breeders since the late 1960s [24,25]. These mutants are widely used for the development of F1 hybrids [2] that, today, represent the 95% of the cultivated fennel varieties [26]. Strikingly, despite its economic relevance (more than 1 million tons produced in 2018, with a gross production value of \$6.5 billion [27]), the genetic mechanism underlying the cytoplasmic male sterility in fennel is far from being understood and no candidate gene has been identified yet. What is more, the only biological data currently available for this species in public databases (e.g., GenBank) are the plastome [28], an early genome draft [2] and a leaf transcriptome [29].

It is now well understood that one of the most effective strategies to identify CMS candidate genes relies on the comparison between mitochondrial genomes of both fertile and sterile genotypes, to search for differences in terms of sequences or structure organization [30,31]. In the last few years the identification of such genes has been greatly facilitated by the advent of Next Generation Sequencing (NGS): To date 311 plant mitochondrial genomes are available in the NCBI organelle genome database [32] and this number is constantly increasing thanks to the ongoing evolution of sequencing technology and the decrease in sequencing costs [33].

Following this approach, in the present study we performed a fennel genome mitochondrial assembly based on data collected from public repositories and we then looked for nucleotide variants comparing sequences from both fertile and sterile lines.

2. Results

2.1. Mitochondrial Genome Sequence and Annotation

486,047,668 reads Illumina HiSeq 2500 (150 bp PE, Illumina, Inc., San Diego, CA, USA) derived from the Whole Genome Sequencing (WGS) of a cytoplasmic male sterile (CMS) accession (SRX7730623) of *Foeniculum vulgare*, were assembled in a single mtDNA molecule of 296,483 bp with a GC contents of 45.9% (Supplementary File 1). Mapping back the reads to validate the mtDNA assembly, the proportion of properly aligned reads was 0.3% with an average coverage depth of 758×. From the annotation step performed through GeSeq [34] 24 mtDNA genes, 26 ORFs, 25 tRNAs, and 9 rRNAs were identified (Figure 1, Supplementary File 2).





Figure 1. Annotation of *Foeniculum vulgare* mitochondrial genome assembled from a cytoplasmic male sterile (CMS) accession. The annotation was created with GeSeq [34] and visualized with OGDRAW. Genes marked with an asterisk (*) contain introns.

In parallel, 17,377,226 reads Ion GeneStudio S5 System (400 bp SE, Thermo Scientific, Pittsburgh, PA, USA) were retrieved from the WGS of a male fertile (MF) fennel accession (SRX2770225) and assembled in 94 contigs (Supplementary File 3), ranging from 168 bp to 10,174 bp (total length 166,124 bp, N50 = 4702, L50 = 12) and a GC contents of 46.3%. The percentage of raw reads that were correctly mapped back to the MF mtDNA contigs was 0.4%, with an average coverage depth of 151×. GeSeq [34] allowed for the identification and annotation of 24 mtDNA genes, 24 ORFs, 19 tRNAs, and 9 rRNAs were identified (Supplementary File 4).

From a comparison among the CMS and MF mitochondrial genomes, a one-to-one correspondence was found for three ribosomal protein large subunits (*rpl2*, *rpl5*, *rpl10*), six ribosomal protein small subunits (*rps1*, *rps3*, *rps4*, *rps7*, *rps12*, *rps13*); nine nadh dehydrogenase subunits (*nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *nad7*, and *nad9*, complex I), three cytochrome oxidase subunits (*cox1*, *cox2*, and *cox3*, complex IV), cytochrome b (*cob*, complex III), five atp synthase subunits (*atp1*, *atp4*, *atp6*, *atp8* and *atp9*, complex V), four cytochrome C synthesis related proteins (*ccmB*, *ccmC*, *ccmFc* and *ccmFn*), *matR*, and *mttB*.

2.2. SNP and InDel Detection

In order to identify sequence variations between the MF and CMS accessions, SNPs and InDels were searched aligning the mtDNA MF contigs against the CMS mtDNA. As shown in Table 1, 63 variants, namely 36 SNPs and 27 InDels, were detected.

Table 1. Variant analysis results. Variant location in CMS mtDNA, variant type (Ins = Insertion, Del = Deletion, SNV = Single Nucleotide variant, MNV = Multiple Nucleotide Variant), reference allele (CMS), alternative allele (MF) and annotation (ID) are reported (bold indicates a variant located within the exon).

CMS mtDNA Position	Туре	Ref	Alt	ID	ID CMS mtDNA Position		Ref	Alt	ID
4407^4408	Ins	-	Т		210132	SNV	А	G	atp6
4413^4414	Ins	-	Т		210134210137	MNV	ACCC	GGAA	atp6
9932	Del	А	-		210140210141	MNV	CC	TT	atp6
9937	Del	А	-		210143	SNV	А	G	atp6
17054	SNV	Т	А		210145	SNV	С	Т	atp6
17779	SNV	А	С		210148210150	MNV	CGA	TAG	atp6
19036	Del	Т	-		210152210153	MNV	TA	GG	atp6
2377923780	Del	GG	-	nad4 (intron)	210155	SNV	Т	G	atp6
26757^26758	Ins	-	G	nad4 (intron)	210157210158	MNV	CT	AG	atp6
29433	Del	G	-	nad4 (intron)	210160210163	MNV	ATAT	CCCC	atp6
38972^38973	Ins	-	А		210168210171	MNV	GCTT	ATCC	atp6
85033	Del	А	-		210176	SNV	Т	С	atp6
97184	Del	А	-		210179	SNV	Т	А	atp6
97194	Del	С	-		210193210194	MNV	AA	CC	atp6
103048	Del	Т	-		210218	SNV	G	Т	atp6
106177^106178	Ins	-	Т		210225	SNV	G	С	atp6
109983	Del	G	-		210276	SNV	G	А	atp6
110419	SNV	Т	G		210291	SNV	Т	С	atp6
175249	SNV	Т	С	rps3	210293	SNV	G	А	atp6
175458	SNV	С	G	rps3	210296	SNV	Т	G	atp6
176686	SNV	G	А	rps3	210374	SNV	С	Т	atp6
181657	Del	Т	-	cox2 (intron)	212667	Del	Т	-	
210097210102	Del	CCGCTC	-	atp6	213187	SNV	А	G	
210103210105	MNV	CCA	TTT	atp6	213216	SNV	А	С	
210111	SNV	С	Т	atp6	238616^238617	Ins	-	С	
210113	SNV	G	А	atp6	238643^238644	Ins	-	G	
210117	SNV	А	G	atp6	240379^240380	Ins	-	TTTATTA	AT orf43
210117^210118	Ins	-	AG	atp6	247260	Del	С	-	
210122	SNV	А	G	atp6	248437	Del	А	-	
210124210126	MNV	TAA	GTG	atp6	293347	Del	Т	-	<i>nad2</i> (intron)
210129^210130	Ins	-	CTT	atp6	295317	Del	Т	-	nad2 (intron)
210130	SNV	А	Т	atp6					

Among the sequence variations, 22 occurred in extra genic regions, 6 within intron regions and, as a main finding, *atp6* and *rps3* were the only two CDS regions showing nucleotide variants between fertile and sterile samples (Table 1, in bold). As regards the *rps3* gene (1659 bp), all the 3 SNPs resulted non-synonymous (Figure S1), while for *atp6*, 31 variants were responsible for 28 amino acid mismatches (Figure S2). More interestingly, the first 300 bp of the MF *atp6* copy (from now on defined as *atp6*⁺, 1167 bp) were missing in the *atp6* copy (henceforth *atp6*⁻, 864 bp) found in male sterile mtDNA (Figure 2).

	1	10	20	30 4	0 5	0 6	0 70	8	o 90) 10	0 11	0 12	130
atp6+ atp6-	ATGCCTAT	ATTTCGCCG	C⊤TTGACGAG	AACAGCCTGAAT	ΑĠΤΑΑΤĊΑΑΑ	CATATACTACT	ΓΑΑΑΤĊTTTTA	CTAGTCAACO	TGAACCAACG	ACGACTATGA	СТАТТААТБА	TTATGCTTT	AACCCATCTGC
	1	40	150	160 17	0 180) 190	0 200	210	220	230	240	250	260
atp6+ atp6-	CGCTTCTG	GCGCTCAGG	CTCAGGCGTT	GAATGAACGTGC	AGCGGAACTTO	GGGGCTGGAAT	FACACGACGTA	TTAGGAGAGT	TAATGGTTAA	CAGAAGTCCC	GACGAGGTTT	TAAGAGCGGC	GGAGGCGGTGC
	27	70 2	80 2	90 300	310	320	330	340	350	360	370	380	390
atp6+ atp6-	ATGAGTCA	AGCAACGGG	ATCGCTTTTC	TGGAGTGCCTTT	TAA <mark>ATGA</mark>	TCCCA	TAAATCGAGT ACAGATCAT	GGAGAGTCAA GGAAATAAAA	CCTTTTGCGG		TATAGOGOG TACGAOTAGT		CCATCCCACC
	400	41	0 42	0 430	440	450	460	470	480	490	500	510	520
atp6+ atp6-	CCACTTGA	GCAATTTGC		ТТGATTCCTATG ТТGATTCCTATG	AATATAGGACA AAGATAGGAGA		ΤCΑΤΤCΑCAA		GTTTATGCTA GTTTATGCTA	CTCACTCTCA CTCACTCTCG	стттестсст стттестсст	АСТПСПАСТО АСТПТПССТТ	CATTTTGTTAC CATTTTGTTAC
	530	540	550	560	570	580	590	600	610	620	630	640	650
atp6+	TAAAAACO	GAGGAGGAA	ACTCAGTACC	AAATGCTTGGCA	АТССТТОСТАС	GAGCTTATTTA	ATGATTTTGTG	CCGAACCCG	TAAACGAACA	AATAGGTGGT	CTTTCCGGAA	ATGTTAAACA	AAAGTTTTCCC
atp6-	TAAAAACO	GAGGAGGAA	ACTCAGTACC	AAATGCTTGGCA	ΑΤϹϹΤΤGGTAG	GAGCTTATTT	ATGATTTCGTG	CCGAACCCGG	TAAACGAACA	AATAGGTGGT	CTTTCCGGAA	ATGTTAAACA	AAAGTTTTCCC
	660	670	680	690	700	710	720	730	740	750	760	770	780
atp6+	CTCGCATO	TCGGTCACT	ΤΤΤΑCΤΤΤΤΤ	CGTTCTTTCGTA	ATCCACAGGGT	ΓΑΤGΑΤΑCCΤΊ	TATAGCTTCAC.	AGTTACAAGT	CATTTTCTCA	TTACTTTGGG	CCTCTCATTT	TCAATTTTTA	TTGGCATTACT
атрь-	CICGCATO 700	PAD	910	2GII (2III (2GIA	P20			AGI LACAAGI	970	PRO 000		BOO	010
ato6+										CTTGAGCTAA		300	
atp6-	CTCGTGGG	ATTTCAAAG	AAATGGGCTT	CATTTTTTAAGC	TTTTCATTAC	CGCAGGAGT	CCACTGCCGT	TAGCACCTTT	TTTAGTACTC	CTTGAGCTAA	TCCCTCATTG	TTTTCGCGCA	TTAAGCTCAGG
	920	930	940	950	960	970	980	990	1,000	1,010	1,020	1,030	1,040
atp6+	AATACGTT	TATTTGCTA	ATATGATGGC	CGGTCATAGTTC	AGTAAAGATTI	TAAGTGGGT	IC GCTTGGACT.	ATGCTATGTA	TGAATGATCT	TTTATATTTC.	ATAGGAGATC	TGGGTCCTTT	ATTTATAGTTC
atp6-	AATACGTI	TATTTGCTA	ATATGATGGC	CGGTCATAGTTC	AGTAAAGATTI	TAAGTGGGT	TCGCTTGGACT	ATGCTATGTA	TGAATGATCT	TTTATATTTC.	ATAGGAGATC	TGGGTCCTTT	ATTTATAGTTC
	1,050	1,060	1,070	1,080	1,090	1,100	1,110 1	,120	1,130 1	1,140 1	,150 1	1,160	1,173
atp6+ atp6-	TTGCATTA TTGCATTA	ACCGGTCCG	GAATTAGGTG GAATTAGGTG	ΤΑΘΟΤΑΤΑΤΟΑΟ ΤΑΘΟΤΑΤΑΤΟΑΟ	AAGCTCATGT AAGCTCATGT	ΓΤΟΤΑΘΑΑΤΟ ΓΤΟΤΑΘΑΑΤΟ	TCAATCTGTAT TCAATCTGTAT	TTACTTGAAT	GATGCTACAA GATGCTACAA	АТСТССАТСА. АТСТССАТСА.	AAGTGAAGTG AAGTGAAGTG	GTTATTTATA GTTATTTATA	ATTGA

Figure 2. Nucleotide acid alignment of *atp6* genes sequences retrieved from the mtDNA of male fertile (MF, *atp6*⁺) and cytoplasmic male sterile (CMS, *atp6*⁻) accessions of *Foeniculum vulgare*.

2.3. Selection of Candidate Gene Controlling Cytoplasmic Male Sterility

Following the results obtained for *atp6*, raw sequencing reads from both MF and CMS GenBank accessions were mapped back to *atp6*⁺ and *atp6*⁻. As reported in Figure 3, panel A, MF reads were successfully mapped only to the *atp6*⁺ copy, while CMS reads aligned both against *atp6*⁺ and *atp6*⁻. However, the alignment coverage of these latter reads against the *atp6*⁺ copy was considerably lower than the coverage exhibited against the *atp6*⁻ (Figure 3, panel A).



Figure 3. (**A**) Nucleotide alignment between the two *atp6* copies retrieved within the mitochondrial genomes of *Foeniculum vulgare. atp6*⁺ indicates the wild type gene (1167 bp) found in the male fertile (MF) mtDNA while *atp6*⁻ is an *atp6*-like sequence lacking the first 300 bp (864 bp) retrieved in the cytoplasmic male sterile (CMS) mtDNA. Raw sequencing reads from MF and CMS accessions were mapped to *atp6*⁺ and *atp6*⁻. (**B**) PCR profiles generated in two CMS commercial lines and their MF maintainers with specific primers for *atp6*⁺(left) and *atp6*⁻ (right). (**C**) Protein alignment of *atp6*⁺, *atp6*⁻, and the *atp6* protein from *Daucus carota* (Dc_atp6).

The results obtained from the reads mapping were further confirmed by Sanger sequencing two different CMS lines and their MF maintainers by means of two different gene-specific primer sets. As shown in Figure 3, panel B, male fertile samples exhibited PCR amplicons only for *atp6*⁺, while the CMS accessions were successfully amplified at both loci. Although *atp6*⁺ primers binding sites were found conserved in both CMS and MF accessions, the intensity of the band for *atp6*⁺ was very low for male sterile individuals (Figure 3B, the original picture of the agarose gel is available as Figure S3). The same result was also confirmed in a further amplification experiment conducted on four additional different commercial CMS accessions and their MF maintainers (Figure S4).

Finally, from the alignment of the resulting proteins of $atp6^+$ (388 amino acids) and $atp6^-$ (287 amino acids) with atp6 of *Daucus carota* (Dc_atp6, 383 amino acids), the percentage of pairwise identical residues was 92.3% between $atp6^+$ and Dc_atp6 and 90.2% between $atp6^-$ and Dc_atp6 (Figure 3, panel C). This finding, along with the comparable sequence length of $atp6^+$ and Dc_atp6 (i.e., 388 vs. 383 amino acids), led to the hypothesis that $atp6^+$ represents the wt allele.

2.4. Expression Analyses of atp6⁺, atp6⁻, atp1, and atp9

The expression patterns of four genes, *atp6*⁺, *atp6*⁻ *atp1*, and *atp9* were successfully analyzed in leaves and flowers at three different developmental stages (0% flowering, 50% flowering and 100% flowering) of two CMS commercial lines and their MF maintainer (Figure 4). *atp6*⁺ and *atp6*⁻ were found to be constantly and exclusively expressed only in MF or CMS, respectively (Figure 4, panels A and B). Vice versa, *atp1* and *atp9*, used here as positive controls, were constantly expressed in all the tissues considered without significant differences between CMS and MF samples (Figure 4, panels C and D).



Figure 4. Evaluation of $atp6^+$ (**A**), $atp6^-$ (**B**), atp1 (**C**), and atp9 (**D**) mRNA expression by qPCR analysis. Expression analysis were conducted on male fertile (MF, n = 4, red triangles) and male sterile samples (CMS, n = 4, blue squares). Four different tissues were evaluated for each gene: leaf (L), flower before anthesis (0% flowering, FBA), flower in half anthesis (50% flowering, FHA) and flower in full anthesis (100% flowering, FFA). Ct values were normalized for *gapdh*. Statistical significance was evaluated with unpaired Student's *t*-test (* p < 0.05). All data are shown as mean \pm SD.

3. Discussion

In the last five years, the knowledge about extranuclear genomes of the Apiaceae family, whose aromatic plants are extensively used for several purposes, including food products (e.g., celery and fennel), beverages (e.g., carrot), flavoring components (e.g., cumin), cosmetics (e.g., tōki), and pharmaceuticals (e.g., carrot and fennel) [35,36], has considerably advanced thanks to the advent of NGS technologies. Several studies have greatly contributed not only to uncover the sequence, but also to comprehend the structure, expression, and variation of these organelles [37–39]. However, it must be acknowledged that in the face of 42 chloroplast (cp) genomes deposited in NCBI from as many Apiaceae species, only two mitochondrial (mt) genomes are available (i.e., Daucus carota and Bupleurum falcatum). A similar discrepancy is also observed at a wider level (i.e., plant level), being the number of plastid genomes (3.720) accessible in NCBI ten times larger than that of mitochondrial assemblies (311). This numerical disproportion can be attributed to several factors. First, in contrast to the cpDNA genome, mtDNA is more often characterized by heteroplasmy phenomena (i.e., co-existence of additional mitochondrial genome types called mitotypes) [40] and by a variety of different arrangements and configurations (circular, linear or branched molecules) [41–43] that can strongly affect their correct and exhaustive assembly. Moreover, when compared to cpDNA genomes, mtDNA has proven to be bigger and to contain more repeat regions as well as significant amounts of nuclear and chloroplast genomic sequences [44]. All these peculiarities, along with technical issues linked to the amount and the type of available reads (short vs. long), may limit the accurate assembly of this organelle genome. These observations would explain the different findings achieved in this study for the mtDNA assembly of male sterile and male fertile accessions of Foeniculum vulgare. We successfully managed to assembly the CMS mtDNA in a single molecule while the MF mtDNA was assembled only to a contig level. The main reason behind this discrepancy lies on the huge gap in terms of raw sequences available for the two contrasting phenotypes, being the MF reads accessible in GenBank less than a twentieth of the CMS ones. In fact, the poor coverage characterizing MF reads has complicated the assembly of extra genic regions and repetitive structures leading to the translation into gaps and incomplete results. On the contrary, the assembly of genic regions were comparable between the two mtDNAs and, in both cases, we were able to reconstruct the full sequences of 33 genes (excluding ORFs and tRNAs). This number is in agreement with what reported by Kubo and Newton [45], and Morley and Nielsen [46]: Excluding tRNAs and all the species-specific ORFs, the number of mitochondrial genes usually ranges from 27 (Beta vulgaris) to 39 (Zea Mays).

Based on the most complete assembly (the CMS one), the estimated size of the fennel mtDNA assembled in this study is 296,483 bp, 1.94 greater than the fennel plastome (153 kb) assembled by Peery et al. [28]. This is in line with what observed for *Daucus carota* (from the same family, Apiaceae) whose mtDNA is 281,132 bp in length and the ratio mtDNA/cpDNA is 1.80.

The SNP calling analysis, performed mapping the MF contigs to the CMS mtDNA, allowed us to detect 42 polymorphic loci able to discriminate the two antagonist phenotypes (i.e., MF and CMS plant genotypes) and, most importantly, to reveal the presence of two different *atp6* copies. *atp6*⁺, considered the wild type copy based on the highest similarity with *atp6* in *Daucus carota* was detected only in the male fertile mtDNA, *atp6*⁻, the putative mutant copy with a 300 bp truncation at the 5'-end, was available only in the mtDNA of male sterile samples. As already observed for *Rhazya stricta* [47] and *Arabidopsis thaliana* [48], *atp6*, compared to other mt genes, is often characterized by extended repeat sequences located at the upstream region suggesting that this gene is more likely involved in genomic recombination events. Despite the data available for the MF mtDNA are not sufficient to demonstrate the presence of repeated regions in the 5' region of *atp6*⁺, it is reasonable to assume that the truncation of 300 bp in *atp6*⁻ is attributable to a mitochondrial rearrangement.

Following the validation analyses (i.e., mapping back the sequencing reads to both *atp6* copies and performing PCR based on copies-specific primers) two main findings were achieved. First, *atp6⁻* was confirmed to be present only in the CMS samples and, thus, the F2/R2 primer couple (namely the one used to specifically amplify the *atp6⁻* copy) can be directly used in a molecular assay suitable

for breeding purposes and for selecting only CMS samples. Second, $atp6^+$ was detected also in CMS samples. However, the intensity of the $atp6^+$ PCR band and the reads coverage was lower when compared to the mutant ($atp6^-$) counterpart (Figure 3A,B and Figures S3 and S4), suggesting that $atp6^+$ is present as single copy in the nuclear genome, while $atp6^-$ is present in the mitochondrial genome, as supported by the CMS mtDNA assembly, the higher sequencing coverage and greater band intensity. A possible transfer of the atp6 gene to the nuclear genome would not be exclusive to *Foeniculum vulgare:* Genes duplication and/or relocation from the mitochondrion to the nucleus is a recurrent and consistent feature of eukaryotic genome evolution [49] and this phenomenon has already been described both in plants [50] and animals [51]. This was also confirmed in *Daucus carota* from an *ex-post* analysis that we performed aligning the mitochondrial atp6 gene (NC_017855.1:6101-7252) to the entire genome: An almost identical copy of this gene was found also within the chromosome 7 (NC_030387.1:11491922-11492723, E-value = 0.0, % identity = 99.50). It is also reasonable to think that a nuclear copy of atp6 is also present in MF genotypes, but, unfortunately, this hypothesis cannot be verified because the only whole genome sequencing available for fennel comes from a CMS individual [2].

The transcription analyses added another piece to this increasingly complex puzzle. While other ATP synthase subunits (i.e., *atp1* and *atp9*) were constantly and equally expressed (i.e., no significant expression difference observed between CMS and MF), *atp6*⁺ was ubiquitously expressed in all the tissues of the MF samples and never detected in any CMS tissues. Thus, it can be postulated that the nuclear *atp6*⁺ copy is probably a pseudogene. Moreover, despite the 300 bp truncation at the 5'-end, *atp6*⁻ mRNA was found to be constantly available in all the tissues analyzed for the CMS. The main findings are summarized in Figure 5.



Gene duplication and transfer event
 mtDNA rearrangement and atp6 truncation

Figure 5. Involvement of *atp6* in cytoplasmic male sterility in fennel. The *atp6*⁺ gene duplication and transfer from the mtDNA to the nuclear genome is a recurrent and consistent feature of eukaryotic genome evolution that seems to have characterized also the evolutionary path of *Foeniculum vulgare*. Subsequently, a mitochondrial rearrangement event produced a truncation of 300 bp in the mtDNA *atp6* copy (*atp6*⁻) leading to a cytoplasmic male sterility phenotype. Expression analyses revealed that (i) *atp6*⁺ gene is constantly expressed in male fertile plants (MF), where it is available both in the mtDNA and, probably, in the nuclear DNA (ii) *atp6*⁺ gene is never expressed in male sterile plants, where is available only in the nuclear DNA (probably as pseudogene) (iii) *atp6*⁻ is constantly expressed in male sterile plants.

It is well known that *atp6* is a key component of the transmembrane F_0 portion of the ATP synthase and it is frequently associated with CMS phenomena both in monocotyledon species such as rice [52], maize [53], and dicotyledon species such as pepper [5], sugar beet [13], brown mustard [12], carrot [39], kenaf [54], and sunflower [55].

Our results fit with the "energy deficiency model" according to which, the partial or total loss of function of one or more subunits belonging to the mtETC (complexes I-IV) or to the ATP synthase (complex V), can determine mitochondrial deficiencies [1]. This is supported by several evidences available in the scientific literature. In CMS sugar beet a truncated copy of *cox2* (a subunit of complex IV) showed a decreased activity when compared to the wt copy and proved a molecular link between mtETC defect and male sterility [56]. In *pol* CMS of *Brassica napus* the mutant *atp6* protein is thought to diminish the functionality of the complex V [57] while in CMS sunflower a chimeric protein orf522 seems to reduce the ATPase activity [58]. In *Foeniculum vulgare*, based on our findings, we assumed that *atp6*⁻, being the only gene expressed in CMS fennel plants, must be also translated into protein. It can be assumed that *atp6*⁻ proteins may affect ATP production by forming less efficient ATPase complexes that guarantee essential cellular processes (and thus the plant growth) but not a high energy demand process such as pollen development. Further analyses of post-translation events are necessary to determine the role of *atp6*⁻ on the male-sterility induction.

In conclusion, our study provides useful findings and novel insights into the fennel mtDNA genome and its *atp6* genes, and paves the way for functional studies in order to shed light on their roles in the determination of male sterility.

4. Materials and Methods

4.1. Assembly of Mitochondrial DNA

Two SRA sequence datasets, namely SRX2770225, and SRX7730623, were downloaded from the National Center for Biotechnology Information (NCBI) GenBank, under the taxonomy ID: 48038 (*Foeniculum vulgare*). The first archive was obtained from a male fertile sample (hereafter named "MF reads"), whereas SRX7730623 reads derived from a cytoplasmic male sterile accession (hereafter named "CMS reads").

CMS reads were assembled using GetOrganelle software [59], a fast de novo genome organelle assembler. The program was run providing as seed the mitochondrial genome of *D. carota* (GenBank accession NC_017855.1), using the option -fast and setting the k-mers length used by SPAdes assembler [60] to 21, 65, and 105.

Since GetOrganelle is not optimized to run with Ion GeneStudio S5 sequences, an alternative approach was used instead to assemble the MF mitochondrial genome. At first, the reads set was aligned to the CMS mitochondrial genome using Bowtie2 program [61]. All the mapping reads were retrieved, and de novo assembled using SPAdes [60]. Spades was run setting the option –iontorrent and adjusting the k-mers length to 21,33,55,77,99 and 127.

Both genomes were finally annotated by means of GeSeq [34] using the mtDNA of *Daucus carota* (NC_017855.1) as a reference.

4.2. Variant Calling and Sanger DNA Sequencing Validation

The two new assembled genomes were compared to detect any variant. Contigs assembled from MF reads were mapped against the CMS mtDNA, using the "basic variant detection" tool (CLC Genomics Workbench 11.0.1, CLC Bio).

Following the results obtained from the variant calling analysis, raw reads from both MF and CMS accessions were mapped against the *atp6* gene copy found within the mtDNA fertile (defined as *atp6*⁺) and the *atp6* gene copy found within the mtDNA sterile (named *atp6*⁻). Mapping was performed using the "map reads to contigs" tool of CLC Genomics Workbench 11.0.1 software (CLC Bio, Århus, Denmark) at default settings.

The two *atp6* gene copies were also validated through a Sanger sequencing. Geneious 11.1.4 software (http://www.geneious.com [62]) was used to design primers, for the specific amplification of both *atp6* copies. In details, for the analysis of $atp6^+$, the following primers were used: F1: ATCGACCTGAACAACATATACGGA R1: GCTGGCGATTTCCGACAAGT while for $atp6^-$, F2: TCACTGAGCACTGTCTG and R2: CCTAGAGTCTTTCGATACTATA were employed. PCR reactions were performed on genomic DNA purified (DNAeasy Plant Mini Kit, Qiagen, Valencia, CA, USA), from 4 samples: Two CMS individuals from as many commercial lines and two samples from their two isogenic MF maintainers. Reactions were set up in a 20 μ L reaction volume, containing 1× Platinum[®] Multiplex PCR Master Mix (Thermo Scientific, Pittsburgh, PA, USA), 0.5 µM of primers, 15 ng of genomic DNA and water up to volume A Veriti 96-Well Thermal Cycler (Applied Biosystems, Foster City, CA, USA) was programmed with the following conditions: initial denaturation at 95 °C for 5 min followed by 40 cycles at 95 °C for 30 s, 58 or 53 °C (*atp6*⁺ or *atp6*⁻, respectively) for 30 s, 72 °C 90 s, and final extension of 10 min at 72 °C. Amplification was evaluated on a 1.5% (w/v) agarose gel stained with 1 x SYBR SafeTM DNA Gel Stain (Life Technologies, Carlsbad, CA, USA). Amplicons were purified through ExoSap-IT (Applied Biosystems), sequenced (ABI 3730 DNA Analyzer, Applied Biosystem) using both forward and reverse gene specific primers, analyzed in Geneious 11.1.4 and deposited to GenBank (MT199323 and MT199324).

4.3. RNA Isolation and Quantitative Real-Time PCR

The same four lines used for DNA sequencing validation (i.e., two CMS commercial lines and their two isogenic MF maintainers) were used for expression analyses. From each line we collected (in two biological replicates) four tissues: leaf, flower before anthesis (0% flowering), flower in half anthesis (50% flowering) and flower in full anthesis (100% flowering). Samples were ground in liquid nitrogen, and total RNA was extracted using the SpectrumTM Plant Total RNA Kit (Sigma-Aldrich, St. Louis, MO, USA) following the manufacturer's instructions. RNA quality and quantity were checked by means of conventional electrophoresis and spectrophotometry using a NanoDrop-1000 (Thermo Scientific). cDNA was synthetized starting from 500 ng of RNA using the SuperScriptTM IV VILO Master Mix (Thermo Scientific) according to the manufacturer's instructions.

Five genes, namely $atp6^+$, $atp6^-$, atp1, and atp9 and gapdh (this latter used as housekeeping gene) were selected and primers were designed using Geneious 11.1.4 (Table S1). All the qPCRs were performed on a QuantStudio 3 machine (Thermo Scientific) following the PowerUp SYBR Green Master Mix method (Applied Biosystems). Each reaction was carried out in a volume of 10 µL containing 5 µL of 2× SYBR Green, 0.5 µM of primers, 2 µL of 1:10-diluted cDNA and water up to volume. The run method set was as follows: initial denaturation 95 °C for 20 s, followed by 40 cycles of denaturation at 95 °C for 3 s and primer annealing, extension and gathering the fluorescence signal at 60 °C for 30 s. Subsequently, the melting curve analysis was achieved to verify the specificity of the primer with the following program: 95 °C/15 s, 60 °C/1 min, and 95 °C/15 s. The output data were analyzed according to Livak et al. [63]. The baseline and threshold cycles (Ct) were automatically determined by the software of the system and the three technical replicates taken for each biological replicate were averaged. The resulting mean Ct values were normalized as difference in Ct values (Δ Ct) between the analyzed mRNA and *gapdh* reference gene. The Δ Ct values of each sample were then normalized with respect to the Δ Ct values of the MF samples (Δ \DeltaCt). Finally, the variation was reported as fold change ($2^{-\Delta\Delta}Ct$).

Supplementary Materials: Supplementary materials can be found at http://www.mdpi.com/1422-0067/21/13/ 4664/s1.

Author Contributions: Conceptualization, F.P.; Methodology, F.P. and G.M.; Software, N.V. and F.P.; Writing—original draft preparation, F.P. and A.V.; Writing—review and editing, F.P., A.V., G.M., N.V. and G.B.; Funding acquisition, G.B. All authors have read and agreed to the published version of the manuscript.

Acknowledgments: We wish to thank Alice Patella and Domiziana Battaggia for their help with some PCR experiments performed at the Laboratory of Genomics for Breeding, University of Padova (Italy). We wish to thank also Ian Dry and Elena Gonnelli for many helpful discussions on cytoplasmic male sterility.

Conflicts of Interest: The authors declare no conflicts of interest

References

- 1. Chen, L.; Liu, Y.-G. Male Sterility and Fertility Restoration in Crops. *Annu. Rev. Plant Biol.* **2014**, *65*, 579–606. [CrossRef] [PubMed]
- Palumbo, F.; Galla, G.; Vitulo, N.; Barcaccia, G. First draft genome sequencing of fennel (*Foeniculum vulgare* Mill.): Identification of simple sequence repeats and their application in marker-assisted breeding. *Mol. Breed.* 2018, *38*, 122. [CrossRef]
- 3. Tester, M.; Langridge, P. Breeding Technologies to Increase Crop Production in a Changing World. *Science* **2010**, *327*, 818–822. [CrossRef] [PubMed]
- 4. Prasannan, A. Hybrid Seeds Market by Crop, Duration, Seed Treatment, Farm Type and Acreage: Global Opportunity Analysis and Industry Forecast 2017-2023; Allied Market Research: Portland, OR, USA, 2017.
- 5. Wang, P.; Lu, Q.; Ai, Y.; Wang, Y.; Li, T.; Wu, L.; Liu, J.; Cheng, Q.; Sun, L.; Shen, H. Candidate gene selection for cytoplasmic male sterility in pepper (*Capsicum annuum* L.) through whole mitochondrial genome sequencing. *Int. J. Mol. Sci.* **2019**, *20*, 578. [CrossRef]
- 6. Budar, F.; Pelletier, G. Male sterility in plants: Occurrence, determinism, significance and use. *Comptes Rendus l'Academie des Sci. Ser. III* **2001**, *324*, 543–550. [CrossRef]
- 7. Feng, X.; Kaur, A.P.; MacKenzie, S.A.; Dweikat, I.M. Substoichiometric shifting in the fertility reversion of cytoplasmic male sterile pearl millet. *Theor. Appl. Genet.* **2009**, *118*, 1361–1370. [CrossRef]
- 8. Priyadarshan, P.M. Male sterility. In *Plant breeding: Classical to Modern;* Priyadarshan, P.M., Ed.; Springer: Singapore, 2019; pp. 105–129. ISBN 9789811370953.
- 9. Kim, B.; Yang, T.J.; Kim, S. Identification of a gene responsible for cytoplasmic male-sterility in onions (*Allium cepa* L.) using comparative analysis of mitochondrial genome sequences of two recently diverged cytoplasms. *Theor. Appl. Genet.* **2019**, *132*, 313–322. [CrossRef]
- 10. Park, J.Y.; Lee, Y.P.; Lee, J.; Choi, B.S.; Kim, S.; Yang, T.J. Complete mitochondrial genome sequence and identification of a candidate gene responsible for cytoplasmic male sterility in radish (*Raphanus sativus* L.) containing DCGMS cytoplasm. *Theor. Appl. Genet.* **2013**, *126*, 1763–1774. [CrossRef]
- 11. Nakajima, Y.; Yamamoto, T.; Muranaka, T.; Oeda, K. A novel orfB-related gene of carrot mitochondrial genomes that is associated with homeotic cytoplasmic male sterility (CMS). *Plant Mol. Biol.* **2001**, *46*, 99–107. [CrossRef]
- 12. Jing, B.; Heng, S.; Tong, D.; Wan, Z.; Fu, T.; Tu, J.; Ma, C.; Yi, B.; Wen, J.; Shen, J. A male sterility-associated cytotoxic protein ORF288 in Brassica juncea causes aborted pollen development. *J. Exp. Bot.* **2012**, *63*, 1285–1295. [CrossRef]
- Yamamoto, M.P.; Kubo, T.; Mikami, T. The 5'-leader sequence of sugar beet mitochondrial atp6 encodes a novel polypeptide that is characteristic of Owen cytoplasmic male sterility. *Mol. Genet. Genomics* 2005, 273, 342–349. [CrossRef] [PubMed]
- Wang, Z.; Zou, Y.; Li, X.; Zhang, Q.; Chen, L.; Wu, H.; Su, D.; Chen, Y.; Guo, J.; Luo, D.; et al. Cytoplasmic Male Sterility of Rice with Boro II Cytoplasm Is Caused by a Cytotoxic Peptide and is restored by Two Related PPR Motif Genes via Distinct Modes of mRNA Silencing. *Plant Cell* 2006, 18, 676–687. [CrossRef] [PubMed]
- 15. Tang, H.V.; Pring, D.R.; Shaw, L.C.; Salazar, R.A.; Muza, F.R.; Yan, B.; Schertz, K.F. Transcript processing internal to a mitochondrial open reading frame is correlated with fertility restoration in male-sterile sorghum. *Plant J.* **1996**, *10*, 123–133. [CrossRef] [PubMed]
- 16. National Library of Medicine National Center for Biotechnology Information. Nucleotide Database. Available online: https://www.ncbi.nlm.nih.gov/nucleotide/ (accessed on 30 June 2020).
- 17. Liu, C.G.; Hou, N.; Liu, L.K.; Liu, J.C.; Kang, X.S.; Zhang, A.M. A YA-type cytoplasmic male-sterile source in common wheat. *Plant Breed.* **2006**, *125*, 437–440. [CrossRef]
- 18. Huang, J.Z.; Zhi-Guo, E.; Zhang, H.L.; Shu, Q.Y. Workable male sterility systems for hybrid rice: Genetics, biochemistry, molecular biology, and utilization. *Rice* **2014**, *7*, 1–14. [CrossRef]

- Gabay-Laughnan, S.; Laughnan, J.R. Male sterility and restorer genes in maize. In *The Maize Handbook*; Freeling, M., Walbot, V., Eds.; Springer-Verlag: New York, NY, USA, 1994; pp. 418–423.
- Zhao, N.; Xu, X.; Wamboldt, Y.; Mackenzie, S.A.; Yang, X.; Hu, Z.; Yang, J.; Zhang, M. MutS HOMOLOG1 silencing mediates ORF220 substoichiometric shifting and causes male sterility in *Brassica juncea*. *J. Exp. Bot.* 2016, 67, 435–444. [CrossRef]
- Heng, S.; Liu, S.; Xia, C.; Tang, H.Y.; Xie, F.; Fu, T.; Wan, Z. Morphological and genetic characterization of a new cytoplasmic male sterility system (oxa CMS) in stem mustard (*Brassica juncea*). *Theor. Appl. Genet.* 2018, 131, 59–66. [CrossRef]
- 22. Wang, Y.; Wang, Q.; Hao, W.; Li, J.; Qi, M.; Zhang, L. Mitochondrial Genome Sequencing Reveals orf463a May Induce Male Sterility in NWB Cytoplasm. *Genes* **2020**, *11*, 74. [CrossRef]
- 23. Hans Köhler, R.; Horn, R.; Lössl, A.; Zetsche, K. Cytoplasmic male sterility in sunflower is correlated with the co-transcription of a new open reading frame with the atpA gene. *Mol. Gen. Genet.* **1991**, 227, 369–376. [CrossRef]
- 24. Quagliotti, L.; Accati, E.; Milanesio, M. Preliminary report on a functional type of male-sterility in fennel. In Proceedings of the 5th Eucarpia Congress, Milan, Italy, 30 September–2 October 1968; pp. 359–363.
- 25. Ramanujam, S.; Tewari, V.P. Heterosis breeding in fennel. Indian J. Genet. Plant Breed. 1970, 30, 732–737.
- 26. Bruins, M. Modernising fennel. Available online: https://european-seed.com/2015/10/modernising-fennel/. (accessed on 30 June 2020).
- 27. Food and Agriculture Organization (FAO), FAO Food and Agriculture Organization of the United Nations: Value of Agricultural Production. Available online: http://www.fao.org/faostat/ (accessed on 30 June 2020).
- Peery, R.; Kuehl, J.V.; Boore, J.L.; Raubeson, L. Comparisons of three Apiaceae chloroplast genomes—coriander, dill and fennel. In Proceedings of the Botany 2006 Meeting, Chico, CA, USA, 2006, 28 July–2 August 2006.
- 29. Palumbo, F.; Vannozzi, A.; Vitulo, N.; Lucchin, M.; Barcaccia, G. The leaf transcriptome of fennel (*Foeniculum vulgare* Mill.) enables characterization of the t-anethole pathway and the discovery of microsatellites and single-nucleotide variants. *Sci. Rep.* **2018**, *8*, 1–12. [CrossRef] [PubMed]
- 30. Li, S.; Chen, Z.; Zhao, N.; Wang, Y.; Nie, H.; Hua, J. The comparison of four mitochondrial genomes reveals cytoplasmic male sterility candidate genes in cotton. *BMC Genomics* **2018**, *19*, 1–15. [CrossRef]
- Allen, J.O.; Fauron, C.M.; Minx, P.; Roark, L.; Oddiraju, S.; Guan, N.L.; Meyer, L.; Sun, H.; Kim, K.; Wang, C.; et al. Comparisons among two fertile and three male-sterile mitochondrial genomes of maize. *Genetics* 2007, 177, 1173–1192. [CrossRef]
- 32. National Center for Biotechnology Information (NCBI), Organelle Genome Resources. Available online: https://www.ncbi.nlm.nih.gov/genome/organelle/ (accessed on 30 June 2020).
- Wu, Z.; Hu, K.; Yan, M.; Song, L.; Wen, J.; Ma, C.; Shen, J.; Fu, T.; Yi, B.; Tu, J. Mitochondrial genome and transcriptome analysis of five alloplasmic male-sterile lines in *Brassica juncea*. *BMC Genomics* 2019, 20, 1–15. [CrossRef] [PubMed]
- 34. Tillich, M.; Lehwark, P.; Pellizzer, T.; Ulbricht-Jones, E.S.; Fischer, A.; Bock, R.; Greiner, S. GeSeq–Versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* **2017**, *45*, W6–W11. [CrossRef]
- 35. Ngahang Kamte, S.L.; Ranjbarian, F.; Cianfaglione, K.; Sut, S.; Dall'Acqua, S.; Bruno, M.; Afshar, F.H.; Iannarelli, R.; Benelli, G.; Cappellacci, L.; et al. Identification of highly effective antitrypanosomal compounds in essential oils from the Apiaceae family. *Ecotoxicol. Environ. Saf.* **2018**, *156*, 154–165. [CrossRef] [PubMed]
- Sayed-Ahmad, B.; Talou, T.; Saad, Z.; Hijazi, A.; Merah, O. The Apiaceae: Ethnomedicinal family as source for industrial uses. *Ind. Crops Prod.* 2017, 109, 661–671. [CrossRef]
- Lee, H.O.; Joh, H.J.; Kim, K.; Lee, S.C.; Kim, N.H.; Park, J.Y.; Park, H.S.; Park, M.S.; Kim, S.; Kwak, M.; et al. Dynamic chloroplast genome rearrangement and DNA barcoding for three apiaceae species known as the medicinal herb "bang-poong". *Int. J. Mol. Sci.* 2019, 20, 2196. [CrossRef]
- 38. Yao, X.Y.; Chen, Z.X.; Wang, Q.Z. The complete chloroplast genome of *Torilis scabra* (Apiaceae). *Mitochondrial DNA Part B Resour.* **2019**, *4*, 2914–2915. [CrossRef]
- 39. Tan, G.F.; Wang, F.; Zhang, X.Y.; Xiong, A.S. Different lengths, copies and expression levels of the mitochondrial atp6 gene in male sterile and fertile lines of carrot (*Daucus carota* L.). *Mitochondrial DNA Part A DNA Mapp. Seq. Anal.* **2018**, *29*, 446–454. [CrossRef]
- 40. Gualberto, J.M.; Newton, K.J. Plant Mitochondrial Genomes: Dynamics and Mechanisms of Mutation. *Annu. Rev. Plant Biol.* **2017**, *68*, 225–252. [CrossRef]

- Kozik, A.; Rowan, B.A.; Lavelle, D.; Berke, L.; Eric Schranz, M.; Michelmore, R.W.; Christensen, A.C. The alternative reality of plant mitochondrial DNA: One ring does not rule them all. *PLoS Genet.* 2019, 15, 1–30. [CrossRef] [PubMed]
- 42. Sloan, D.B. One ring to rule them all? Genome sequencing provides new insights into the "master circle" model of plant mitochondrial DNA structure. *New Phytol.* **2013**, 200, 978–985. [CrossRef] [PubMed]
- 43. Mower, J.P.; Case, A.L.; Floro, E.R.; Willis, J.H. Evidence against equimolarity of large repeat arrangements and a predominant master circle structure of the mitochondrial genome from a monkeyflower (*Mimulus guttatus*) lineage with cryptic CMS. *Genome Biol. Evol.* **2012**, *4*, 670–686. [CrossRef]
- 44. Zhang, F.; Li, W.; Gao, C.W.; Zhang, D.; Gao, L.Z. Deciphering tea tree chloroplast and mitochondrial genomes of *Camellia sinensis* var. *assamica. Sci. Data* **2019**, *6*, 209. [CrossRef]
- 45. Kubo, T.; Newton, K.J. Angiosperm mitochondrial genomes and mutations. *Mitochondrion* **2008**, *8*, 5–14. [CrossRef]
- 46. Morley, S.A.; Nielsen, B.L. Plant mitochondrial DNA. Front. Biosci. 2017, 22, 1023–1032.
- 47. Park, S.; Ruhlman, T.A.; Sabir, J.S.M.; Mutwakil, M.H.Z.; Baeshen, M.N.; Sabir, M.J.; Baeshen, N.A.; Jansen, R.K. Complete sequences of organelle genomes from the medicinal plant *Rhazya stricta* (Apocynaceae) and contrasting patterns of mitochondrial genome evolution across asterids. *BMC Genomics* **2014**, *15*, 1–18. [CrossRef]
- Marienfeld, J.; Unseld, M.; Brandt, P.; Brennicke, A. Genomic recombination of the mitochondrial atp6 gene in *Arabidopsis thaliana* at the protein processing site creates two different presequences. *DNA Res.* 1996, *3*, 287–290. [CrossRef]
- 49. Brandvain, Y.; Wade, M.J. The functional transfer of genes from the mitochondria to the nucleus: The effects of selection, mutation, population size and rate of self-fertilization. *Genetics* **2009**, *182*, 1129–1139. [CrossRef]
- 50. Funes, S.; Davidson, E.; Gonzalo Claros, M.; Van Lis, R.; Pérez-Martínez, X.; Vázquez-Acevedo, M.; King, M.P.; González-Halphen, D. The typically mitochondrial DNA-encoded ATP6 subunit of the F1F0-ATPase is encoded by a nuclear gene in *Chlamydomonas reinhardtii*. *J. Biol. Chem.* **2002**, 277, 6051–6058. [CrossRef] [PubMed]
- Kohn, A.B.; Citarella, M.R.; Kocot, K.M.; Bobkova, Y.V.; Halanych, K.M.; Moroz, L.L. Rapid evolution of the compact and unusual mitochondrial genome in the ctenophore, *Pleurobrachia bachei*. *Mol. Phylogenet. Evol.* 2012, 63, 203–207. [CrossRef] [PubMed]
- 52. Itabashi, E.; Kazama, T.; Toriyama, K. Characterization of cytoplasmic male sterility of rice with Lead Rice cytoplasm in comparison with that with Chinsurah Boro II cytoplasm. *Plant Cell Rep.* **2009**, *28*, 233–239. [CrossRef] [PubMed]
- 53. Dewey, R.E.; Timothy, D.H.; Levings, C.S. Chimeric mitochondrial genes expressed in the C male-sterile cytoplasm of maize. *Curr. Genet.* **1991**, *20*, 475–482. [CrossRef] [PubMed]
- 54. Zhao, Y.; Liao, X.; Zhou, B.; Zhao, H.; Zhou, Y.; Zhou, R. Mutation in the coding sequence of atp6 are associated with male sterile cytoplasm in kenaf (*Hibiscus cannabinus* L.). *Euphytica* **2016**, 207, 169–175. [CrossRef]
- 55. Makarenko, M.S.; Usatov, A.V.; Tatarinova, T.V.; Azarin, K.V.; Logacheva, M.D.; Gavrilova, V.A.; Kornienko, I.V.; Horn, R. Organization features of the mitochondrial genome of sunflower (*Helianthus annuus* L.) with ANN2-type male-sterile cytoplasm. *Plants* **2019**, *8*, 439. [CrossRef] [PubMed]
- 56. Ducos, E.; Touzet, P.; Boutry, M. The male sterile G cytoplasm of wild beet displays modified mitochondrial respiratory complexes. *Plant J.* **2001**, *26*, 171–180. [CrossRef]
- 57. An, H.; Yang, Z.; Yi, B.; Wen, J.; Shen, J.; Tu, J.; Ma, C.; Fu, T. Comparative transcript profiling of the fertile and sterile flower buds of pol CMS in *B. napus. BMC Genomics* **2014**, *15*, 2–11. [CrossRef]
- 58. Sabar, M.; Gagliardi, D.; Balk, J.; Leaver, C.J. ORFB is a subunit of F1F0-ATP synthase: Insight into the basis of cytoplasmic male sterility in sunflower. *EMBO Rep.* **2003**, *4*, 381–386. [CrossRef]
- 59. Jin, J.-J.; Yu, W.-B.; Yang, J.-B.; Song, Y.; DePamphilis, C.W.; Yi, T.-S.; Li, D.-Z. GetOrganelle: A fast and versatile toolkit for accurate de novo assembly of organelle genomes. *bioRxiv* **2019**, 1–31. [CrossRef]
- 60. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Prjibelski, A.D.; et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **2012**, *19*, 455–477. [CrossRef] [PubMed]
- 61. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [CrossRef] [PubMed]

- 62. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [CrossRef]
- Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔCT} method. *Methods* 2001, 25, 402–408. [CrossRef]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).