Theor Appl Genet (2002) 104:698–702 © Springer-Verlag 2002

T. Engelke · T. Tatlioglu

A PCR-marker for the CMS_1 inducing cytoplasm in chives derived from recombination events affecting the mitochondrial gene *atp*9

Received: 6 June 2001 / Accepted: 3 August 2001

Abstract The complete coding and 3´-flanking region of the mitochondrial gene *atp*9 of chives (*Allium schoenoprasum* L.) was determined in order to develop primers that allow the identification of *atp*9-related sequences in subsequent PCR-amplifications. One of these sequences is of a chimerical nature, consisting of *atp*9-homologous regions on its end, interrupted by an insertion that is composed of one *atp*6-homologous part and one part of unknown origin. This PCR-fragment is 762 bp in size and exclusively amplified in the sterility inducing cytoplasm of CMS_1. Thus it can be used as a PCR-marker in order to distinguish this cytoplasm type from the remaining cytoplasm types of chives. The chimerical marker sequence forms a putative open reading frame (*orf*A501), from which CMS_1 might originate.

Keywords *Allium schoenoprasum* · Cytoplasmic male sterility

Introduction

In *Allium schoenoprasum* we have previously described two CMS systems, differentiated by their phenotypical expression and histological behaviour during microsporogenesis, as well as the restorer genes involved, that have been distinguished by extended segregation analyses. In the sterility inducing cytoplasm of the first CMS system a stable restoration is obtained by the gene *X* (Tatlioglu 1982). A second restorer gene, *T*, only leads to pollen production at high temperatures (24°C/24°C, day/night), while the same genotypes remain male-sterile at lower temperatures (20°C/12°C, day/night) (Tatlioglu 1987). A third gene, *a*, causes pollen production in combination with tetracycline treatment (Tatlioglu and Wricke 1988). This CMS system will be referred to as CMS_1 below. The restoration in CMS_2 is controlled by the dominant gene *St2* and influences from different temperatures or tetracyline treatment on this sterility system have not yet been observed (Engelke and Tatlioglu 2000a). Both sterility inducing cytoplasms, as well as three different normal cytoplasms, have been distinguished by means of RFLP-analyses using mitochondrial gene probes (Engelke and Tatlioglu 2000b). The cytoplasm types (1), (2) and (3) will be called (N_1), (N_2) and (N_3) below to indicate their character as normal cytoplasm types, while the cytoplasm types (S) (Tatlioglu 1982) and (5) (Engelke and Tatlioglu 2000a) will be referred to as (S_1) and (S_2) to indicate their sterility inducing functions in CMS_1 and CMS_2 respectively.

As a result of the previously performed RFLP-analyses we discussed a possible function of the *atp*9 gene in generating the different cytoplasm types in chives, due to the fact that the heterologous probe containing the *atp*9-gene (*Bam*HI fragment from *Arabidopsis thaliana*, gene bank accession NC001284, Unseld et al. 1997), detected more fragments than all other mitochondrial gene probes used (Engelke and Tatlioglu 2000b). This gives a hint that at least parts of the coding or flanking regions of *atp*9 are repeated in the mitochondrial genome of *A. schoenoprasum*. As evidence of these assumptions, we have now isolated a complete sequence of the *atp*9-gene of *A. schoenoprasum* and subsequently designed PCR-primers to amplify different *atp*9-related sequences. The occurrence of these sequences in the above-mentioned different cytoplasm types has then be tested to develop a PCR-marker that allows the differentiation of the CMS_1-inducing cytoplasm from the other cytoplasm types.

Communicated by R. Hagemann

T. Engelke (✉) · T. Tatlioglu
Abteilung Angewandte Genetik, Universität Hannover,
Herrenhäuser Strasse 2, D-30419 Hannover, Germany
e-mail: thomas.engelke@genetik.uni-hannover.de
Tel.: +49-511-62-2672, Fax: +49-511-762-3608

Materials and methods

Plant materials

Each of the cytoplasm types, that were previously distinguished by RFLP-analyses, is presented in this study by three single plants

```
-1-45   5`AG ATG TTA GAA GGA GCT AAA TTA ATA GGT GCC GGA GCT GCT ACA ATT
46-93        GCT TCA GCG GGA GCG GCT GTC GGT ATT GGA AAC GTG TTC AGT TCT TTG
94-141       ATT CAT GGA GTG GCT AGA AAT CCA TCA TTG GCA AAA CAA TCA TTT GGT
142-189      TAT GCC ATT TTG GGC TTT GCT CTA ACC GAA GCT ATT GCA TTG TTT GCC
190-237      CTA ATG ATG GCC TTT CTG ATC TCA TTC GTC TTC CAA AGT GAA CCA GAT
238-285      GCT TAA AAG CGC TTT GAA TAA GTC ATT AAT TTT TCC TTT TTG AAA AAA
286-333      AAG AAG AAT AAT GGC AAA CAC ATA TAC CGG AAG GGG AGG GCC TCT TCC
334-381      ACC AGA GTA CTT TAC TTT CTG GAA TTC ATT CTT AAC TTC AAA AGT ACA
382-429      GGA ACT GGG TAA AAA AAT GGA CCG AGA ACT ACA TAT GGT CTG AGA CTC
430-474      ACT TCC TTG AAA AGA GGA GGC GAG CTA TGT AGG CTG TGG CCG TCA A3`
```

Fig. 1 Coding and 3´-flanking region of *atp*9, isolated from (S_1)-cytoplasm of *A. schoenoprasum.* **Bold letters** coding region of *atp*9. *Underlined primer-pair* (position 59–79 and 189–210) that has been used in subsequent PCR-amplifications

to show PCR-amplification in a first approach. The amplification product, which might be used as a marker has subsequently been tested in plants from a backcross programme to show the maternal inheritance and to exclude segregation as would be expected from a marker for nuclear DNA. Furthermore, the influence from the different restorer genes, *X*/*x*, *T*/*t* and *A*/*a*, on the occurrence of the marker has been examined in segregating offspring. For the number of plants that have been used in these investigations see the corresponding figures.

Database analyses and the primer-design

Primers, that allow the amplification of the *atp*9 gene in chives, were developed from homologous regions of known *atp*9 sequences from monocotyle plants. Therefore, the following gene bank accessions were compared: *Hordeum vulgare*: X74365; *Lolium perenne*: Z50200; *Secale cereale*: X99020; *Sorghum bicolor*: U61165; *Triticum aestivum*: S47364; *Oryza sativa*: D38481; and *Zea mays*: M18339. Different primers were designed by using the programme Primer Premier (Premier Biosoft International) and have been synthesised by MWG Biotech AG (data not shown). Subsequently, the sequence information from the isolated *atp*9 gene of *A. schoenoprasum* was used to develop primer pairs that allow the amplification of *atp*9-related sequences. For the interesting primer sequences see Table 1.

Isolation of DNA and PCR-amplifications

Highly purified total DNA was isolated as previously described (Engelke and Tatlioglu 2000b). When the PCR-protocols had been optimised, an alternative DNA-isolation protocol, originally described by Edwards et al. (1991) was used, which is much less time and cost intensive and needs less plant material.

PCR was performed using 50 ng of total DNA in a final volume of 20 µl, containing 0.25 µM of each primer, 150 µM of each dNTP, and 0.25 units of Dynazyme in the reaction buffer as recommended by the supplier (Biometra). The reaction mixture was incubated in a thermocycler (Biometra) for 2 min at 94°C, followed by 40 cycles: 1 min at 94°C, 1 min for annealing, 2 min at 72°C, and a final extension for 30 min at 72°C, to allow the non-template-dependent terminal adding of a single deoxyadenosine (A) to the 3´-ends of PCR products that is needed for subsequent cloning. The annealing temperatures were calculated by the given T_M-values of the different primers and optimised, using the temperature gradient of the thermocycler. For the interesting annealing temperature of the developed marker see below. PCR-products were separated by flatbed electrophoresis using 1.5% agarose gels in 1×TAE buffer. Fragments for cloning were cut out of the gels and transferred to a 1% LM-agarose gel from which they were cut out again. LM-agarose was liquefied and digested by Gelase (Epicenter) according to the manufacturer's instructions.

Cloning and sequencing of amplified products

For cloning of the PCR-products a TOPO TA Cloning Kit was used as recommended by the supplier (Invitrogene). Analysis of positive clones were directly carried out by dipping the toothpick with a picked colony into a PCR-mix before preserving it on a new plate. The PCR-mix and the PCR-conditions were performed as described above, with T3/T7 primers and an annealing temperature of 50°C, so that the cloned fragments were specifically amplified. PCR-products were separated on 1% agarose gels in 1×TAE and the colonies with the correct insertion increased in liquid LB-medium. Plasmids were recovered using the QuiaPrep-Mini kit (Quiagen). Automatic sequencing was done by using the Abi Prism 377 DNA Sequencer (Applied Biosystems).

Results

The complete coding and 3´-flanking sequence of *atp*9 from *A. schoenoprasum*

Until now *atp*9 sequences of monocotyledons have been represented exclusively by members of the Poaceae (see Materials and methods) as the nearest relatives to chives. It was possible to design primers which allow PCR-amplification in *A. schoenoprasum*, since the sequence of *atp*9 is conserved over a wide range of the gene. Different primers have been designed anchoring in the beginning of the coding region and in the 3´-end or -flanking region, where homology between the different known sequences is strongly decreased. Using these primers, amplification products of expected sizes and additional fragments were amplified (data not shown). One of these (around 1,650 bp), that has been isolated from (S_1)-cytoplasm, contains the complete coding region and a large 3´-flanking region of *atp*9 (Fig. 1). The coding region shows high homology to corresponding *atp*9-sequences from different plant species. The highest homology was found to *S. cerale* (genebank accession number X99020) with 207 identical bases out of 223 (92%) at the 5´-end of the gene. The stop codon in *A. schoenoprasum* (TAA) is found in position 241, thus the predicted ATP9-protein consists of 80 amino acids.

A subsequently identified *atp*9-related sequence with a chimerical structure

The sequence information from *atp*9 of *A. schoenoprasum* was subsequently used to develop primer pairs that allow the amplification of *atp*9-related sequences. PCR was performed with DNA representing the previously distinguished five cytoplasm types. One of the primer pairs has been derived from position 59–79 and 189–210 of the coding region (Fig. 1) and shows an amplification product of 152 bp in all five cytoplasm types, that reflects the expected fragment of the *atp*9-gene. An additional

Fig. 2 Sequence of the CMS_1-specific 762-bp fragment. *Underlined primers* as given from the original *atp*9-gene from Fig. 1. *Framed 623 bases*, inserted in comparison with the original part of the *atp*9-gene from Fig. 1. *Crossed out* 13 bases, deleted in comparison with the original part of the *atp*9-gene from Fig. 1. *Italic letters atp*6-homologous sequence (138 bases). **Bold letters** *orf*A501

```
59-108    CG GCT GTC GGT ATT GGA AAC GTG TTC AGT TCT TTG ATT CAT GGA GTG GCT
109-156  AGA AAT CCA TCA TTG GCA AAA CAA TCA TTT GGT TAT GCG AAT GTC TCG
157-204  GGG ATC GCA CGA GTC CGC GGA CGG GCC ACT TCG CCC ACC AGG CGG GGT
205-252  TTC TCC ACA ACT CCG GAC CCA GAA ACC CCG GAG ACT AGG ATG GCT CGC
253-300  CTT GAA AGA GAG CAT GCT CGC CTT GAA AGA GAG CAT GCT CTC TTA CAA
301-348  GAA AAG TGG AAT TCC TCG TTA TGG GGG TTT CTA GGC TTG TTA TCA AGC
349-396  AGG TTT CGC CCC TGG TTT CTA TTA GGA TTC AGT TCG TTT TTA CTT CTT
397-444  TTG TGG GGG CAT TTA GTA CCT ATA TGG TAC TAC AGC GAG CGG CTT TCC
445-492  GCT GCC ATG AAT TAT TTT ATG TCG GAA GGC ATT CTT TTT GAG TTG GAA
493-540  TTC GGG TGG GAC AAA GTG GCG ATT CGC CCA GTT CAG TCG GTG GCG ACT
541-588  GTT ATC AAA GCA GGG GAA GTA CCT CCT GTT CCA CCA GTG ACG GGT ATA
589-636  GTA CCA CCC GTA GAA TCC CAC TGC CCA CTA GAG TGG AAC TTC TTC ACA
637-684  AAT CCA TCC TTG TCT ATG CTG CTC ACT CTC GGT TTG GTC CTA CTG ATG
685-732  ATT TTT GT-G TTA CGA AAA AGG GAG GGG GAA AGT CAG CGC CAA ATG CTT
733-769  GGC AAT CCT TGG TAG AGC TTA TTC ATG ATT TCG TGC GC ATT TTG GGC TTT
770-820   GC TCT AAC CGA AGC TAT TGC ATT GTT TGC CCT AAT GAT GGC CTT TCT GAT C
```

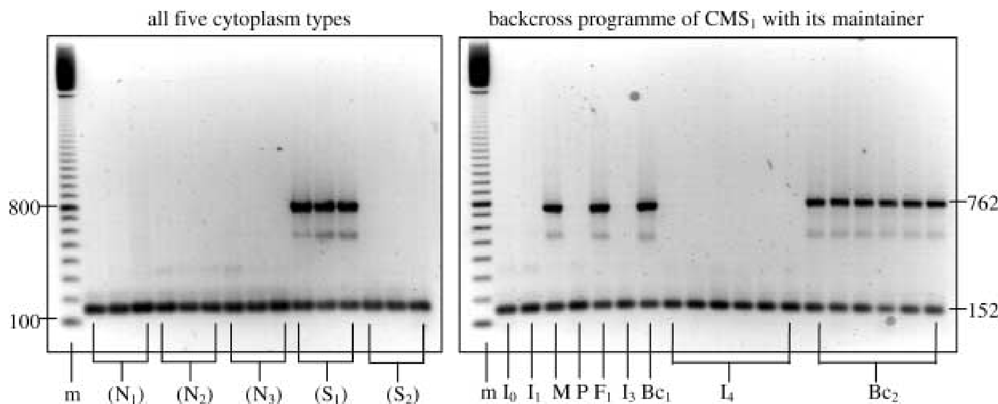

Fig. 3 PCR-marker for (S_1)-cytoplasm of CMS_1 (762 bp). *m*= 100 bp-ladder. (N_1), plants with normal cytoplasm type 1. (N_2), plants with normal cytoplasm type 2. (N_3), plants with normal cytoplasm type 3. (S_1), plants with sterility inducing cytoplasm of CMS_1. (S_2), plants with sterility inducing cytoplasm of CMS_2. I, inbreeding line of a maintainer; cytoplasm type: (N_3). M, male-sterile maternal parent; cytoplasm type: (S_1). P, paternal parent (second inbreeding generation of the maintainer). F_1, first filial generation. Bc, backcross generation. 152-bp fragment, part of the *atp*9-gene (for sequence see Fig. 1). 762-bp fragment, CMS_1-specific fragment (for sequence see Fig. 2)

762-bp fragment was exclusively amplified in cytoplasm-type (S_1). Both PCR fragments were solidly amplified over a wide range of annealing temperatures (35°C–53°C) and DNA template concentrations (2.5 ng→250 ng); we recommend 52°C and 100 ng.

The sequence of this fragment (Fig. 2) is of a chimerical nature, since the *atp*9 homologous sequence is interrupted at position 147. The following insertion of 623 bp in size is of unknown origin at its 5´-end, while 138 bp at its 3´-end are *atp*6-homologous. The highest homology was found to *S. bicolor* with 128 identical bases (92%) and one gap out of 139 bases beginning in position 548 of the *Sorghum atp*6 gene (gene bank accession number X57100).

The 3´-end of the 762-bp fragment corresponds to the remaining part of the *atp*9 homologous sequence of Fig. 1, except for a deletion of 13 bp. A new putative open reading frame of 501 bp is formed within the chimerical sequence, starting at position 244 and ending with TAG at position 745 (Fig. 2).

The chimerical fragment as a PCR-marker for the CMS_1 inducing cytoplasm

To be confident that the above-mentioned chimerical sequence (Fig. 2) is exclusively amplified in the CMS_1-inducing cytoplasm, we have compared all five cytoplasm types and followed the marker in a backcross programme of a sterile line with its maintainer (Fig. 3). The backcross programme shows the maternal inheritance, which was tested to provide evidence that the marker sequence is really of cytoplasmic nature and not part of the nuclear DNA. In the latter case the marker would segregate in a ratio of 1:1 in Bc_1, and would follow this segregation or would be lost in all of the subsequent backcross generations. The total loss of the marker as well as segregation in Bc_2 ($P_{1:1}$=1.56%, due to the sample size of six plants in Bc_2) can be excluded; therefore the fragment shows a constantly maternal inheritance.

Since it is known that restorer genes might have an influence on the mitochondrial genome organisation, offspring segregating for the genes *X*, *T* and *a*, which are involved in CMS_1, have been tested for the occurrence of the marker fragment. As can be seen in Fig. 4, an influence from the restorer genes on the existence of the marker can be ruled out.

The primer pair developed from Fig. 1 can be directly used to amplify the 762-bp sequence shown in Fig. 2 as a marker for (S_1)-cytoplasm. Moreover, the 762-bp marker fragment and the 152-bp fragment representing part of

Table 1 Primer pair and amplification products of the marker for (S_1)-cytoplasm

Sense-primer in 5´–3´	Antisense-primer in 5´–3´	Product-size
CGGCTGTCGGTATTGGAAACG	GATCAGAAAGGCCATCATTAGG	
Position 59–79 in Fig. 1 and Fig. 2	Position 189–210 in Fig. 1	152 bp in all cytoplasms
	Position 799–820 in Fig. 2	762 bp in (S_1)

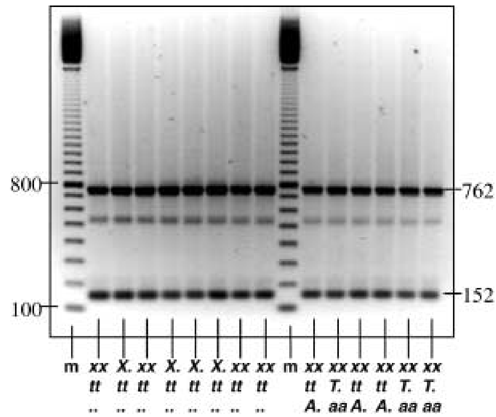

Fig. 4 PCR-marker for (S_1)-cytoplasm in progenies segregating for the restorer genes *X*/*x*, *T*/*t* and *A*/*a*; the *dot* (.) indicates the unknown allele (dominant or recessive); *m*=100-bp ladder

the *atp*9-gene (compare Fig. 1) is amplified in all cytoplasm types and might serve as a control for successful amplification during PCR (Table 1).

Discussion

Apart from the *Poaceae* (see Materials and methods) the described complete coding and 3´-flanking region of *atp*9 from *A. schoenoprasum* is the first available information of this gene from a monocotyledon. As briefly reviewed in our former paper (Engelke and Tatlioglu 2000b), the *atp*9-region might be considered as a hotspot region for recombination events not only in chives, but also in a wide range of different plant species. The *atp*9-related sequence described here confirms these previous assumptions. This sequence is of a chimerical nature, consisting of *atp*9-homologous regions on its end that are interrupted by an insertion which is composed of one *atp*6-homologous part and one part of unknown origin (Fig. 2). Since it is known that parts of the mitochondrial genome can also be repeated in the nuclear DNA, as is the case for chromosome 2 of *A. thaliana* as the most prominent example (Lin et al. 1999), we confirmed the cytoplasmic nature of our sequence by segregation studies during a backcross programme (Fig. 2). The nuclear background or a restorer gene might be capable of eliminating parts of the mitochondrial genome or reducing them to substoichiometric levels, as was described for the restorer gene *Fr* in the CMS system of *Phaseolus vulgaris* (He et al. 1995). An influence from the nuclear background, with special regard to the different restorer genes on the occurrence of the marker fragment in this study, has not been observed (Fig. 4) confirming our previously described RFLP-experiments (Engelke and Tatlioglu 2000b).

The *atp*9-related sequence can be used as a PCR-marker in order to determine the character of the cytoplasm, whether it is a CMS_1-inducing cytoplasm or a normal cytoplasm type. This might be helpful to reduce the efforts in a breeding programme when new sterile lines and the corresponding maintainer lines can be developed. If the sterility inducing cytoplasm is detected by the PCR-marker, the selected plant can never be used as a maintainer for this CMS system; however, the plant might be sterile and can be used as the starting point for a sterile line. The selection of sterile and corresponding maintainer plants from one and the same provenance might reduce the number of backcrosses for reaching homozygous male sterile lines with corresponding maintainers, which are needed in a successful hybrid breeding programme. Up to now, a differentiation of the cytoplasm types could be realised by RFLP-analyses using mitochondrial gene probes (Engelke and Tatlioglu 2000b). In order to distinguish (N)- and (S)-cytoplasms, the characterization of polymorphisms in the mitochondrial DNA by RFLPs is a significantly faster procedure than test crossing, which is normally used in plant breeding. Nevertheless it is still time-consuming to complete RFLP analysis, especially due to the fact that large amounts of highly purified DNA are necessary but not available from small plants. In contrast, the PCR-marker described here allows a quick and confident identification of the cytoplasm of individual plants, even if they have just reached a small size. Such a PCR-marker for CMS_1 cytoplasm has not yet been available in chives, despite one that was described by Dewal (2000). However, an open reading frame has not been found within this marker (Dewal 2000), in contrast to the chimerical marker sequence described here. This putative open reading frame of 501 bp will be named *orf*A501. If translated, the size of the expected protein comes very close to the previously described CMS_1-specific protein, with an apparent molecular weight of 18 kDa, that was deduced from its mobility in gel electrophoresis (Potz and Tatlioglu 1993). However, evidence of the transcription of *orf*A501 has to be provided in further investigations, and an influence from the restorer genes on the expected transcripts has to be examined.

Acknowledgements This research project was supported by the Bundesministerium für Wirtschaft und Technologie (KF 0233401KRF0).

References

Dewal G (2000) Beiträge zur molekularen Charakterisierung der unterschiedlichen Formen der männlichen Sterilität bei *Allium schoenoprasum* L.. PhD thesis, University of Hannover, Hannover, Germany

Edwards K, Johnstone C, Thompson C (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic Acids Res 19:1349

Engelke T, Tatlioglu T (2000a) Genetic analyses supported by molecular investigations provide evidence of a new genic (st1) and a new cytoplasmic male sterility (st2) in *Allium schoenoprasum* L. Theor Appl Genet 101:478–486

Engelke T, Tatlioglu T (2000b) Mitochondrial genome diversity in connection with male sterility in *Allium schoenoprasum* L. Theor Appl Genet 100:942–948

He S, Lyznik A, Mackenzie S (1995) Pollen fertility restoration by nuclear gene Fr in CMS bean: nuclear directed alteration of a mitochondrial population. Genetics 139:955–962

Lin X, Kaul S, Rounsley S, Shea TP, Benito MI, Town CD, Fujii CY, Mason T, Bowman CL, Barnstead M, Feldblyum TV, Buell CR, Ketchum KA, Lee J, Ronning CM, Koo HL, Moffat KS, Cronin LA, Shen M, Pai G, Van Aken S, Umayam L, Tallon LJ, Gill JE, Venter JC, et al. (1999) Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. Nature 402:761–768

Potz H, Tatlioglu T (1993) Molecular analysis of cytoplasmic male sterility in chives (*Allium schoenoprasum* L.). Theor Appl Genet 87:439–445

Tatlioglu T (1982) Cytoplasmic male sterility in chives (*Allium schoenoprasum* L.). Z Pflanzenzuechtg 89:251–262

Tatlioglu T (1987) Genetic control of temperature-sensitivity of cytoplasmic male sterility (cms) in chives (*Allium schoenoprasum* L.). Plant Breed 99:65–76

Tatlioglu T, Wricke G (1988) Genetic control of tetracycline sensitivity of cytoplasmic male sterility (cms) in chives (*Allium schoenoprasum* L.). Plant Breed 100:34–40

Unseld M, Marienfeld JR, Brandt P, Brennicke A (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. Nature Genet 15:57–61