Isolation and molecular characterization of the *RecQsim* gene in Arabidopsis, rice (*Oryza sativa*) and rape (*Brassica napus*)

M.B. Bagherieh-Najjar\(^a\)*, S. Navabpour\(^b\), J. Hille\(^c\), P.P. Dijkwel\(^c\)

\(^a\)Department of Biology, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran
\(^b\)Department of Agronomy and Plant Breeding, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran
\(^c\)Department of Molecular Biology of Plants, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, The Netherlands.

*Corresponding author. Email: bagherieh@gau.ac.ir

Accepted 17 January 2007; Published online 24 February 2007.

**Abstract**

In any organism that reproduces sexually, DNA Recombination plays vital roles in the generation of allelic diversity as well as in preservation of genome fidelity. Genome fidelity is particularly important in plants because mutations occurring during the development of flowering plants are heritable and can be passed onto the next generation. One of the gene families that play crucial roles in the regulation of DNA recombination and repair is the *RecQ* family of DNA helicases. In flowering plants, *RecQ* family members have only been characterized in Arabidopsis. Among all *RecQ* family members, the Arabidopsis *RecQsim* is distinct in that there is a substantial insertion (of around 100 amino acids) inside its helicase domain. We previously showed that this novel *RecQ* homologue is expressed in various organs of Arabidopsis and rice. We also showed that the Arabidopsis *RecQsim* gene when ectopically expressed in yeast *RecQ* deficient (*sgs1*) cells, can partially compensate for the absence of *SGS1*. Here, we perform an evolutionary analysis of *RecQsim* orthologues isolated from rice and rape together with other known plant *RecQ* family members. Furthermore, isolation and molecular characterization of two Arabidopsis *recQsim* knockout mutants is explained. The position of T-DNA integration suggests that the *recQsim-1* is likely to be a real knockout while *recQsim-2* is expected to be a knockdown mutant. Segregation analysis of the T-DNA selectable marker together with Southern hybridization revealed that in both isolated mutant lines a single copy of the T-DNA is inserted into the genome. Analysis of these mutant lines will provide evidence on the roles that *RecQsim* plays in DNA recombination and the regulation of leaf senescence in plants and may open new insights into how plants respond to various environmental challenges.

**Keywords:** Arabidopsis; Brassica; DNA recombination; *RecQ*; Rice.
Introduction

DNA Recombination is one of the most important biological processes in any organism that reproduces sexually. Homologous recombination is important both in the generation of allelic diversity and in preservation of genetic information (Walbot, 1996; Lucht et al., 2002). In addition to reshuffling genetic information to pass on through reproduction, recombination has other highly important roles in the cell. Not least amongst these is the role which recombination plays in DNA repair (Puchta and Hohn, 1996). Genome fidelity is particularly important in plants because of the late differentiation of their germ cells. This means that mutations occurring during the development of the plant are heritable, and can be passed onto the germ line (Walbot, 1996). Possibly in an evolutionary response to this, plant cells are highly resistant to DNA-damaging treatments, and have a tendency to induce higher levels of recombination than do animal cells (Puchta et al., 1995; Walbot, 1996).

One of the gene families with an important role in DNA repair, which might be activated in response to DNA damaging agents, is RecQ helicases (Harmon et al., 1998; Kawabe et al., 2000). Genes belonging to this family have been identified in most model organisms including *Escherichia coli* (Harmon et al., 1998), *Saccharomyces cerevisiae* (Gangloff et al., 1994), *Drosophila* (Kusano et al., 1999), *Neurospora crassa* (Cogoni et al., 1999) and *Homo sapiens*. While in the genomes of *E. coli* and *S. cerevisiae* only one RecQ homologue is present, this number rises to five different RecQ homologues in humans (Ellis et al. 1995; Puranam et al., 1994; Seki et al., 1994; Gray et al., 1997; Kitao et al., 1998), and recently, seven homologues have been identified in the model plant *Arabidopsis thaliana* (Hartung et al., 2000; Bagherieh-Najjar et al., 2003).

In humans, mutations in the RecQ genes have aroused keen interest since they have been shown to be the underlying causes of three clinical disorders. Bloom’s syndrome, Werner syndrome and Rothmund-Thomson syndrome are associated with a mutation in the genes BLM, WRN and RECQ4, respectively. Mutations in all of these genes lead to an increase in the incidence of cancer, high rates of DNA recombination and in a majority of cases early aging (reviewed in Bachrati and Hickson, 2003).

All RecQ family members studied share a central helicase domain of around 400 amino acids which contains seven conserved motifs including two ATP binding Walker A and B boxes (Karow et al., 2000). To date, the biochemical properties of five members of the RecQ helicase family isolated from various organisms (BLM, WRN, RecQ, Sgs1 and RECQL) have been studied (Karow et al., 2000). These homologues have been shown to unwind DNA helices in a 3’ to 5’ direction, using an ATP-dependent reaction. Additionally, RecQ has been shown to act as a repressor of illegitimate recombination in *E. coli* (Hanada et al., 1998).

In flowering plants, RecQ family members have only been characterized in Arabidopsis (Hartung et al., 2000; Bagherieh-Najjar et al., 2003, 2005). Among all RecQ family members, the Arabidopsis RecQsim is distinct in that there is a substantial insertion (of around 100 amino acids) inside the helicase domain. Because of this, the gene was postulated to be non-functional, and as such was named RecQsim rather than RecQl6 (Hartung et al., 2003). We previously showed that this novel homologue is expressed at detectable levels in all organs of the plant and is also present in rape and rice, indicating that it may perform a plant specific function. In addition, we showed that the Arabidopsis
RecQsim gene when ectopically expressed in yeast RecQ deficient (sgs1) cells, can partially compensate for the absence of SGS1 suggesting that RecQsim can function in the regulation of DNA recombination (Bagherieh-Najjar et al., 2003). It has been widely reported that DNA recombination in plants is induced by various stress conditions (Lebel et al., 1993; Puchta et al., 1995), thus the RecQsim gene was postulated to be involved in plant responses to stress.

Here we describe the isolation of RecQsim orthologues from rice and rape and perform a phylogenetic analysis of all known plant RecQ family members. Furthermore, isolation of two Arabidopsis recQsim knockout mutants is explained. Analysis of these mutant lines will provide evidence on the roles that RecQsim plays in DNA recombination and the regulation of leaf senescence in plants and may open new insights into how plants respond to various environmental challenges.

Materials and Methods

Plant growth conditions

All experiments were performed using Arabidopsis thaliana ecotype Wassilewskija (WS). Seeds were surface sterilized in 20% bleach /80% EtOH for 15 minutes, then washed twice with 96% EtOH and 4 times with sterile water. Plants were grown on GM medium (Murashige & Skoog, MS, salts supplemented with 1% sucrose). For T-DNA selection kanamycin was added at a concentration of 50 µg/ml. To synchronize the germination all seeds were incubated at 4 °C for 48 hours, then transferred to a climate room at 22°C with 16h light (60 µmol m⁻² s⁻¹)/ 8h dark cycles, with 65% relative humidity. Alternatively seeds were sown directly on soil, germinated and grown as above.

Molecular cloning and sequencing of RecQsim orthologues

Total RNA from Arabidopsis was extracted from 5-week-old flowering plants (using tissue from the whole plant), as described by Sambrook and Russell (2001). Rice total RNA was generously provided by A. Meijer (Ouwerkerk et al., 2001). Total RNA from Brassica napus was isolated with TRIzol reagent (BRL Life Technologies) as described previously (Bagherieh-Najjar et al., 2003).

RT-PCR was performed using M-MLV Reverse Transcriptase (Promega), following the supplier’s instructions. 3’ RACE and 5’ RACE were performed using a 5’/3’ RACE kit (Roche), following supplier’s protocols. For the amplification of internal Arabidopsis and rice RecQsim cDNAs, the PCR reactions (50 µl) contained 0.2 mM of each dNTPs, 0.2 µM of each primer, 4 µl of the 4 × diluted RT reactions, 1× Taq polymerase buffer, and 1 unit of Taq DNA polymerase (Roche Diagnostics). PCR products were cloned into pGEMT-EASY (Promega), introduced into DH5α E.coli competent cells by heat-shock transformation, and were sequenced. The full length cDNA of Brassica napus was obtained as described previously (Bagherieh-Najjar et al., 2003).

Alignment: The protein alignment was conducted by using the CloneManager program ver.6 set to scoring matrix of PAM250 and the program preset parameters for Multi-Way alignment type.
Identification of recQsim Knockout mutants

The knockout mutants were obtained from the Arabidopsis Knockout Facility at the University of Wisconsin (Krysan et al., 1999). The appropriate seed pools were located using a combination of either the following RecQsim gene specific primers: P284-5'TAG CGG ATA ATT CGG AAA ATT AAG GAA AC; P285-5'AAT TCG ATG GATT TGT CTT CTG ATC AAC T together with the T-DNA border primer: JL-202-5'CAT TTT ATA ATA ACG CTG CGG ACA TCT AC, as described at http://www.biotech.wisc.edu/arabidopsis/default.htm.

DNA extraction and Southern blot analysis

DNA was extracted from leaf tissue collected from 7-day-old seedlings as described by Dellaporta et al. (1983). For Southern blot analysis, two μg of DNA was digested with 20 units of EcoRI in a final volume of 50 μl for 16h. Digested DNA samples were directly electrophoresed in a 0.8% agarose/TAE gel over night. The gels were blotted onto a positively charged nylon membrane (Hybond N+, Amersham Pharmacia), as described (Sambrook and Russell, 2001). The membrane was hybridized in phosphate-SDS buffer at 60 °C and washed in phosphate-SDS washing solutions at 58°C, as described by Sambrook and Russell, 2001. A 1.7 kb PCR fragment amplified from the GUS gene of the T-DNA present in the recQsim mutants using primers (5'-CGTCCTGTAGAAACCCCAACC) and (5'-AGGTCGCAAAATC GGCGAAA) was labeled by [α-32p]ATP and used as the probe in Southern analysis.

Results

RecQsim homologues exist in O.sativa and B.napus

The rapidly growing sequence databases for plant species such as A.thaliana and O.sativa provide a powerful tool for the identification of members of a protein family or particular homologues. Database searching using the TBLASTN algorithm with the sequence corresponding to both the conserved helicase domain and the insertion contained in the Arabidopsis RecQsim gene identified new putative homologues in O.sativa (rice) and B.napus (rape). Total RNA from rice was obtained and starting from the most conserved region, partial cDNAs were amplified by RT-PCR, followed by 5’ and 3’ RACE. These partial cDNAs were cloned into pGEMT-EASY vector and sequenced. The sequence of the whole cDNA was deduced from these cDNA fragments. The exon-intron composition of the O.sativa gene was then determined by comparison of the cDNA sequence to the genomic BAC clones (Accession number AY180331). By screening a cDNA library, the whole sequence of the B.napus cDNA was obtained (Accession number AY180332). Interestingly, the unusual insert inside the helicase domain was present in all three RecQsim proteins (Figure 1A). Within the insert, the deduced amino acid sequences of the Arabidopsis and rape RecQsim proteins are 69% identical while the insertion of either proteins has 27% identity with that of rice RecQsim. The deduced amino acid sequences of the whole three RecQsim orthologues and other RecQ-like proteins from Arabidopsis were
used in a phylogenetic analysis, as described in materials and methods, and the results are shown in Figure 1B. According to this analysis, the RecQ15 and RecQ13 proteins of Arabidopsis are the most divergent members of the group while Arabidopsis RecQ14A and RecQ14B are very similar. As expected, the RecQsim orthologues belong to a subgroup in which the proteins from Arabidopsis and rape are very close. These data indicate that RecQsim exists in mono- and di-cotyledons and most likely is a functional gene with plant specific functions.

Figure 1. (A) Amino acid sequence alignment of the insert inside the helicase domain of the RecQsim proteins from Arabidopsis (AtRecQsim), rape (BnRecQsim) and rice (OsRecQsim) using ClustalW. (B) A phylogram of known plant RecQ proteins generated by using the Phylib method. Branch distances correspond to sequence divergence and the length of each pair of branches represents the distance between sequence pairs. At, Arabidopsis thaliana; Bn, Brassica napus; Os, Oryza sativa.
To study the roles of RecQsim in plant growth and metabolism, we aimed to isolate T-DNA knockout mutants of the gene. Since in plants T-DNA integration occurs mainly randomly via nonhomologous end joining, as of today it is not easily possible to target a T-DNA specifically into the gene of interest (reviewed in Ray and Langer, 2002). Therefore, pools of randomly transferred knockout plant lines have been made, by screening of which, one can isolate plants that carry a T-DNA in a specific gene. One of these pools is available at The Arabidopsis knockout facility at the University of Wisconsin (Krysan et al., 1999). Two gene specific primers corresponding to both ends of the RecQsim gene were employed in PCR reactions in combination with a T-DNA border primer, to screen the Wisconsin T-DNA insertional pools. In the first round of PCR 68 pools were examined and many putative lines were found (Figure 2). To confirm which of the amplified fragments correspond to the RecQsim gene Southern blot analysis was performed using a RecQsim internal fragment as the probe. Consequently, five positive bands were found. This indicates that among many amplified fragments in the first round of PCR only five fragments were RecQsim specific. Among these five seed pools two were selected for further experiments (Figure 3). Finally, after several subsequent PCR reactions two independent mutant lines were identified. The amplified fragments corresponding to these two T-DNA lines were cloned and sequenced and revealed the insertion position of the T-DNA (Table 1). In the first knockout line (recQsim-1) the T-DNA is inserted at position 748 bp from the putative start codon in the cDNA sequence, inside the helicase domain. As the total RecQsim cDNA is 2574 base pairs long, the insertion leads to the production of a truncated protein of 249 amino acids, instead of the expected 858 amino acids, which is unlikely to be functional. On the other hand, the insertion in the recQsim-2 line is located at the end of the gene outside of the helicase domain. Although amino acids corresponding to the last exon will be removed in the recQsim-2 protein, it might still have some helicase activity.

Figure 2. Gel electrophoresis of the products amplified in the first round of PCR. M, Molecular marker.
Figure 3. Southern blot analysis performed on the gel shown in figure 2. RecQsim specific bands are indicated by circles. Pools indicated with arrows were chosen for further analysis.

Table 1. Molecular characterization of the recQsim mutant lines.

<table>
<thead>
<tr>
<th>RecQsim1 (5/5)</th>
<th>RecQsim2 (5/2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAC clone accession</td>
<td>AC007478</td>
</tr>
<tr>
<td>Gene size</td>
<td>4893</td>
</tr>
<tr>
<td>Position of insertion in BAC clone</td>
<td>107109</td>
</tr>
<tr>
<td>PCR fragment size (T-DNA+gene)</td>
<td>973 (151+882)</td>
</tr>
<tr>
<td>Predicted size of encoded protein</td>
<td>241</td>
</tr>
</tbody>
</table>

Both mutant lines were self-fertilized and DNA from their progeny was isolated and used in PCR reactions to identify homozygous, hemizygous or wild type siblings. A pair of one gene specific primer and a T-DNA border primer was used to distinguish between wild type plants from mutants. Concurrently, two gene specific primers that span the insertion site of the T-DNA were used to identify hemizygous and homozygous plants (Figure 4A). In this PCR, amplification of the wild type gene is indicative of the presence of at least one copy of the RecQsim wild type gene. Consequently, homozygous mutant plants were identified for each line (Figure 4B).

Prior to phenotypic analysis of the mutants, it is extremely important to identify the number of T-DNAs integrated into each mutant line. The T-DNA that has been used to knockout the RecQsim gene contains a kanamycin resistance gene that can be employed to check if the T-DNA in each line was unique. Homozygous recQsim mutant lines were crossed to WS wild type plants and the resulting F1 progenies were allowed to self fertilize. The F2 segregating populations were grown on GM plates containing kanamycin and the
ratio of resistant to sensitive plants was calculated (Figure 5). For the recQsim-1 mutant, 354 kanamycin resistant plants and 138 sensitive plants were counted which corresponds to a 3:1 ratio ($\chi^2_{\text{calc}} = 0.016, \chi^2_{(0.05, df=1)} = 3.84$). Similarly, for the recQsim-2 line, 407 kanamycin resistant plants and 170 sensitive plants were found (3:1, $\chi^2_{\text{calc}} = 2.96, \chi^2_{(0.05, df=1)} = 3.84$). These indicate that in both mutant lines there is only one locus with kanamycin resistance gene; however, a possible integration of a truncated T-DNA without kanamycin gene cannot yet be excluded. To confirm the numbers of the T-DNAs present in each line, Southern blot analyses were conducted using an internal fragment of the T-DNA as the probe. The results presented in Figure 6 demonstrate that both mutant lines contain a single locus of T-DNA integration.

The recQsim knockout mutation is not lethal and when grown under normal conditions, (either on agar or soil), both homozygous recQsim alleles were fertile and exhibit no obvious phenotypic depletion.
Figure 5. Appearance of plants grown on plates containing kanamycin. Pictures were taken 5 days after germination. Seedling of the self-pollination of wild-type (A), homozygous mutant (B) and hemizygous (C) are shown.

Figure 6. DNA gel blot of the wild type Wassilewskija (lane 1), recQsim-1 (lane 2) and recQsim-2 (lane 3) plants, using a 1.7 kb GUS fragment of the T-DNA as the probe.

Discussion

One important feature distinguishing plants from other multicellular organisms is that plants are sessile and thus have to endure environmental challenges. Although various stress conditions elicit specific plant responses, they also activate some common reactions such as DNA damage response. In addition, DNA assaults are easily heritable in plants, due to the late development of their germ line cells. Therefore, it seems logical that plants often have more family members for genes devoted to DNA repair and recombination (Arabidopsis Genome Initiative, 2000).

Here we showed that the RecQsim gene is present in Arabidopsis, rape and rice and comprise an evolutionary subgroup, suggesting that RecQsim is a plant specific member of the family. This plant specificity proposes that the gene has a specialized role in plants with a function adapted to the conditions that plants face. On the other hand, we previously showed that the Arabidopsis RecQsim gene can partially complement the hypersensitivity
of the yeast sgs1 mutant cells implying that the gene may perform some helicase activity similar to that of SGS1 (Bagherieh-Najjar et al., 2003). It is possible that additional domains present in the RecQsim protein and especially in its insert grant some plant specific functions for the protein. In support of this notion, the expression profile of the RecQsim gene in various organs of the plant and in response to environmental stimuli is different from that of other Arabidopsis RecQ genes (Bagherieh-Najjar et al., 2003).

The *O.sativa*, *B.napus* and *A.thaliana* RecQsim proteins seem to have a high level of conservation. This is particularly true for the intron-exon boundaries, including those in the inserted domain. Interestingly the sequence of the insert shows no homology to any other part of the Arabidopsis genome, or to genes of other organisms. Although there seems to be some conserved amino acids inside the insert, the lack of an obvious functional domain in the insert makes it hard to assess the role that the insert plays.

Despite the similarity, there are probably significant differences between the *A.thaliana* and *O.sativa* homologues. The *A.thaliana* RecQsim has a putative nuclear localization signal, which the *O.sativa* homologue does not have. This could potentially mean that the rice homologue is targeted to an alternative location in the cell, or that the *A.thaliana* nuclear localization signal is not functional. The latter, seems unlikely given the function of the RecQ family members in DNA recombination. It is also possible that the RecQsim is targeted to the chloroplast to perform DNA related functions. This scenario is however very speculative specially that no chloroplast targeting signal has been identified in the deduced protein.

We showed that the *recQsim* mutant plants do not show any phenotypic depletion, when they are grown under standard growth conditions. It is likely that in the *recQsim* knockout mutants, expression of the other RecQ family members is increased, leading to some redundancy of function, masking the full effect of the knockout. It would thus be interesting to look at the expression profiles of the other RecQ homologues in *recQsim* mutant plants as well as to silence all RecQ homologues in Arabidopsis. Surprisingly, preliminary data suggest that homozygous *recQsim* mutant lines grow faster than the wild type plants when grown under standard sterile conditions. Future detailed quantitative experiments will hopefully uncover the role that RecQsim protein plays in plants growth and development under various environmental conditions.

**Acknowledgements**

We are grateful to Annemarie Meijer and Pieter Ouwerkerk for providing the rice RNA samples as well as to Onno de Vries and Christina Hammond for excellent technical assistance.

**References**


