Degenerated recognition property of a mitochondrial homing enzyme in the unicellular green alga *Chlamydomonas smithii*

Sayuri Kurokawa[1], Tomohito Yamasaki[1], Teruaki Komatsu[1], Kazuo I. Watanabe[2], Takeshi Ohama[1,*]

[1]Graduate School of Engineering, Department of Environmental Systems Engineering, Kochi University of Technology (KUT), Tosayamada, Kochi 782-8502, Japan

[2]Institute for Cellular and Molecular Biology, The University of Texas at Austin, Moffett Molecular Biology Bldg, A4800, 2500 Speedway, Austin, TX 78712-1095, USA

Key words: selfish genetic element, group I intron, coconversion, horizontal transmission, adapted recognition

Running title: Intron invasion assay for degenerate recognition of intronic homing enzyme using chimeric diploid cells

4920 words in total

*For correspondence (fax +81-887-57-2520; tell + 81-887-57-2512; e-mail ohama.takeshi@kochi-tech.ac.jp)

Summary

Target sequence cleavage is the essential step for intron invasion into an intronless allele. DNA cleavage at a specific site is performed by an endonuclease, termed a homing enzyme, which is encoded by an open reading frame within the intron. The recognition properties of them have only been analyzed *in vitro*, using purified, recombinant homing enzyme and various mutated DNA substrates, but it is unclear whether the homing enzyme behaves similarly *in vivo*. To answer this question, we determined the recognition properties of I-*Csm*I *in vivo*. I-*Csm*I is a homing enzyme encoded by the ORF of the alpha-group I-intron, located in the mitochondrial COB gene of the green alga *Chlamydomonas smithii*. The *in vivo* recognition properties of it were determined as the frequency of intron invasion into a mutated target site. For this purpose, we utilized hybrid diploid cells developed by crossing alpha-intron-plus *C. smithii* to intron-minus *C. reinhardtii* containing mutated target sequences. The intron invasion frequency was much higher than the expected from the *in vitro* cleavage frequency of the respective mutated substrates. Even the substrates that had very little cleavage in the *in vitro* experiment were efficiently invaded *in vivo*, and were accompanied by a large degree of coconversion. Considering the ease of the homing enzyme invading into various mutated target sequences, we propose that the principle bottleneck for lateral intron transmission is not the sequence specificity of the homing enzyme, but instead is limited by the rare occurrence of inter-specific cell fusion.

Introduction

Molecular phylogenetic analyses has demonstrated that group I introns are mobile elements that can be laterally transmitted to a wide range of organisms (Belfort and Roberts, 1997; Cho *et al*., 1998; Lambowitz, 1989; Watanabe *et al.,* 1998). They are thought to be parasitic genetic elements that have persisted for an evolutionarily long time (Goddard and Burt, 1999). Site-specific double-stranded DNA cleavage is the essential initial step for regular intron invasion. Without cleavage of the target DNA sequence, the intron cannot invade the intronless genome. DNA cleavage is performed by a sequence specific endonuclease (termed a homing enzyme) encoded in an open reading frame (ORF) found within the intron. Cleavage of the target DNA triggers a series of reactions resulting in the transmission of the intervening sequence unidirectionally from the intron-containing DNA strand to the cleaved, intron-lacking recipient DNA strand. *In vitro* analysis has shown that the homing enzyme recognizes a non-palindromic 16 to 30 base pair (bp) sequence encoding highly conserved amino acid residues (for reviews, Belfort and Perlman, 1995; Chevalier and Stoddard, 2001). On the other hand, the homing enzyme also cleaves substrates containing a silent or tolerated amino acid change more efficiently than non-silent or notolerated ones (e.g., Aagaard *et al*., 1997; Dalgaard *et al*., 1994; Kurokawa *et al*., 2005; Pellenz *et al*., 2002; Turmel *et al*., 1997; Wernette 1998;).

So far, the analysis of homing enzyme properties has been limited to *in vitro* assay systems using recombinant enzymes and artificial substrates. Unfortunately, no procedure is available to address the differences between the *in vitro* and *in vivo* behaviors of the homing enzymes. Therefore, the efficiency of homing enzyme cleavage *in vivo* remains unknown. It is also unknown why most homing enzymes are active *in vitro* around pH 9.0 (Geese *et al*., 2003; Monteilhet *et al*., 2000), which is an alkaline environment compared to the physical pH value of pH 7.5 in yeast

mitochondria (Wernette *et al*., 1990). The lack of answers to these basic questions is probably due to the difficulty in developing strains bearing the intended mutations in the target region.

The unicellular green alga *Chlamydomonas smithii* has a 1075-bp long group I intron (alpha-intron or *Cs* COB-1) located in the apocytochrome b (COB) gene of the mitochondrial genome (Colleaux *et al*., 1990). The alpha-intron contains an ORF encoding a homing enzyme, I-*Csm*I (Ma *et al*., 1992), which has the two typical LAGLIDADG motifs. Using two types of N-terminally truncated recombinant I-*Csm*I polypeptides, we previously performed systematic *in vitro* assays to determine the cleavage efficiency of various mutated target sequences (Kurokawa *et al*., 2005). Furthermore, we recently developed a practical method for mitochondrial transformation of *Chlamydomonas reinhardtii* (Randolph-Anderson *et al*., 1993; Yamasaki *et al*., 2005). Using this technique, we developed various *C. reinhardtii* strains containing different point mutations at the I-*Csm*I target site.

These strains have been very useful for probing the recognition properties of I-*Csm*I *in vivo*. The two species are interfertile and have very similar mitochondrial genome organization and DNA sequences, except that *C. reinhardtii* lacks the alpha-intron located within the COB gene. Therefore, the *in vivo* recognition properties of I-*Csm*I were investigated by analyzing the invasion of the alpha-intron into the *C. reinhardtii* strains containing the various mutated target sites within their mitochondrial genomes. The analysis is complicated because *Chlamydomonas* mitochondrial DNA is transmitted to the meiotic progeny from the paternal (mt^-) parent but not from the maternal (mt^+) parent (Boynton *et al*., 1987; Matagne *et al*., 1988). Maternal mitochondrial genomes are subjected to degradation in the course of meiosis (Beckers *et al.,* 1991) preceding the fusion of mt^+ and mt^- mitochondria (Nakamura *et al*., 2003). However, in the diploid cells, which are rare zygotes and account for about 5% of the mated cells, inheritance of the mitochondrial genome is

no longer strictly uniparental. This means that the analysis of diploid cells is essential when investigating the sequences susceptible to I-*Csm*I cleavage *in vivo*, as measured by intron invasion into the intronless *C. reinhardtii* mitochondrial genome. In the hybrid diploid cell, fusion of *C. smithii* and *C. reinhardtii* mitochondria occurs, and the diploid cell divides mitotically escaping the maturation steps of meiosis (Gillham, 1978).

Results

Creation of C. reinhardti*i strains containing single point mutations at the* I-CsmI *target site*

We prepared 33 types of DNA constructs and used them to introduce various single point mutations into the I-*Csm*I target region of the *C. reinhardtii* mitochondrial COB gene (Table 1). These mutations correspond to the six highly conserved amino acid residues at the target site, Trp-Gly-Gln-Met-Ser-Phe. We attempted to transform respiratory deficient *dum*-1 strains two or three times by biolistic-bombardment with each type of DNA construct. As a result, we obtained 14 types of respiratory active strains that contain the desired point mutation. However, we were not able to obtain the other 19 types of DNA constructs using the same approach. It became clear that our success was related to the type of mutation present in the 5.0 kb DNA construct. 17 of the 19 unsuccessful transformation constructs contained non-synonymous/non-silent mutations. On the other hand, nine of the 14 successfully transformed constructs contained silent mutations, four had synonymous ones (one Met to Leu mutation and three Met to Ile mutations at different positions), and only one construct had non-synonymous/non-silent mutations [TGG(Trp)/TGT(Cys)]. These results suggest that our inability to obtain transformants was caused by the inactive COB

peptides produced from the mutated COB gene of the 5.0 kb DNA constructs. Supporting the above results, some of the transformants containing silent mutations showed slower growth, i.e., four of the five randomly selected transformants containing GGC/GGG(Gly), CAA/CAG(Gln), CCA/CCC (Pro) or TCT/TCC(Ser) mutations showed about 20% slower growth than the wild-type, with the exception of the GGC/GGA (Gly) mutation (data not shown).

Zeocin resistance was conferred to each type of successfully transformed *C. reinhardtii* (mt^-) strain by introducing the BLE gene into the nucleus, followed by mating to *C. smithii* (mt^+, spc R) to obtain chimeric diploid cells. More than 10 hybrid diploids were developed from each possible mating combination, except for the *C. reinhardtii* strain, which contains the mutation ATG/ATC (Ile). A cross between this strain and *C. smithii* resulted in very limited mating because of unknown reasons, and we were unable to analyze the diploid cells. All together, we successfully obtained 13 types of hybrid diploid cells as a result of crosses between the intron-containing *C. smithii* and various mutated intronless *C. reinhardtii* strains.

We confirmed the diploidy of the double-antibiotic resistant cells by checking for the presence of mating type specific genes by PCR. Among 134 of 183 randomly chosen zeocin and spectinomycin resistant cells, three (2.2 %) were actually false-positive cells. The mt^- specific gene *mid1* was detected, however the mt^+ specific gene, *fus1*, could not be amplified by PCR. Therefore, these cells were highly likely haploid cytoductants (Matagne *et al*., 1991) that possess the nuclear of *C. reinhardtii* (mt^-) and the chloroplast *C. smithii*. Such pseudo-diploids were not analyzed further.

Intron invasion into the mutated target site of* C. reinhardtii *mitochondrial genome

Alpha-intron invasion into the *C. reinhardtii* target site was analyzed using a set of primers flanking the I-*Csm*I target site (Fig.1). With this primer set, the expected PCR product is 3 kb long when the target site contains the alpha-intron and 2 kb long when

the intron is missing (Fig. 2). The intron invasion frequency was calculated as a ratio of the two PCR products (see Experimental procedures for the details) and is summarized in Table 2.

The +6 base, relative to the intron insertion site, was the most critical base among the mutated sites investigated in this study: Two types of mutations positioned at +6, ATG(Met) to ATA(Ile) and ATT(Ile), resulted in a low invasion frequency, 6% and 35% respectively. Position +6 is near the cleavage point of the coding strand, located between bases +5 and +6. This result is consistent with our previous *in vitro* assay, which showed that the bases located around the coding strand cleavage point are critical for intron invasion (Kurokawa *et al*., 2005). Another striking feature of the *in vivo* assay is that all of the mutated sequences that are cleaved moderately or efficiently *in vitro* [i.e., the cleavage frequency of the substrate is more than 50 % (designated as '++' or '+++' in Table 2)] showed a much higher frequency of alpha-intron invasion (83-100%) than expected from the *in vitro* ratio. On the other hand, one apparent inconsistency between the *in vitro* and *in vivo* data concerned the mutation GGC/GGG(Gly) at position -1. This mutant target site showed only a low amount of cleavage in the *in vitro* experiment, while here it was efficiently invaded *in vivo* at a frequency as high as 92%. Intron invasion ratio was 100% when wild-type *C. smithii* and *C. reinhardtii* were crossed (Table 2).

Coconversion of genetic markers associated with intron invasion

Next, we determined whether there is a correlation between the coconversion of the target site-flanking regions and intron invasion. For this purpose, we took advantage of the presence of restriction enzyme sites for *Nhe*I, *Nco*I, and *Hpa*I (Fig. 3). The *C. smithii* mitochondrial genome contains at least one unique restriction site for each of these restriction enzymes, while they do not exist in the *C. reinhardtii* genome in the region surrounding the alpha-intron insertion site. Diagnostic *Nhe*I recognition site

(hereafter *Nhe*I*) is located about 500 bases upstream of the target site between the alpha-intron and the left-arm of the *C. smithii* mitochondrial genome, while *Nco*I* and *Hpa*I* are located downstream of the alpha-intron, about 1.4 kb and 5.0 kb from the intron insertion site, respectively (Fig. 1). For the intron-invaded mitochondrial genome, the upstream and downstream regions surrounding the alpha-intron were amplified by PCR. The upstream 1 kb PCR product was incubated with the restriction enzyme *Nhe*I, while the 5.5 kb downstream PCR product was incubated with *Nco*I or *Hpa*I (Fig. 3). Restriction enzyme analysis indicated a high frerquency of coconversion for *Nhe*I* and *N*coI*, irrespective of the type of mutation at the target site, while the coconversion of *Hpa*I*, which is the most downstream marker, showed a significantly lower frequency of coconversion (Table 2). This suggests that exonucleolytic degradation progresses bi-directionally away from the cleavage point before repair synthesis begins in the mitochondrial matrix. Interestingly, when the intron invasion frequency was low, coconversion of the farthest genetic marker *Hpa*I* was not detected [i.e., the *Hpa*I* coconversion ratio was null for ATG(Met)/ATA(Ile) (6% intron invasion frequency) and TGG(Trp)/TGT(Cys) (31% intron invasion frequency)].

Discussion

The native I-*Csm*I ORF is fused to its preceding exon and the mature N-terminus of I-*Csm*I has not yet been identified because of the difficulties in isolating this very low expression-level enzyme. Therefore, the enzymes used in the previous *in vitro* assay were recombinant polypeptides containing different N-termini [I-*Csm*I(200) and I-*Csm*I(217)] (Kurokawa *et al*., 2005). I-*Csm*I(200) is composed of 200 amino acid residues, while I-*Csm*I(217) is 217 amino acid residues. These recombinant I-*Csm*I's demonstrated that their recognition specificity is tuned to efficiently cleave the most likely encountered target sequences in nature (e.g., most of the tolerated variations

occur at the codon third positions in the target site). Therefore, it is likely that our previous *in vitro* data on recognition properties of I-*Csm*I reflect the basic characteristics of I-*Csm*I. It is also likely that apparent inconsistency in substrate cleavage and intron invasion frequencies is a result of the differences in the method of assay, not the nature of the participating homing enzyme, whether recombinant or native. The assembly of some accessory proteins with I-*Csm*I in the mitochondrial matrix might explain this inconsistency. We have no solid data to exclude this possibility, however the interpretation described below seems reasonable enough to explain the inconsistency without involving any accessory elements. Close comparison of the results obtained by *in vitro* cleavage assay and *in vivo* assay seems very informative to assess the inconsistency.

The *in vitro* cleavage assay was performed in 50 μl of a reaction mixture containing 25 mM NaCl, 5 mM $MgCl_2$, 50 mM Hepes (pH 7.0), 1 mM dithiothreitol, 0.01% (w/v) bovine serum albumin, 1.0 μg of purified recombinant I-*Csm*I(200) or I-*Csm*I(217) [this is in excess for complete digestion of some kinds of substrate], and eighty-one kinds of substrate DNA fragment (104 bp) containing single point substitutions in the core recognition region (between -12 and +15) of the enzyme (Kurokawa *et al*., 2005). After incubating the reaction at 30 ^{o}C for 6-8 h, the cleaved and uncleaved substrates were separated on an agarose gel and quantified. It is possible that substrates remained uncleaved even after 6-8 h incubation in the mixture could remain uncleaved even after an infinitely long incubation time, because I-*Csm*I may lose its cleavage activity before binding to the target region. The substrate DNA trapped by the inactive enzyme could remain uncleaved. Actually, at high concentrations of both the enzyme and substrate, we observed the accumulation of high molecular weight aggregates composed of substrate DNA and homing enzyme in our previous *in vitro* assay (Kurokawa *et al*., 2005). On the other hand, newly translated I-*Csm*I might be continuously supplied to the mitochondrial matrix, and

inactive I-*Csm*I bound to the target region might be removed from the DNA when the mitochondrial genome replicates. This scenario may explain how a target sequence cleaved at only 50% was almost entirely invaded by the alpha-intron *in vivo*. For example, substrate DNA carrying the mutations CCA/CCT(Pro), CCA/CCG(Pro), and CCA/CCC(Pro) were cleaved at a frequency of about 50% (as designated '++' in Table2) *in vitro*, but was invaded by the alpha-intron with a frequency of 83%, 100%, 91%, respectively. On the other hand, four kinds of transformants containing TGG(Trp)/TGT(Cys)/(-), ATG(Met)/TTG(Leu)/(-), ATG(Met)/ATT(Ile)/(-), or TCT(Ser)/TCG(Ser)/(-) mutations ['-' designates almost no cleavage *in vitro*], showed an intermediate intron invasion frequency *in vivo* (31-50%). It is likely that these target sequences were cleaved inefficiently even in the context of the mitochondrial matrix, because of the very limited recognition by I-*Csm*I.

Coconversion of genetic markers from *C. smithii* to *C. reinhardtii* occurs in associated with alpha-intron invasion. The extent of coconversion was often as far as 5.0 kb from the cleavage site (Table 2). This coincides with earlier results obtained for wild-type *C. reinhardtii* and *C. smithii* (Remacle and Matagne, 1993). The large region of coconversion in *Chlamydomonas* mitochondria is in contrast to the shorter region of coconversion observed for the omega-intron in *Saccharomyces cerevisiae*. In the case of the omega group I intron, found in the mitochondrial large subunit of the rRNA gene, coconversion region between omega-plus and omega-minus *S. cerevisiae* strains is limited to a few hundred bases long (Jacquier and Dujon, 1985; Muscarella and Vogt, 1989). The region of coconversion appears to depend on the size of the gap produced after cleavage of the target sequence by the homing enzyme. Therefore, the implicated exonucleases (e.g., DNA exonuclease A-like enzyme) (Huang *et al*., 1999) might be much more active in *C. reinhardtii* than in *S. cerevisiae*.

In *C. reinhardtii* chloroplast, exon coconversion associated with a group I intron (*Cr.psbA4*) homing has been reported by transformation of an intronless-*psbA* strain with integrated DNA construct containing the intron region (Odom *et al*., 2001). Moreover, the coconversion tract associated with Ce LSU·5 intron (a *C. eugametos* specific group I intron) into the *C. moewusi* chloroplast genome has analyzed in detail by reciprocal crosses between the two species (Bussieres *et al*., 1996). Conconversion of genetic markers from the intron towards the largest single-copy region is observed to a distance of at least 5 kb and 37 kb, while towards the small single-copy region coconversion is observed to a distance of 55 kb. These investigations show large region of coconversion is not limitted to *C. reinhardtii* mitochondria. If the wide range of coconversion observed in *Chlamydomonas* is common for many other organisms, then the methods used to determine the relationship of organisms based on the comparison of organellar DNA sequences may be flawed. This effect may be amplified when the organellar genome being analyzed contains many group I introns, as in fungi.

In this study, we performed crosses between *C. smithii* (mt^+, $alpha^+$, *Nhe*I*, *Nco*I*, *Hpa*I*) and *C. reinhardtii* (mt^-, $alpha^-$) to probe the frequency of intron invasion. Therefore, the apparent mt^+ type mitochondrial genome in the hybrid diploid, which contains all of the mt^+ genetic markers ($alpha^+$, *Nhe*I^*, *Nco*I^*, *Hpa*I^*), has two possibilities concerning its origin; if the *C. smithii* mitochondrial genome itself or the *C. reinhardtii* genome was subjected to a larger degree of coconversion accompanying alpha-intron invasion. Remacle and Matagne (1993) have shown that successful cleavage of mt^- mitochondrial genomes is essential in developing a diploid cell that is composed of only mt^+ type mitochondrial genomes by crossing alpha-plus mt^+-*C. smithii* with alpha-minus mt^--*C. reinhardtii*. Considering this, the 5.5 kb PCR product containing the physical markers, *Nhe*I^*, *Nco*I^*, *Hpa*I^*, and alpha-intron is likely the result of a large amount of coconversion associated with intron invasion,

but not a result of the remaining *C. smithii* mt^+ mitochondrial genome itself. Therefore, the corresponding PCR products were considered results of intron invasion in this study.

In diploids that are a result of artificial somatic cell fusion between wild-type $alpha^+$-*C. smithii* and $alpha^-$-*C. reinhardtii,* the alpha-intron is transmitted to *C. reinhardtii* as efficiently as that observed in sexually developed diploid cells (Remacle and Matagne, 1993). This demonstrates that the expression level of the homing enzyme in the somatic cell is not very different from that in the zygote. Considering the limited variation in the target sequence and the ability to recognize degenerated target sequences, the principle bottleneck for lateral intron transmission is not likely related to the presence of a cleavable target sequence in the genome. Furthermore, in this investigation we showed that inter-specific transmission of a group I intron is easier than expected from the results of the *in vitro* cleavage assay. Also, the tendency to exclude the intron-invaded mitochondrial genomes must be very weak, judging from the heteroplasmy of intron-plus and minus molecules even after two-rounds of single-colony isolation. All together, it is plausible that the bottleneck of intron invasion across species is the rare occurrence of inter-specific cell fusion. The discovery of a common group I intron sequence in both a plant-parasitic fungus and its host (Nishida and Sugiyama, 1995) supports the direct cell to cell transmission model of group I introns.

Experimental procedures

Algal strains and culture conditions

Wild type *Chlamydomonas smithii* CC-1373, mating type plus (mt^+), *Chlamydomonas reinhardtii* CC-124, mating type minus (mt^-), and the respiratory deficient strain CC-2654 *dum*-1 (mt^-) were obtained from the *Chlamydomonas* Genetic Center (Dept. of

Botany, Duke Univ. NC 27706, USA). The strains were cultured in Tris-acetate-phosphate (TAP) medium (Gormman and Levine, 1965) at 25 ^{o}C under constant cool white fluorescent light (84 μmol m^{-2} s^{-1}) with vigorous shaking, unless indicated otherwise.

Developing* C. reinhardtii *strains containing various single point mutations

The 5.0 kb DNA construct, which contains a wild-type *C. reinhardtii* COB gene and upstream region was used as starting material (Yamasaki *et al*., 2005). I-*Csm*I's target sequence in the 5.0 kb DNA construct was modified using the mega primer method (Sambrook and Russell, 2001). In total, 83 types of single point mutation-containing 5.0 kb DNA constructs were prepared and ligated into the *E. coli* vector pT7-Blue (Novagen, Milwaukee, USA). The respiratory deficient *C. reinhardtii* strain *dum*-1, which lacks the entire COB gene and the upstream region, was transformed using a mutated 5.0 kb DNA construct to compensate the deleted region as described elsewhere (Yamasaki *et al*., 2005). After biolistic bombardment of *dum*-1, the plates were kept in the dark for 3-4 weeks until respiratory-active colonies appeared. Single colonies were re-streaked, isolated, and sequenced to verify whether they contained the correct mutation. All plasmids used in this experiment are available on request.

***Isolation of hybrid diploid cells by crossing* C. smithii *and* C. reinhardtii**

We conferred spectinomycin resistance (spc R) on mt^{+}-*C. smithii* by introducing plasmid pEX-50-AAD (Takahashi *et al*., 1996) into the plastid, while zeocin (or bleomycin) resistance (ble R) was conferred to each one of mt^{-}-*C. reinhardtii* transformants by re-introducing pSP124S (Lumbreras *et al*., 1998) into the nucleus.

The mating and isolation of rare diploid cells were performed as follows: The *C. smithii* (mt^{+}, spc R) and single point mutation-carrying *C. reinhardtii* (mt^{-}, ble R) were grown at 25 ^{o}C under constant light on NH_4Cl deprived TAP agar plates for one week.

Then, the cells were suspended into gamete-inducing medium (Martin and Goodenough, 1975; Harris, 1989) for about 3 h at 25 ^{O}C under constant white fluorescent light (120 μmol m^{-2} s^{-1}). Two kinds of gametes were mixed for 30 min and allowed to mate, and spread on TAP plates containing 100 μg/ml spectinomycin and 10 μg/ml zeocin (Invitrogen, California, USA). Colonies that appeared after 5-7 days were chosen, and two rounds of single-colony isolation was performed using plates containing spectinomycin and zeocin. The diploidy of these cells was confirmed by the existence of mating type-specific genes within the cell, *fus1* (mt^+ specific) and *mid1* (mt^- specific), using the PCR method of Zamora *et al* (2004).

PCR product analysis to indicate the presence of alpha-intron in the target region

Each diploid cell was grown to late log phase in 3 ml of TAP medium. Then, total DNA was prepared using DNeasy Plant Mini Kit (Qiagen, California, USA) after disrupting the cells with glass beads. PCR was performed using a set of primers that flank the I-*Csm*I target sequence (the location of primer set A is shown in Fig. 1). With this primer set, a 3 kb PCR product is expected when the alpha-intron is inserted at the target site, and a 2 kb PCR product is expected for the intronless target site. At least 10 hybrid diploids were analyzed for each type of cross.

Coconversion frequency of the diagnostic restriction enzyme sites associated with intron invasion

To probe the range of coconversion using genetic markers, we digested the PCR product with diagnostic restriction enzymes, which exist only in the *C. smithii* mitochondrial genome. Following the confirmation of intron invasion, two types of PCR were performed to amplify the upstream and downstream regions flanking the alpha-intron.

A 1 kb PCR fragment was produced from the following set of primers: one that anneals at the 3'-terminus of the alpha-intron, and one that anneals to the left arm of the genome (primer set B in Fig.1). This PCR product contains a diagnostic *Nhe*I restriction enzyme site that is located 500 bp upstream of the I-*Csm*I target site. To generate the 5.5 kb PCR fragment, one primer hybridizes to the alpha-intron's 5' end and the other to the N-terminus of the ND2 gene (primer set C in Fig.1). This product contains diagnostic *Nco*I and *Hpa*I restriction enzyme sites, which are located 1.4 kb and 5.0 kb downstream of the forward primer site, respectively (Fig. 1). PCR products were purified using a PCR purification kit (Qiagen, California, USA) and digested with the appropriate restriction enzymes.

Calculation of the frequencies of intron invasion and coconversion of the flanking genetic markers

In this study, the intron invasion frequency is defined as the ratio of intron-containing diploid clones to the total number of examined diploid clones. A diploid clone that is heteroplasmic for the presence of the intron was counted as both an intron-containing and intronless diploid clone. The same approach was used to calculate the coconversion frequency of the flanking genetic markers when we observed similar heteroplasmy.

DNA primer sequences

Primers used for PCR-A are 5'-GTTAGTACTACCGTATTGGTGCAAG-3' and 5'-GATTCCAGAGGCCCATTATCATATGGT-3'; PCR-B are 5'-TGTAGCTAGTCTTGCCCTGGTTAC-3' and 5'-TTATGTTATTTTGGTCGAATTCTC-3'; PCR-C are 5'-GGAGCAATCCAACAGCGATAT-3' and 5'-CTAGTGGAGGCACCACGAGCCAATAGACCGCACAAT-3'. Primers used to amplyfy a part of *fus1* are 5'-GCTGGCATTCCTGTATCCTTGACGC -3' and 5'-

GCGGCGTAACATAAAGGAGGGTCG-3'; and for *mid1* are 5'-CGACGACTTGGCATCGACAGGTGG-3' and 5'-CTCGGCCAGAACCTTTCATAGGGTGG-3'.

Acknowledgements

We thank B. Engs. T. Komatsu, C. Nakagoshi and S. Minakuchi (KUT) for their technical support in making point mutation bearing DNA-constructs and Ms. Mariya Takeuchi for her encouragement to perform this study.

References

Aagaard, C., Awayez, M.J. and Garrett, R.A. 1997. Profile of the DNA recognition site of the archaeal homing endonuclease I-*Dmo*I. Nucl. Acids Res. 25: 1523-1530.

Beckers, M.C., Munaut, C., Minet, A. and Matagne, R.F. 1991. The fate of mitochondrial DNAs of mt^+ and mt^- origin in gametes and zygotes of *Chlamydomonas*. Curr. Genet. 20: 239-243.

Belfort, M. and Perlman, P.S. 1995. Mechanisms of intron mobility. J. Biol. Chem. 270: 30237-30240.

Belfort, M. and Roberts, R.J. 1997. Homing endonucleases: keeping the house in order. Nucl. Acids Res. 25: 3379-3388.

Boynton, J.E., Harris, E.H., Burkhart, B.D., Lamerson, P.M. and Gillham, N.W. 1987. Transmission of mitochondrial and chloroplast genomes in crosses of *Chlamydomonas*. Proc. Natl. Acad. Sci. USA 84: 2391-2395.

Bussieres, J., Lemieux, C., Lee, R.W. and Turmel, M. 1996. Optional elements in the chloroplast DNAs of *Chlamydomonas eugametos* and *C. moewusii*: unidirectional gene conversion and co-conversion of adjacent markers in high-viability crossses. Curr. Genet. 30:356-365.

Chevalier, B.S. and Stoddard, B.L. 2001. Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. Nucl. Acids Res. 29: 3757-3774.

Cho, Y., Qiu, Y.L., Kuhlman, P. and Palmer, J.D. 1998. Explosive invasion of plant mitochondria by a group I intron. Proc. Natl. Acad. Sci. USA 95: 14244-14249.

Colleaux, L., Michel-Wolwertz, M.R., Matagne, R.F. and Dujon, B. 1990. The apocytochrome b gene of *Chlamydomonas smithii* contains a mobile intron related to both *Saccharomyces* and *Neurospora* introns. Mol. Gen. Genet. 223: 288-296.

Dalgaard, J.Z., Garrett. R.A. and Belfort, M. 1994. Purification and characterization of two forms of I-*Dmo*I, a thermophilic site-specific endonuclease encoded by an archaeal intron. J. Biol. Chem. 269: 28885-28892.

Geese, W.J., Kwon, Y.K., Wen, X. and Waring, R.B. 2003. In vitro analysis of the relationship between endonuclease and maturase activities in the bi-functional group I intron-encoded protein, I-*Ani*I. Eur. J. Biochem. 270: 1543-1554.

Gillham, N.W. 1978. Organelle heredity. Raven Press, New York.

Goddard, M.R. and Burt, A. 1999. Recurrent invasion and extinction of a selfish gene. Proc. Natl. Acad. Sci. USA 96: 13880-13885.

Gormman, D.S. and Levine, R.P. 1965. Cytochlome F and plastocyanin: Their sequence in the photosynthetic electron transport chain of *Chlamydomonas reinhardtii*. Proc. Natl. Acad. Sci. USA 54: 1665-1669.

Harris, E.H. 1989. The *Chlamydomonas* Sourcebook. Academic Press, New York.

Harris, E.H., Burkhart, B.D., Gillham, N.W. and Boynton, J.E. 1989. Antibiotic resistance mutations in the chloroplast 16S and 23S rRNA genes of *Chlamydomonas reinhardtii*: correlation of genetic and physical maps of the chloroplast genome. Genetics 123: 281-292.

Huang, Y.J., Parker, M.M. and Belfort, M. 1999. Role of exonucleolytic degradation in group I intron homing in phage T4. Genetics 153: 1501-1512.

Jacquier, A. and Dujon, B. 1985. An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. Cell 41: 383-394.

Kurokawa, S., Bessho, Y., Higashijima, K., Shirouzu, M., Yokoyama, S., Watanabe, K.I. and Ohama, T. 2005. Adaptation of intronic homing endonuclease for successful horizontal transmission. FEBS J. 272: 2487-2496.

Lambowitz, A.M. 1989. Infectious introns. Cell 56: 323-326.

Lumbreras, V., Stevens, D.R. and Purton, S. 1998. Efficient foreign gene expression in *Chlamydomonas reinhardtii* mediated by an endogenous intron. Plant J. 14: 441-448.

Ma, D.P., King, Y.T., Kim, Y. and Luckett, W.S. Jr. 1992. The group I intron of apocytochrome b gene from *Chlamydomonas smithii* encodes a site-specific endonuclease. Plant Mol. Biol. 18: 1001-1004.

Martin, N.C. and Goodenough, U.W. 1975. Gametic differentiation in *Chlamydomonas reinhardtii*. I. Production of gametes and their fine structure. J. Cell Biol. 67: 587-605.

Matagne, R.F., Rongvaux, D. and Loppes, R. 1988. Transmission of mitochondrial DNA in cross involving diploid gametes homozygous or heterozygous for the mating-type locus in *Chlamydomonas*. Mol. Gen. Genet. 214: 257-262.

Matagne, R.F.,Remacle, C. and Dinant, M. 1991.Cytoduction in *Chlamydomonas reinhardtii*.Proc. Natl. Acad. Sci. (USA) 88: 7447-7450.

Monteilhet, C., Dziadkowiec, D., Szczepanek, T. and Lazowska, J. 2000. Purification and characterization of the DNA cleavage and recognition site of I-*Sca*I mitochondrial group I intron encoded endonuclease produced in *Escherichia coli*. Nucl. Acids Res. 28: 1245-1251.

Muscarella, D.E. and Vogt, V.M. 1989. A mobile group I intron in the nuclear rDNA of *Physarum polycephalum.* Cell 56: 443-454.

Nakamura, S., Aoyama, H. and van, Woesik, R. 2003. Strict paternal transmission of mitochondrial DNA of *Chlamydomonas* species is explained by selection against maternal nucleoids. Protoplasma 221: 205-210.

Nishida, H. and Sugiyama, J. 1995. A common group I intron between a plant parasitic fungus and its host. Mol. Biol. Evol. 12: 883-886.

Odom, O.W., Holloway, S.P., Deshpande, N.N., Lee, J. and Herrin, D.L. 2001. Mobile self-splicing group I introns from the *psb*A gene of *Chlamydomonas reinhardtii*: highly efficient homing of an exogenous intron containing its own promoter. Mol. Cell Biol. 21: 3472-3481.

Pellenz, S., Harington, A., Dujon, B., Wolf, K. and Schafer, B. 2002. Characterization of the I-*Spom* I endonuclease from fission yeast: insights into the evolution of a group I intron-encoded homing endonuclease. J. Mol. Evol. 55: 302-313.

Randolph-Anderson, B.L., Boynton, J.E., Gillham, N.W., Harris, E.H., Johnson, A.M., Dorthu, M.P. and Matagne, R.F. 1993. Further characterization of the respiratory deficient *dum-1* mutation of *Chlamydomonas reinhardtii* and its use as a recipient for mitochondrial transformation. Mol. Gen. Genet. 236: 235-244.

Remacle, C. and Matagne, R.F. 1993. Transmission, recombination and conversion of mitochondrial markers in relation to the mobility of a group I intron in *Chlamydomonas*. Curr. Genet. 23: 518-525.

Sambrook, J. and Russell, D.W. 2001. Molecular Cloning a laboratory manual Third Edition. Cold Spring Harbor: Cold Spring Hatbor Laboratory Press.

Takahashi, Y., Utsumi, K., Yamamoto, Y., Hatano, A. and Satoh, K. 1996. Genetic engineering of the processing site of D1 precursor protein of photosystem II reaction center in *Chlamydomonas reinhardtii*. Plant Cell Physiol. 37: 161-168.

Turmel, M., Otis, C., Cote, V. and Lemieux, C. 1997. Evolutionarily conserved and functionally important residues in the I-*Cue*I homing endonuclease. Nucl. Acids Res. 25: 2610-2619.

Watanabe, K.I., Ehara, M., Inagaki, Y. and Ohama, T. 1998. Distinctive origins of group I introns found in the *COXI* genes of three green algae. Gene 213: 1-7.

Wernette, C.M., Saldahna, R., Perlman, P.S. and Butow, R.A. 1990. Purification of a site-specific endonuclease, I-*Sce* II, encoded by intron 4 alpha of the mitochondrial *cox*I gene of *Saccharomyces cerevisiae*. J. Biol. Chem. 265: 18976-18982.

Wernette, C.M. 1998. Structure and activity of the mitochondrial intron-encoded endonuclease, I-*Sce*IV. Biochem. Biophys. Res. Commun. 248: 127-133.

Yamasaki, T., Kurokawa, S., Watanabe, K.I., Ikuta, K. and Ohama, T. 2005. Shared molecular characteristics of successfully transformed mitochondrial genomes in *Chlamydomonas reinhardtii*. Plant Mol. Biol. 58: 515-527.

Zamora, I., Feldman, J.L. and Marshall, W.F. 2004. PCR-based assay for mating type and diploidy in *Chlamydomonas*. Biotechniques 37: 534-536.

Table 1. List of mutations introduced into 5.0 kb-DNA construct and resulting assigned amino acid changes.

	Pro			Trp			Gly			Gln			Met			Ser		
	#-9	-8	-7	-6	-5	-4	-3	-2	-1	+1 ↓	+2	+3	+4	+5 ⋮	+6 ↓	+7	+8	+9
	C	C	A	T	G	G	G	G	C	C	A	A	A	T	G	T	C	T
A				×			×		○	×				×	○			×
				Arg			Ser		Gly	Lys				Lys	Ile			Ser
T			○		×	○		×	○		×	×	○		○			
			Pro		Leu	Cys		Val	Gly		Leu	His	Leu		Ile			
G			○	×					○	×	×	○	×	×				○
			Pro	Gly					Gly	Glu	Arg	Gln	Val	Arg				Ser
C			○		×	×	×				×			×	○*			○
			Pro		Ser	Cys	Arg				Pro			Thr	Ile			Ser

*: very limitted mating ability (see text), #: the numbering is in relation to the intron insertion site, ○: transformant was obtained, ×: transformants was not obtained. An arrow with a dotted line shows the cleavage site of the non coding strand, while an arrow with solid line denotes the cleavage site for the coding stand.

Table 2. Intron invasion ratio with each type of mutation in the I-*Csm*I's recognition site.

Position of mutation	Type of	Result of *in vitro* cleavage	Type of PCR product and intron invasion ratio			Specific genotypes and coconversion frequency			
		I-*Csm*I(200)	3kb	2kb	2&3kb	*Nhe*I*	*Nco*I*	*Hpa*I*	No. of diploid
-7	**CCA(Pro)/CCT(Pro)**	++[1]	10	2	0	-[2]	+	+	1
						-/+	-/+	-	1
						+	-	-	1
						+	+	-	1
						+	+	+	6
				83% [3]		82% [4]	82%	70%	10 [5]
-7	**CCA(Pro)/CCG(Pro)**	++	15	0	0	-	-	-	1
						-	-/+	-	1
						-	+	-	1
						-/+	-	-	1
						-/+	-/+	-	2
						-/+	+	-	1
						-/+	+	+	1
						+	-/+	-	1
						+	-/+	-/+	1
						+	+	-	1
						+	+	-/+	3
						+	+	+	1
				100%		60%	65%	32%	15
-7	**CCA(Pro)/CCC(Pro)**	++	10	1	0	-	-/+	-/+	1
						-/+	-	-	1
						-/+	-	-/+	1
						-/+	+	-/+	1
						-/+	+	+	1
						+	-	-	1
						+	-/+	-/+	1
						+	+	-/+	1
						+	+	+	2
				91%		64%	58%	53%	10
-4	**TGG(Trp)/TGT(Cys)**	-	4	9	0	-	-	-	2
						+	-/+	-	1
						+	+	-	1
				31%		50%	40%	0%	4
-1	**GGC(Gly)/GGA(Gly)**	+++	10	0	0	-	-/+	-/+	1
						-	+	+	2
						+	+	-	2
						+	+	-/+	1
						+	+	+	4
				100%		70%	91%	67%	10
-1	**GGC(Gly)/GGT(Gly)**	+++	13	0	0	-	-	-	2
						-	+	-	1
						+	-	-	3
						+	+	-	1
						+	+	-/+	3
						+	+	+	3

			100%			77%	62%	38%	13
continued									
-1	**GGC(Gly)/GGG(Gly)**	+	11	0	1	-	-	-	2
						-	-/+	-	1
						-/+	-	-	1
						-/+	-/+	-/+	1
						-/+	+	+	1
						+	-	-	3
						+	-/+	-/+	1
						+	+	+	2
				92%		60%	40%	36%	12
+3	**CAA(Gln)/CAG(Gln)**	++	9	0	1	-/+	-	-	2
						-/+	-/+	-	1
						-/+	-/+	-/+	1
						-/+	+	-/+	1
						+	-/+	-	2
						+	-	-	2
						+	+	+	1
				91%		67%	43%	25%	10
+4	**ATG(Met)/TTG(Leu)**	-	7	8	0	+	+	+	7
				47%		100%	100%	100%	
+6	**ATG(Met)/ATA(Ile)**	-	0	14	1	+	-/+	-	1
				6%		100%	50%	0%	
+6	**ATG(Met)/ATT(Ile)**	-	6	13	2	-/+	-	-	1
						-/+	-/+	-	1
						+	-	-	4
						+	-/+	-/+	1
						+	+	+	1
				35%		80%	30%	22%	8
+9	**TCT(Ser)/TCG(Ser)**	-	5	5	1	-/+	-	-	1
						+	-	-	2
						+	+	-/+	1
						+	+	+	2
				50%		86%	50%	43%	6
+9	**TCT(Ser)/TCC(Ser)**	+++	11	0	1	-/+	-/+	-	1
						-/+	-	-/+	2
						-	+	+	1
						+	-	-	1
						+	+	+	7
				92%		73%	69%	71%	12
	***C. reinhardtii* Wild** × ***C. smithii* Wild**	++	10	0	0	-	+	-	1
						-	+	+	2
						+	-/+	-	1
						+	+	-	2
						+	+	+	4
				100%		70%	91%	60%	10

(1) '+++' designates about 80% to complete digestion of the substrate DNA by *in vitro* cleavage assay (Kurokawa et al., 2005), '++' about 50%-60% cleavage, '+' about 5%-25%, '-' no digestion. (2) + indicates existence of the genetic marker, - ; absence; -/+; heteroplasmy for the genetic marker. (3) Intron invasion ratio is calculated as (no. of 3 kb + no. of 2 & 3 kb)/[no. of 3 kb + no. of 2 kb + 2x(no. of 2 & 3 kb)] (see Materials and Methods for the definition). (4) Coconversion frequency of the genetic marker is calculated as [no. of (+)-diploids + no. of (-/+)-diploids]/[no. of (+)-diploids + no. of (-)-diploids + no. of 2x(-/+)-diploids] . *Nhe*I*, *Nco*I*, *Hpa*I* show the diagnostic restriction sites in *Chlamydomonas smithii* mitochondrial genome. (5) sub total no. of the diploids.

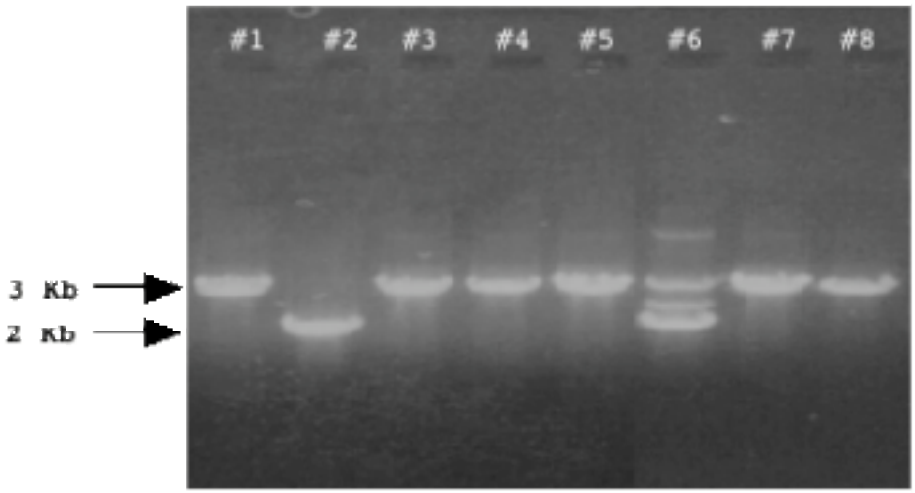

Fig. 2

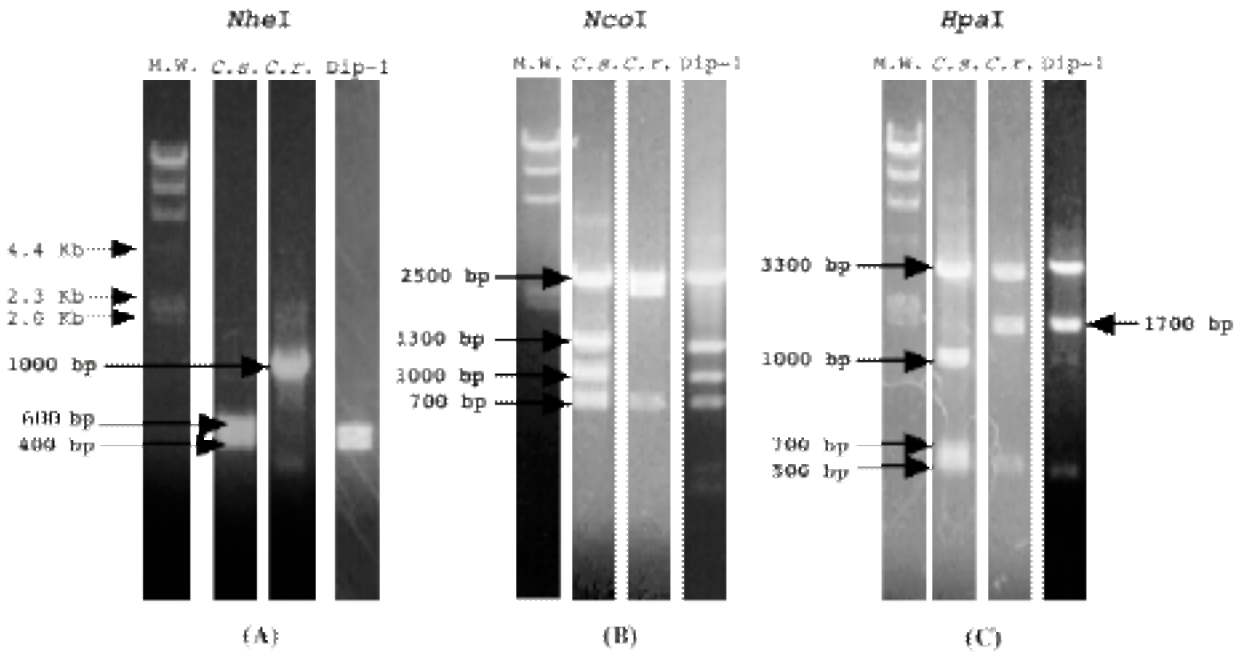
NheI
M.W. C.s. C.r. Dip-1
4.4 Kb
2.3 Kb
2.0 Kb
1000 bp
600 bp
400 bp
(A)
NcoI
M.W. C.s. C.r. Dip-1
2500 bp
1300 bp
1000 bp
700 bp
(B)
HpaI
M.W. C.s. C.r. Dip-1
3300 bp
1000 bp
700 bp
500 bp
1700 bp
(C)

Fig. 3