# The RecQ gene family in plants


## Frank Hartung*, Holger Puchta

*Botanisches Institut II, Universität Karlsruhe (TH), Kaiserstr. 12, 76128 Karlsruhe, Germany*



**KEYWORDS**
DNA repair;
Holliday junction;
Recombination;
RecQ helicases



## Summary

RecQ helicases are conserved throughout all kingdoms of life regarding their overall structure and function. They are 3′ 5′ DNA helicases resolving different recombinogenic DNA structures. The RecQ helicases are key factors in a number of DNA repair and recombination pathways involved in the maintenance of genome integrity. In eukaryotes the number of RecQ genes and the structure of RecQ proteins vary strongly between organisms. Therefore, they have been named RecQ-like genes. Knockouts of several RecQ-like genes cause severe diseases in animals or harmful cellular phenotypes in yeast. Until now the largest number of RecQ-like genes per organism has been found in plants. Arabidopsis and rice possess seven different RecQ-like genes each. In the almost completely sequenced genome of the moss *Physcomitrella patens* at least five RecQ-like genes are present. One of the major present and future research aims is to define putative plant-specific functions and to assign their roles in DNA repair and recombination pathways in relation to RecQ genes from other eukaryotes. Regarding their intron positions, the structures of six RecQ-like genes of dicots and monocots are virtually identical indicating a conservation over a time scale of 150 million years. In contrast to other eukaryotes one gene (RecQsim) exists exclusively in plants. It possesses an interrupted helicase domain but nevertheless seems to have maintained the RecQ function. Owing to a recent gene duplication besides the AtRecQl4A gene an additional RecQ-like gene (AtRecQl4B) exists in the *Brassicaceae* only. Genetic studies indicate that a AtRecQl4A knockout results in sensitivity to mutagens as well as an hyper-recombination phenotype. Since AtRecQl4B was still present, both genes must have non-redundant roles. Analysis of plant RecQ-like genes will not only increase the knowledge on DNA repair and recombination, but also on the evolution and radiation


*Abbreviations:* CO, crossover; HRDC, homologous region RNase D C-terminal; RQC, RecQ C-terminal; SCE, sister chromatid exchange


*Corresponding author. Tel.: +49 721 6084875; fax: +49 721 6084874.
*E-mail address:* frank.hartung@bio.uka.de (F. Hartung).

## Introduction

RecQ helicases are proteins involved in the maintenance of genome stability of virtually all organism like bacteria, yeast, animals and last but not least plants. In *Escherichia coli* the RecQ protein consists of 610 amino acids and was originally described as a suppressor of illegitimate recombination (Nakayama et al., 1985; Umezu et al., 1990; Hanada et al., 1997). In general, the protein size and the complexity of the different RecQ homologues increased from bacteria to higher organisms during evolution (Ellis et al., 1995; Cogoni and Macino 1999; Hartung et al., 2000). The largest known RecQ homologue was recently discovered in *Schizosaccharomyces pombe* and has more than 2000 amino acids (Mandell et al., 2005). Moreover, gene duplications must have occurred several times during evolution of higher eukaryotes because they contain up to seven different RecQ-like genes (Hartung et al., 2000; Hickson 2003; Hartung and Puchta, 2004). Ample work has been spent to elucidate the role of the respective homologues in different recombination pathways. We will give here a short summary of the RecQ functions described so far in eukaryotic organisms and then focus on what is known about plant RecQ-like genes. A main aspect of this review will be a summary of the structural analysis of important protein domains of the RecQ-like helicases.

## RecQ homologues in eukaryotes

In yeast, a knockout line of SGS1, the single RecQ homologue of *Saccharomyces cerevisiae*, is viable but it shows increased rates of mitotic and meiotic recombination, chromosomal rearrangements and chromosome loss (Watt et al., 1995; Sinclair and Guarente, 1997). The SGS1 mutant is hypersensitive to methylmethane sulfonate (MMS), hydroxyurea and UV light (Watt et al., 1996; Yamagata et al., 1998; Frei and Gasser, 2000). The SGS1 mutation was originally isolated as a suppressor mutation of the slow growth phenotype of a Topoisomerase 3 mutant, (abbreviation of Slow Growth Suppression). As it has been shown in vitro and in vivo, SGS1 and Topoisomerase 3 build up a tight protein complex which stabilizes and moves (by branch migration) stalled or collapsed replication forks, thus suppressing recombinational crossover (CO) events (see overview in Wu and Hickson 2003). The mechanism by which RecQ homologues and Top3 act to prevent CO seems to consist of the migration and final dissolution of DNA intermediates like Holliday junctions before CO events can take place (Karow et al., 2000; Wu and Hickson, 2003).

In animals, as in yeasts, defects in most of the RecQ homologues are leading to an increase in recombination and genome instability. This might explain the predisposition for cancer that has been found in animals. In detail, each of the different RecQ homologues of higher eukaryotes seems to have a special function because the mutation of one given RecQ gene cannot be complemented by any other of the RecQ genes. Three severe but different diseases are known to originate from mutations in human RecQ homologues: the Werner- Blooms- and Rothmund Thomson-Syndrome, caused by mutations in the WRN, BLM and RecQ4 gene, respectively (Ellis et al., 1995; Yu et al., 1996; Kitao et al., 1999; Shen and Loeb, 2000). Furthermore, mutations in a fourth RecQ-like gene, RecQ5 in mammalians causes increased sister chromatid exchange (SCE), similar to the BLM mutation but in a synergistic manner (Wang et al., 2003). This points to a different, BLM independent pathway of SCE in mammals (Hu et al., 2005). In several recent publications, a strand annealing activity of the BLM protein and the long isoform of the human RecQ5 protein has been described (Garcia et al., 2004; Machwe et al., 2005; Cheok et al., 2005). Strand annealing is an activity counterbalancing the standard DNA unwinding activity of helicases. The presence of both activities in the same RecQ protein seems to be counterproductive at first glance. Cheok et al. speculate that the strand annealing activity might help to migrate double Holliday junctions in an atypical manner to reach a hemicatenane which is then dissolved (Cheok et al., 2005).

The RecQ-like proteins have a conserved structural organization which is described in detail later (see Fig. 1). However, many of them were enlarged during evolution. *S. cerevisiae* is the only known eukaryotic organism so far which contains only one RecQ homologue, SGS1, whereas the fungi *Schizosaccharomyces pombe* and *Neurospora crassa*, possess two homologues (Cogoni and Macino, 1999; Mandell et al., 2005). The second RecQ homologue of *S. pombe* is encoded in a subtelomeric area which is also the case in the fungi *Metarhizium anisopliae, Aspergillus fumigatus* as well as *Emericella nidulans* (Hofmann and Harris, 2001; Mandell et al., 2005; Inglis et al., 2005). This is remarkable as a number of studies showed that RecQ-like proteins play a pivotal role in the context of telomere stability (Huang et al., 2001; Johnson et al., 2001; Schawalder et al., 2003; Lillard-

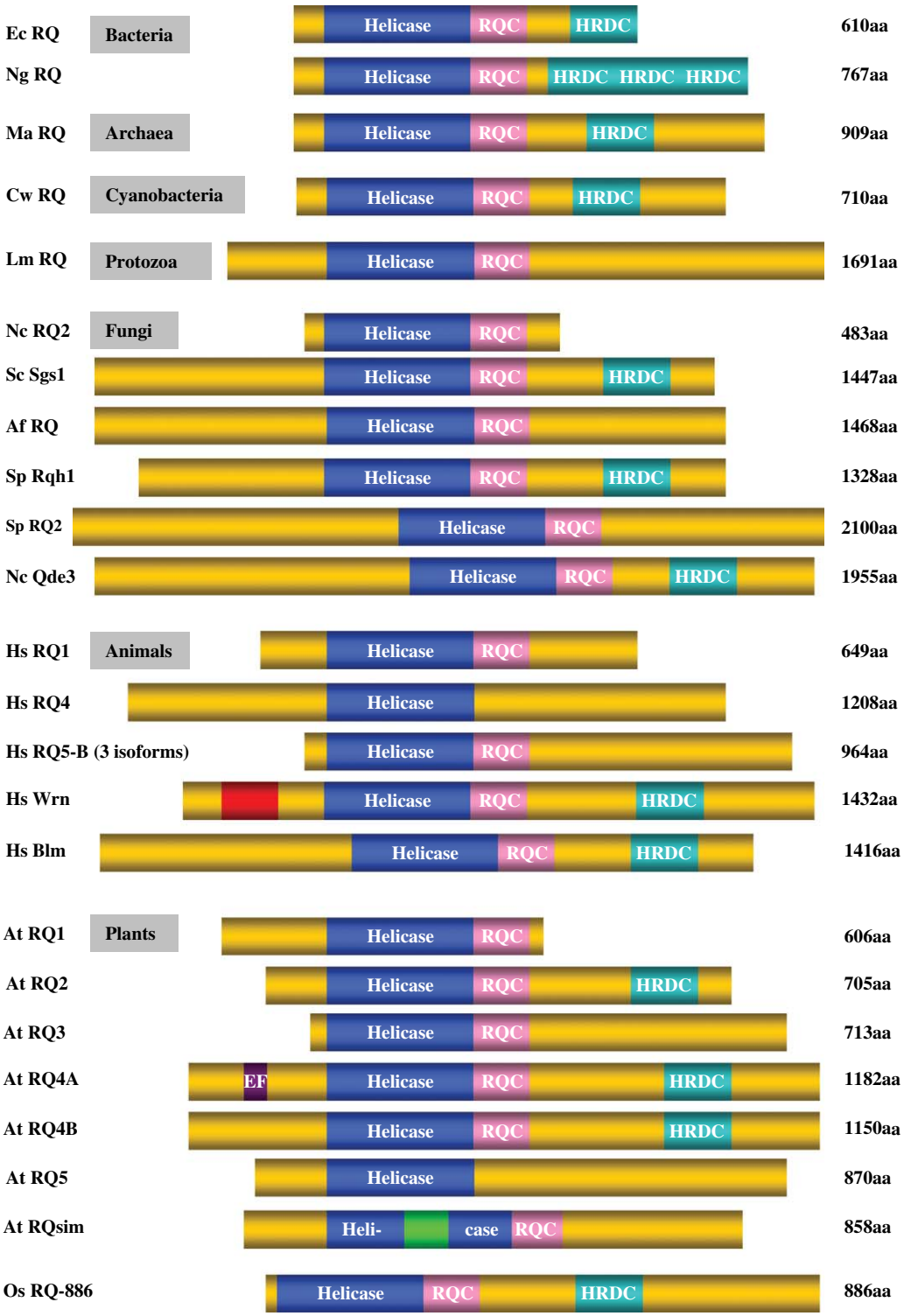


**Figure 1.** Selected RecQ-like proteins from organisms of all kingdoms of life. All proteins are aligned according to their conserved helicase domain. The location of the RQC and HRDC domain is depicted for every protein which contains them. The respective organism is shown on the left, the protein size on the right side. The exonuclease domain of the human WRN protein is shown as a dark red box, the EF hand signature of AtRQ4A as a light purple box and the insertion of 100 amino acids in the RQsim gene as a green box. Abbreviations are: Af *Aspergillus fumigatus*; At *Arabidopsis thaliana*; Cw *Crocosphaera watsonii*; Hs *Homo sapiens*; Lm *Leishmania major*; Nc *Neurospora crassa*; Ng *Neisseria gonorrhoe*, Ma *Methanosarcina acetivorans*; Os *Oryza sativa*; Sc *Saccharomyces cerevisiae*; Sp *Schizosaccharomyces pombe*.

Wetherell et al., 2005; Opresko et al., 2005; review in Bennett and Keck, 2004).

Higher multicellular eukaryotic organisms like animals and plants all possess small families of RecQ-like genes. These gene families contain four members in *Caenorhabditis elegans*, five in *Drosophila melanogaster* and *Homo sapiens* and seven members in *Arabidopsis thaliana* and *Oryza sativa* (Hartung and Puchta, 2004). Additionally, RecQ5 in human and Drosophila exists in three alternatively spliced forms (Sekelsky et al., 1999; Shimamoto et al., 2000). These splice products differ remarkably in length and also seem to fulfil different functions as demonstrated both in vitro and in vivo (Sekelsky et al., 1999; Jeong et al., 2000; Shimamoto et al., 2000; Özsoy et al., 2001; Machwe et al., 2005).

Analysing the biological functions of the RecQ-like proteins one has not only to take into account a larger number of different homologues in eukaryotes but also an enormous number of interaction partners. In various screens for genes, which are synthetically lethal together with SGS1 and/or TOP3 in yeast alone, over 30 different genes were identified which exhibited either a slow growth or a lethal phenotype (Mullen et al., 2001; Ooi et al., 2003). This clearly points to multiple roles of the RecQ homologue SGS1 in distinct biological processes. Among the identified factors are genes which are involved in homologous recombination pathways like MUS81, MMS4 and SRS2 (Gangloff et al., 2000; Kaliraman et al., 2001; Zhang et al., 2005). Very recently, a genetic and direct interaction between the SGS1/TOP3 complex and a new protein named RMI (for RecQ-mediated instability) has been shown (Chang et al., 2005; Mullen et al., 2005). This protein is a structure-specific DNA binding protein which seems to direct the SGS1/TOP3 complex to cruciformous DNA structures like Holliday junctions or replication forks (Mullen et al., 2005). The synthetic lethality of SGS1 and SRS2, another DNA helicase, led to the proposal of the existence of an additional pathway of recombinational repair in *S. cerevisiae* (Gangloff et al., 2000; Maftahi et al., 2002). Interestingly, an SRS2 homologue can be found in Arabidopsis, but not in animals, indicating that the same pathway might operate in plants (Hartung and Puchta, unpublished).

## RecQ homologues in plants – functional analysis

Plants possess the highest number of RecQ homologues in all investigated organisms (Hartung et al., 2000; Bagherieh-Najjar et al., 2003; Hartung and Puchta, 2004). So far, three functional studies of RecQ-like proteins in plants were performed. The first one addressed a homologue named AtRecQsim, which is unusual due to its interrupted helicase domain (Hartung et al., 2000; Bagherieh-Najjar et al., 2003). The protein is able to partially complement the MMS sensitivity of yeast SGS1 mutants, which means that at least this function is conserved (Bagherieh-Najjar et al., 2003). The second study dealt with a small protein named AtWRNexo, which is homologous to the N-terminal part of the human WRN protein (Plchova et al., 2003). The AtWRNexo protein exhibited similar in vitro properties as its human counterpart, the WRN protein, regarding its exonuclease activity on different DNA substrates but also differences were found. Unlike in the human WRN protein, the requirement for $Mg^{2+}$, can be replaced by $Mn^{2+}$ in AtWRNexo and the protein is able to process DNA with 3′-protruding ends (Plchova et al., 2003). The most recent study concentrated on the RecQl4A homologue of Arabidopsis and is the first in planta study on a RecQ homologue (Bagherieh-Najjar et al., 2005). The authors tested all seven RecQ-like proteins of Arabidopsis for their ability to suppress either the MMS hypersensitivity or the hyperrecombination of the SGS1 mutant in yeast. Only one of the homologues could suppress both phenotypes, the AtRecQl4A protein (Bagherieh-Najjar et al., 2003; Bagherieh-Najjar et al., 2005). This finding was very surprising because AtRecQl4A is only one member of a recently duplicated gene pair. The other member, AtRecQl4B shows a 70% identity on amino acid and nucleotide level and could so far only be found in the *Brassicaceae* but not in rice or moss (Hartung et al., 2000; Hartung and Puchta, 2004; unpublished data). Furthermore, different alleles of the knockout mutant of AtRecQl4A show UVC-light and MMS sensitivity and an enhanced recombination rate in planta (Bagherieh-Najjar et al., 2005; Hartung and Puchta, unpublished results). Thus, AtRecQl4A seems to possess important and unique functions in comparison to AtRecQl4B and all other RecQ homologues of Arabidopsis. Interestingly, the overexpression of the *E. coli* RecQ gene in rice caused also an increase in extrachromosomal homologous recombination which is antagonistic to its characterized function in *E. coli* (Li et al., 2004). This contrasting phenotype is probably caused by interference of the heterologous RecQ gene from *E. coli* with the functions of the housekeeping RecQ-like genes in rice (Bagherieh-Najjar et al., 2005).

## RecQ homologues in plants – bioinformatical analysis

In contrast to the relatively small amount of functional studies, a lot of structural and evolutionary data can be obtained analysing several finished or virtually finished plant genomes (*A. thaliana*, *Oryza sativa* and *Chlamydomonas reinhardtii* and the almost complete genome sequence of the moss *Physcomitrella patens*; PHYSCObase; moss.nibb.ac.jp). For a structural analysis of plant RecQ-like proteins a total of seven different RecQ-like genes are available from the Arabidopsis genome, seven from rice and at least five in the moss *P. patens*. Plant RecQ homologues were first characterized in Arabidopsis and designated as AtRecQl 1, 2, 3, 4A, 4B, 5 and AtRecQsim for the gene with an interrupted helicase domain (Hartung et al., 2000; Bagherieh-Najjar et al., 2003; Hartung and Puchta, 2004; see Fig. 1).

A major question is to define common motifs of the RecQ homologues. The extensively described RecQ protein of *E. coli* possesses three domains which are typical for RecQ-like proteins and shared by most of them (Bernstein and Keck, 2003; see also Fig. 2). The most prominent domain of all is the **helicase region** which consists of seven different motifs numbered from I to VI with a total size of usually 350 400 amino acids including a motif Ia (Fig. 2). These regions contain sequences which are necessary for ATP binding, hydrolysis and DNA unwinding as it has been shown in vitro (Lu et al., 1996; Gray et al., 1997; Karow et al., 1997; Bennett et al., 1998; Ahmad et al., 2002; Özsoy et al., 2001, 2003). Moreover, five individual missense mutations in the helicase region of the human BLM protein lead to the Blooms syndrome due to the missing ATPase and helicase functions of the BLM protein (Bahr et al., 1998; Barakat et al., 2000; Rong et al., 2000). Additionally, a point mutation in the helicase region of yeast SGS1 protein leads to hypersensitivity to DNA damaging agents very similar to the SGS1 null mutant (Saffi et al., 2000; Mullen et al., 2000). Thus in general, the helicase region seems to be very important for the cellular functions of RecQ-like proteins.

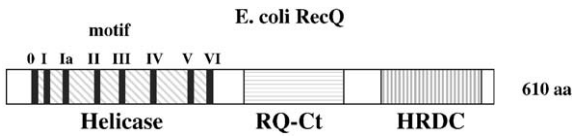


**Figure 2.** Schematic structural organization of the RecQ protein from *E. coli*. All three domains which are common for Rica-like proteins are shown as follows: The helicase domain (diagonal stripes) containing eight different motifs (black bars numbered from 0 to VI); the RQ-Ct domain which is conserved in nearly all RecQ proteins (horizontal stripes) and the HRDC domain which is conserved in more than 50% of the RecQ proteins (vertical stripes) and also in RNase D.

Another region named motif 0 was recently described by Bernstein and Keck (2003). It is a small motif of 14 18 amino acids located just in front of motif I of the classical helicase domain and it is involved in transient DNA binding (Bernstein and Keck, 2003; see also Table 1 and Fig. 2). The motif 0 consists of the four strongly conserved amino acids L, G, F and Q spaced by eight non-conserved and two other conserved amino acids (Table 1). This motif can be found in virtually all RecQ-like proteins described so far and we suggest that it should be included in the definition of the helicase domain. This motif seems to be indispensable for the helicase domain function, due to its high conservation in organisms from all kingdoms (see Table 1). In Table 1, RecQ-like homologues of a broad spectrum of organisms are depicted with a description of the location of the respective motif 0, the RecQ-Ct (RQC) and the Homologous Region RNase D C-terminal (HRDC) domain.

All seven RecQ-like proteins of *A. thaliana* contain the conserved motif 0 but some amino acid changes occurred. AtRecQl2, 4A and 4B possess a changed first position (L to R or N, respectively) and AtRecQl2, sim and 5 an exchange at position 11 or 5, respectively (in two cases F Y and in RecQsim F L). The same holds true for all rice RecQ homologues but here we can see an additional amino acid exchange in OsRQ1 at position 1 (L V) and OsRQsim at position 12 (R K), which is not present in Arabidopsis. The motif 0 can be used for assigning the different RecQ homologues of plants (seven in Arabidopsis and in rice) to each other as the sequence is not highly conserved and changes in amino acid positions can give us valuable phylogenetic information. This is not the same for the classical motifs I VI as the conservation during evolution was much higher.

The second domain conserved between most of the RecQ-like proteins is called RQC and can be found in virtually all RecQ-like proteins. We found only two genes in which the RQC region is not detectable, RecQ5 from Arabidopsis and rice and RQ4 from human and Drosophila (see Table 1 and Fig. 1). Little is known about the function of this domain but like in the helicase domain, mutations in this region are also impairing the RecQ activity (Foucault et al., 1997; Neff et al., 1999; Onoda et al., 2000; Ui et al., 2001; Yankiwski et al., 2001). Interestingly, these mutations in the BLM protein changed conserved cystein residues which are part

**Table 1.** Sequence and location of motif 0 and the location of the RQC and HRDC domain in RecQ homologues from different organisms

| Organism | Consensus<br>LxxxFG-xxxFR-xxQ | Pos.<br>[aa] | RQC<br>[aa] | HRDC<br>[aa] | Size<br>[aa] | Acc.no. |
|---|---|---|---|---|---|---|
| **Bacteria** | | | | | | |
| Ec RecQ | LQETFG-YQQFR-PGQ | 19-32 | 334-407 | 530-610 | 610 | AAA67618 |
| Ng RecQ | LHEVFG-YPEFR-GRQ | 11-24 | 327-400 | 525-767 (3x) | 767 | YP208755 |
| **Archaea** | | | | | | |
| Mea RecQ | LRQYFG-YTAFR-PLQ | 29-42 | 350-420 | 586-666 | 909 | NP619367 |
| **Cyanobacteria** | | | | | | |
| Ss RecQ | LRRIWG-YDHFR-YPQ | 11-24 | 335-475 | nd | 478 | NP440781 |
| Cw RecQ | LKDHFG-YDQFR-PGQ | 12-25 | 331-404 | 527-607 | 710 | ZP00518182 |
| **Fungi** | | | | | | |
| Sp Rqh1 | LKHKFH-LKGFR-KNQ | 511-524 | 839-916 | 1115-1195 | 1328 | Q09811 |
| Sp RecQ2 | LSQYYGLEAKFRSLKQ | 1361-1376 | 1698-1761 | nd | 2100 | CAE54423 |
| Sc SGS1 | LHEVFK-LPGFR-PNQ | 670-683 | 996-1074 | 1272-1351 | 1447 | P35187 |
| Af RecQ1 | LRDRFH-LRGFR-MNQ | 648-661 | 981-1058 | nd | 1468 | XP749627 |
| Af RecQ Tel3 | LQRVFR-KESFR-PLQ | 17-30 | 382-500 | nd | 538 | XP752103 |
| En MUS-N | LKERFH-LRGFR-PNQ | 692-705 | 1025-1103 | 1310-1395 | 1534 | AAF72650 |
| Nc QDE3 | LKDRFR-MSGFR-QNQ | 897-910 | 1228-1305 | 1582-1662 | 1955 | AAF31695 |
| Nc RecQ2 | LRRQFG-KENFR-PHQ | 16-29 | 367-447 | nd | 483 | CAD70358 |
| Ma RecQ1 | LRAVLRDDSARFRSPQ | 930-945 | 1254-1323 | nd | 1517 | AAL13172 |
| **Protists** | | | | | | |
| Eh RecQ | LHKCFN-IQSFR-PQQ | 28-41 | 352-414 | nd | 509 | XP653505 |
| Lm RecQ | MREVFG-LHDYR-FCQ | 261-274 | 587-686 | nd | 1691 | CAJ05096 |
| **Plants** | | | | | | |
| At RecQ1 | L-VIFG-NKVFR-PLQ | 199-213 | 524-605 | nd | 606 | CAC14163 |
| Os RecQ1 | V-VIFG-NKSFR-PLQ | ~200-215 | ~530-600 | nd | ~600 | AC134924 |
| At RecQ2 | RFNVFG-ISKYR-ANQ | 81-94 | 409-477 | 591-670 | 705 | CAC14866 |
| Os RecQ2 | RFNVFG-ISSYR-QNQ | 81-92 | 456-506 | 653-732 | 759 | AAX95427 |
| At RecQ3 | LRWHFG-HADFR-GKQ | 28-41 | 352-429 | nd | 713 | CAC14867 |
| Os RecQ3 | LKQHFG-YSGFR-GKQ | 30-43 | 292-377 | nd | 692 | XP468107 |
| Os RecQ_886 | LKTYFG-FSGFR-SYQ | 5-18 | 324-397 | 542-622 | 886 | XP479556 |
| At RecQ4A | NKKVFG-NHSFR-PNQ | 439-452 | 763-855 | 1012-1094 | 1182 | CAC14868 |
| Os RecQ4A | NKRVFG-NRSFR-PNQ | 429-442 | 752-841 | 980-1061 | 1164 | CAE03209 |
| Pp RecQ4A | NKKFFG-NKSFR-LNQ | ?? | ?? | ?? | ?? | |
| At RecQ4B | NKLVFG-NHSFR-PNQ | 461-474 | 785-876 | 1029-1111 | 1150 | CAC14869 |
| At RecQ5 | LNLVYG-YDSFR-DGQ | 220-233 | nd | nd | 870 | CAD13472 |
| Os RecQ5 | LNAAYG-HDSFR-QGQ | 277-290 | nd | nd | 874 | XM472951 |
| At RecQsim | LRNRFG-ISSLR-SFQ | 160-173 | 592-662 | nd | 858 | CAC14870 |
| Bn RecQsim | LRNRFG-ISSLR-SFQ | 144-157 | 591-663 | nd | 880 | AAO52679 |
| Os RecQsim | LRKHFG-FSCVK-GFQ | 167-180 | 606-675 | nd | 773 | AAO52678 |
| **Animals** | | | | | | |
| Hs BLM | FHKKFG-LHNFR-TNQ | 659-672 | 985-1068 | 1212-1292 | 1416 | P54132 |
| Xl BLM | FHKKFG-LHRFR-TNQ | 610-624 | 937-1020 | 1164-1244 | 1364 | AAG30928 |
| Ce BLM | LKSKFG-FNQFR-HRQ | 482-495 | 565-645 | 807-888 | 988 | Q18017 |
| Dm BLM | LSYSFG-LKSFR-PNQ | 729-742 | 1054-1137 | 1283-1363 | 1487 | AAF54691 |
| Hs WRN | LKMYFG-HSSFK-PVQ | 540-553 | 860-941 | 1150-1229 | 1432 | Q14191 |
| Xl WRN | LKTYFG-HSSFK-PVQ | 481-494 | 803-884 | 1098-1177 | 1436 | Q93530 |
| Ce WRN | LNEFFG-HKGFR-EKQ | 219-232 | 542-620 | 806-886 | 1056 | Q19546 |
| Hs RecQ4 | L-EQLG-HQAFR-PGQ | 473-485 | nd | nd | 1208 | Q94761 |
| Dm RecQ4 | L-HMFG-HTNFR-KGQ | 814-826 | nd | nd | 1530 | AAF42939 |
| Hs RecQ5-B | LKKVFG-FDSFKTPLQ | 20-34 | 355-436 | nd | 964 | NP004250 |
| Dm RecQ5-B | LKKHFG-HSKFKSDLQ | 11-25 | 352-433 | nd | 1058 | NP729983 |
| Ce RecQ1 | L-ELFC-HKKYRSRLQ | 170-183 | 510-600 | nd | 809 | NP497810 |
| Hs RecQ1 | LQNVFK-LEKFR-PLQ | 83-96 | 410-480 | nd | 649 | NP002898 |
| Ce RecQ2 | LKEQFH-LEKFR-PLQ | 478-491 | 807-877 | nd | 892 | P46064 |
| Dr RecQ1 | LCNIFQ-LSKFR-PLQ | 89-102 | 414-484 | nd | 639 | CAI21096 |

The above table shows the sequence and location of the conserved motif 0 of various selected RecQ-like proteins. The location of the RQC and HRDC domain is depicted for every protein which contains them and the accession numbers of the respective proteins are shown. Conserved amino acids are shown in grey boxes. Abbreviations are: Af = *Aspergillus fumigatus*; At = *Arabidopsis thaliana*; Bn = *Brassica napus*; Ce = *Caenorhabditis elegans*; Cw = *Crocosphaera watsonii*; Dm = *Drosophila melanogaster*; Dr = *Danio rerio*; Eh = *Entamoeba histolytica*; En = *Emericella nidulans*; Hs = *Homo sapiens*; Lm = *Leishmania major*; Ma = *Metarhizium anisopliae*; Nc = *Neurospora crassa*; Ng = *Neisseria gonorrhoe*, Mea = *Methanosarcina acetivorans*; Os = *Oryza sativa*; Pp = *Physcomitrella patens*; Sc = *Saccharomyces cerevisiae*; Sp = *Schizosaccharomyces pombe*; Ss = *Synechocystis speciosae*; Xl = *Xenopus laevis*.

of a described zinc finger (Guo et al., 2005). This zinc finger, located at the end of the RQC domain, is necessary to maintain the three-dimensional structure of the BLM protein. It is not clear whether the zinc finger is only involved in the correct protein folding or also in DNA binding of BLM or both (Guo et al., 2005).

The third conserved region of most RecQ-like proteins can be also found at the C-terminus of the RNase D and therefore it was called **HRDC**, see Fig. 2; Morozov et al., 1997). The HRDC region is not conserved in the majority of the RecQ-like proteins as depicted in Table 1. In more than half of the RecQ-like proteins no HRDC domain can be found and therefore a function for helicase activity cannot be assumed (Fig. 1). A possible function of this domain could be nucleic acid binding because this is the only known common property between RNase D and RecQ. As it has been shown in vivo, the HRDC region is dispensable for the hypersensitivity phenotype of yeast cells (Mullen et al., 2000, 2001) and a HRDC truncated SGS1 protein is still active as helicase and ATPase in vitro (Lu et al., 1996; Bennett et al., 1998). Nevertheless, an HRDC truncated SGS1 protein cannot complement a double mutant of SGS1 and Top3, which points to a function of the HRDC region in the interaction between Sgs1 and Top3 (Mullen et al., 2000). In a very recent paper, Wu and coworkers could show that mutation of a conserved lysine in the HRDC domain leads to inactivation of the dissolution (unwinding of double Holliday junctions) activity of human Blooms protein (Wu et al., 2005). This observation is a bit confusing because human RecQ1 and WRN possess the same lysine in the HRDC domain but cannot unwind double Holliday junctions. So probably there is an overall structural difference between the HRDC domain of BLM and the other RecQ-like proteins.

Except for two members all other plant RecQ-like proteins possess the typical helicase and RQC domains and overall structure mentioned before. These exceptions are AtRecQ5 and AtRecQsim. RecQsim harbors all helicase domains, the RQC but not the HRDC domain and it has an insertion of 100 110 amino acids in between the helicase motifs III and IV (Hartung et al., 2000). An insertion of nearly the same length is also present in the RecQsim homologues of rice and *B. napus* but it seems not to be present in the moss *P. patens* RecQsim homologue (Bagherieh-Najjar et al., 2003, rice: AC084218, *B. napus*: AY180332, moss: gnl-ti-717602022). Nearly one-third (31aa and 28aa in Arabidopsis and rice, respectively) of the approximately 100 amino acids long insertion are glutamic acid or aspartic acid residues. This high content of acidic amino acids points to a regulatory function unrelated to DNA binding but putatively involved in protein interactions. As already mentioned, RecQ5 from Arabidopsis and rice are the only plant RecQ-like proteins which have no RQC domain and therefore also no zinc binding domain. Interestingly, rice contains in its genome an open reading frame (OsRQ886, XP479556) which has no homologue in Arabidopsis. OsRQ886 has only 40% identity over the helicase domain to three different Arabidopsis RecQ-like proteins. Owing to this low identity, which is in the magnitude of homologies between plant and animal or yeast RecQ-like proteins, it is tempting to speculate that this gene is not a recent development in rice, but rather was lost during dicot evolution.

## Evolutionary considerations

It is obvious that all ''classical'' RecQ proteins like RecQ from *E. coli,* SGS1 from *S. cerevisiae,* BLM and WRN from humans possess an HRDC domain whereas the other ones do not. This may point to specialized functions for a number of RecQ-like proteins which obtained new or different tasks after duplication of the ancestral progenitor and subsequent loss or alteration of the HRDC domain. Because of the evolutionary relationship between bacteria, plants, animals and fungi, it is quite clear that the last common ancestor (LCA) of the three major eukaryotic kingdoms already possessed more than one RecQ-like gene. From the available data we cannot conclude how many copies were present in the LCA genome but the phylogenetic comparisons between the RecQ-like proteins demonstrate most often higher homology between proteins from different organisms than between different proteins in one organism. Furthermore, we can compare not only the protein sequences but also the positions of introns in respect to the protein sequence (Hartung et al., 2002). In case of the RecQ-like genes such a comparison is not feasible across the kingdom borders of animals, fungi and plants, but it is fruitful for the plant kingdom itself.

Comparing the intron positions of Arabidopsis and rice, nearly no differences can be found and even more remarkable the same holds true for the limited data from moss genomic sequences. In fact, according to several deviating intron positions of the different Arabidopsis RecQ homologues in conserved regions of the helicase domain, we could assign the rice sequences, and in several cases the partial moss sequences, to their Arabidopsis counterparts. For example, a sequence from BAC

AC134924, which is not annotated as a RecQ gene so far, could be clearly assigned as OsRQ1 by homology of the respective intron positions and a good homology in motif 0 (see Table 1). This means, in an evolutionary time scale of at least 150 million years (the time between monocot and dicot separation) there seem to have occurred only minor changes in the order and position of the intron sequences in RecQ-like genes. Regarding the moss sequences available so far, this seems to be valid even for a time span of 350 million years, the approximate time period between moss and angiosperm separation.

## Perspectives

The main research challenge regarding RecQ-like proteins in higher plants is the elucidation of their respective roles in recombination and repair pathways. One has to assume that the seven RecQ-like genes present in rice and Arabidopsis all possess different functions. Especially striking is the situation of the very recently duplicated gene pair in the Brassicaceae, RecQl4A and 4B, which have diverged functions. Therefore, it is of outstanding interest to analyse this gene pair and to define which differences in the respective amino acid sequence are responsible for functional differences between the two proteins. The described mutants of AtRecQl4A exhibit similar phenotypes to Sgs1 mutants of yeast but which role does AtRecQl4B play? Is it involved in another recombination pathway or does it interact with different proteins than AtRecQl4A? These are questions which can be tested in vitro by protein expression and helicase assays and in vivo by analysis of knockout mutants of AtRecQl4A and 4B and their double knockouts. Another very interesting gene is RecQsim which is obviously an older evolutionary development of plants as we can find it also in monocots. A first study on this gene showed that it is capable to partially complement the Sgs1 mutant phenotype (Bagherieh-Najjar et al., 2003). But so far it has not been demonstrated that the helicase function is indeed preserved. This seems to us rather unlikely because the RecQ helicase domain has not only been conserved in sequence but also in its structure over a billion years. Therefore, an interruption of 100 amino acids in the middle of the helicase domain might affect the protein function in one or the other way.

There is a presumably more ancient RecQ-like gene (OsRecQ886) present in rice but not in Arabidopsis. Is there another family member which took over the function of the gene in Arabidopsis? It would be of course interesting to analyse its function in rice and check whether expression of the protein is able to complement for one of the RecQ-like genes in Arabidopsis. A lot of work has to be done to characterize the biological role of different RecQ-like genes in plants. As stated above, we have to assume that all of them have different specificities and tasks, indicating that quite a number of different recombinogenic DNA structures have to be processed during the life cycle of a plant.

## Acknowledgements

We want to thank Manfred Focke and Stefanie Suer for critical reading of the manuscript and the Deutsche Forschungsgemeinschaft for funding the research on RecQ homologues in plants with the grant A15 in the SFB363 and recently with Grant Pu137/8-1.

## References

Ahmad F, Kaplan CD, Stewart E. Helicase activity is only partially required for *Schizosaccharomyces pombe* Rqh1p function. Yeast 2002;19:1381 98.

Bagherieh-Najjar MB, de Vries OMM, Hille J, Dijkwel PP. Arabidopsis RecQl4A suppresses homologous recombination and modulates DNA damage response. Plant J 2005;43:789 98.

Bagherieh-Najjar MB, de Vries OMM, Kroon JTM, Wright EL, Elborough KM, Hille J, et al. *Arabidopsis* RecQsim, a plant-specific member of the RecQ helicase family, can suppress the MMS hypersensitivity of the yeast *sgs1* mutant. Plant Mol Biol 2003;52:273 84.

Bahr A, De Graeve F, Kedinger C, Chatton B. Point mutations causing Bloom's syndrome abolish ATPase and DNA helicase activities of the BLM protein. Oncogene 1998;17:2565 71.

Barakat A, Ababou M, Onclerq R, Dutertre S, Chadli E, Hda N, et al. Identification of a novel BLM missense mutation (2706T>C) in a Moroccan patient with Bloom's syndrome. Hum Mutat 2000;15:584 5.

Bennett RJ, Keck JL. Structure and function of RecQ DNA helicases. Crit Rev Biochem Mol Biol 2004;39(2): 79 97.

Bennett RJ, Sharp JA, Wang JC. Purification and characterization of the Sgs1 DNA helicase activity of *Saccharomyces cerevisiae*. J Biol Chem 1998;273: 9644 50.

Bernstein DA, Keck JL. Domain mapping of *Escherichia coli* RecQ defines the roles of conserved N- and C-terminal regions in the RecQ family. Nucl Acids Res 2003;31:2778 85.

Chang M, Bellaoui M, Zhang C, Desai R, Morozov P, Delgado-Cruzata L, et al. RMI1/NCE4, a suppressor of genome instability, encodes a member of the RecQ Helicase/Topo III complex. EMBO J 2005;24:2024 33.

Cheok CF, Wu L, Garcia PL, Janscak P, Hickson ID. The Bloom's syndrome helicase promotes the annealing of complementary single-stranded DNA. Nucl Acids Res 2005;33:3932 41.

Cogoni C, Macino G. Posttranscriptional gene silencing in *Neurospora* by a RecQ DNA helicase. Science 1999; 286:2342 4.

Ellis NA, Groden J, Ye TZ, Straughen J, Lennon DJ, Ciocci S, et al. The Bloom's syndrome gene product is homologous to RecQ helicases. Cell 1995;83:655 66.

Foucault F, Vaury C, Barakat A, Thibout D, Planchon P, Jaulin C, et al. Characterization of a new BLM mutation associated with a topoisomerase II alpha defect in a patient with Bloom's syndrome. Hum Mol Genet 1997;6:1427 34.

Frei C, Gasser SM. The yeast Sgs1p helicase acts upstream of Rad53p in the DNA replication checkpoint and colocalizes with Rad53p in S-phase-specific foci. Genes Dev 2000;14:81 96.

Gangloff S, Soustelle C, Fabre F. Homologous recombination is responsible for cell death in the absence of the Sgs1 and Srs2 helicases. Nat Genet 2000;25:192 4.

Garcia PL, Liu Y, Jiricny J, West SC, Janscak P. Human RECQ5beta, a protein with DNA helicase and strand-annealing activities in a single polypeptide. EMBO J 2004;23:2882 91.

Gray MD, Shen JC, Kamath-Loeb AS, Blank A, Sopher BL, Martin GM, et al. The Werner syndrome protein is a DNA helicase. Nat Genet 1997;17:100 3.

Guo RB, Rigolet P, Zargarian L, Fermandjian S, Xi XG. Structural and functional characterizations reveal the importance of a zinc binding domain in Bloom's syndrome helicase. Nucl Acids Res 2005;33:3109 24.

Hanada K, Ukita T, Kohno Y, Saito K, Kato J, Ikeda H. RecQ DNA helicase is a suppressor of illegitimate recombination in *Escherichia coli*. Proc Natl Acad Sci USA 1997;94:3860 5.

Hartung F, Blattner FR, Puchta H. Intron gain and loss in the evolution of the conserved eukaryotic recombination machinery. Nucl Acids Res 2002;30:5175 81.

Hartung F, Plchova H, Puchta H. Molecular characterisation of RecQ homologues in *Arabidopsis thaliana*. Nucl Acids Res 2000;28:4275 82.

Hartung F, Puchta H. What comparative genomics tell us about the evolution of eukaryotic genes involved in recombination. Curr Genom 2004;5:109 21.

Hickson ID. RecQ helicases: caretakers of the genome. Nat Rev Cancer 2003;3:169 78.

Hofmann AF, Harris SD. The *Aspergillus nidulans* musN gene encodes a RecQ helicase that interacts with the PI-3K-related kinase UVSB. Genetics 2001;159: 1595 604.

Hu Y, Lu X, Barnes E, Yan M, Lou H, Luo G. RecQl5 and Blm RecQ DNA helicases have nonredundant roles in suppressing crossovers. Molec Cellular Biol 2005;25: 3431 42.

Huang P, Pryde FE, Lester D, Maddison RL, Borts RH, Hickson ID, et al. SGS1 is required for telomere elongation in the absence of telomerase. Curr Biol 2001;11:125 9.

Inglis PW, Rigden DJ, Mello LV, Louis EJ, Valadares-Inglis MC. Monomorphic subtelomeric DNA in the filamentous fungus, *Metarhizium anisopliae*, contains a RecQ helicase-like gene. Mol Genet Genom [Epub ahead of print], 2005.

Jeong SM, Kawasaki K, Juni N, Shibata T. Identification of *Drosophila melanogaster* RECQE as a member of a new family of RecQ homologues that is preferentially expressed in early embryos. Mol Gen Genet 2000;263: 183 93.

Johnson FB, Marciniak RA, McVey M, Stewart SA, Hahn WC, Guarente L. The *Saccharomyces cerevisiae* WRN homolog Sgs1p participates in telomere maintenance in cells lacking telomerase. EMBO J 2001;20:905 13.

Kaliraman V, Mullen RJ, Fricke WM, Bastin-Shanower SA, Brill SJ. Functional overlap between Sgs1-Top3 and the Mms4-MUS81-endonuclease. Genes Develop 2001;15: 2730 40.

Karow JK, Chakraverty RK, Hickson ID. The Bloom's syndrome gene product is a 3′ 5′ DNA helicase. J Biol Chem 1997;272:30611 4.

Karow JK, Constantinou A, Li JL, West SC, Hickson ID. The Bloom's syndrome gene product promotes branch migration of holliday junctions. Proc Natl Acad Sci USA 2000;97:6504 8.

Kitao S, Shimamoto A, Goto M, Miller RW, Smithson WA, Lindor NM, et al. Mutations in RECQL4 cause a subset of cases of Rothmund Thomson syndrome. Nat Genet 1999;22:82 4.

Li HQ, Terada R, Li MR, Iida S. RecQ helicase enhances homologous recombination in plants. FEBS Lett 2004;574:151 5.

Lillard-Wetherell K, Combs KA, Groden J. BLM helicase complements disrupted type II telomere lengthening in telomerase-negative sgs1 yeast. Cancer Res 2005; 65:5520 2.

Lu J, Mullen JR, Brill SJ, Kleff S, Romeo AM, Sternglanz R. Human homologues of yeast helicase. Nature 1996; 383:678 9.

Machwe A, Xiao L, Groden J, Matson SW, Orren DK. RecQ family members combine strand pairing and unwinding activities to catalyze strand exchange. J Biol Chem 2005;280:23397 407.

Maftahi M, Hope JC, Delgado-Cruzata L, Han CS, Freyer GA. The severe slow growth of Δsrs2 Δrqh1 in *Schizosaccharomyces pombe* is suppressed by loss of recombination and checkpoint genes. Nucl Acids Res 2002;30:4781 92.

Mandell JG, Goodrich KJ, Bahler J, Cech TR. Expression of a RecQ helicase homolog affects progression through crisis in fission yeast lacking telomerase. J Biol Chem 2005;280:5249 57.

Morozov V, Mushegian AR, Koonin EV, Bork P. A putative nucleic acid-binding domain in Bloom's and Werner's syndrome helicases. Trends Biochem Sci 1997;22: 417 8.

Mullen JR, Kaliraman V, Ibrahim SS, Brill SJ. Requirement for three novel protein complexes in the absence of the Sgs1 DNA helicase in *Saccharomyces cerevisiae*. Genetics 2001;157:103 18.

Mullen JR, Kaliraman V, Brill SJ. Bipartite structure of the SGS1 DNA helicase in *Saccharomyces cerevisiae*. Genetics 2000;154:1101 14.

Mullen JR, Nallaseth FS, Lan YQ, Slagle CE, Brill SJ. Yeast Rmi1/Nce4 controls genome stability as a subunit of the Sgs1 Top3 complex. Mol Cell Biol 2005;25: 4476 87.

Nakayama K, Irino N, Nakayama H. The recQ gene of Escherichia coli K12: molecular cloning and isolation of insertion mutants. Mol Gen Genet 1985;200:266 71.

Neff NF, Ellis NA, Ye TZ, Noonan J, Huang K, Sanz M, et al. The DNA helicase activity of BLM is necessary for the correction of the genomic instability of Bloom syndrome cells. Mol Biol Cell 1999;10:665 76.

Onoda A, Seki M, Miyajima A, Enomoto T. Elevation of sister chromatid exchange in *Saccharomyces cerevisiae* sgs1 disruptants and the relevance of the disruptants as a system to evaluate mutations in Bloom's syndrome gene. Mutat Res 2000;459:203 9.

Ooi SL, Shoemaker DD, Boeke JD. DNE helicase gene interaction network defined using synthetic lethality analyzed by microarray. Nat Genet 2003;35:277 86.

Opresko PL, Mason PA, Podell ER, Lei M, Hickson ID, Cech TR, Bohr VA. POT1 stimulates recQ helicases WRN and BLM to unwind telomeric DNA substrates. J Biol Chem 2005;280:32069 80.

Özsoy AZ, Ragonese HM, Matson SW. Analysis of helicase activity and substrate specificity of Drosophila RECQ5. Nucl Acids Res 2003;31:1554 64.

Özsoy AZ, Sekelsky JJ, Matson SW. Biochemical characterization of the small isoform of *Drosophila melanogaster* RECQ5 helicase. Nucl Acids Res 2001;29: 2986 93.

Plchova H, Hartung F, Puchta H. Biochemical characterization of an exonuclease from *Arabidopsis thaliana* reveals similarities to the DNA exonuclease of the huma Werner syndrome protein. J Biol Chem 2003; 278:44128 38.

Rong SB, Valiaho J, Vihinen M. Structural basis of Bloom syndrome (BS) causing mutations in the BLM helicase domain. Mol Med 2000;6:155 64.

Saffi J, Pereira VR, Henriques JA. Importance of the Sgs1 helicase activity in DNA repair of *Saccharomyces cerevisiae*. Curr Genet 2000;37:75 8.

Schawalder J, Paric E, Neff NF. Telomere and ribosomal DNA repeats are chromosomal targets of the bloom syndrome DNA helicase. BMC Cell Biol 2003;4:1 15.

Sekelsky JJ, Brodsky MH, Rubin GM, Hawley RS. Drosophila and human RecQ5 exist in different isoforms generated by alternative splicing. Nucl Acids Res 1999;27:3762 9.

Shen JC, Loeb LA. The Werner syndrome gene: the molecular basis of RecQ helicase-deficiency diseases. Trends Genet 2000;16:213 20.

Shimamoto A, Nishikawa K, Kitao S, Furuichi Y. Human RecQ5beta, a large isomer of RecQ5 DNA helicase, localizes in the nucleoplasm and interacts with topoisomerases 3alpha and 3beta. Nucl Acids Res 2000;28:1647 55.

Sinclair DA, Guarente L. Extrachromosomal rDNA circles a cause of aging in yeast. Cell 1997;91:1033 42.

Ui A, Satoh Y, Onoda F, Miyajima A, Seki M, Enomoto T. The N-terminal region of Sgs1, which interacts with Top3, is required for complementation of MMS sensitivity and suppression of hyper-recombination in sgs1 disruptants. Mol Genet Genom 2001;265:837 50.

Umezu K, Nakayama K, Nakayama H. *Escherichia coli* RecQ protein is a DNA helicase. Proc Natl Acad Sci USA 1990;87:5363 7.

Wang W, Seki M, Narita Y, Nakagawa T, Yoshimura A, Otsuki M, et al. Functional relation among RecQ family helicases RecQL1, RecQL5, and BLM in cell growth and sister chromatid exchange formation. Mol Cell Biol 2003;23:3527 35.

Watt PM, Hickson ID, Borts RH, Louis EJ. Sgs1, a homologue of the Bloom's and Werner's syndrome genes suppress hyperrecombination in yeast Sgs1 mutant: implication for genomic instability in human diseases. Genetics 1996;95:8733 8.

Watt PM, Louis EJ, Borts RH, Hickson ID. Sgs1: a eukaryotic homolog of *E. coli* RecQ that interacts with topoisomerase II in vivo and is required for faithful chromosome segregation. Cell 1995;81: 253 60.

Wu L, Chan KL, Ralf C, Bernstein DA, Garcia PL, Bohr VA, et al. The HRDC domain of BLM is required for the dissolution of double Holliday junctions. EMBO J 2005;24:2679 87.

Wu L, Hickson ID. The Bloom's syndrome helicase suppresses crossing over during homologous recombination. Nature 2003;426:870 4.

Yamagata K, Kato J-I, Shimamoto A, Goto M, Furuichi Y, Ikeda H. Bloom's and Werner's syndrome genes suppress hyperrecombination in yeast Sgs1 mutant: implication for genomic instability in human diseases. Genetics 1998;95:8733 8.

Yankiwski V, Noonan JP, Neff NF. The C-terminal domain of the Bloom syndrome DNA helicase is essential for genomic stability. BMC Cell Biol 2001;2:1 11.

Yu CE, Oshima J, Fu YH, Wijsman EM, Hisama F, Alisch R, et al. Positional cloning of the Werner's syndrome gene. Science 1996;272:258 62.

Zhang R, Sengupta S, Yang Q, Linke SP, Yanaihara N, Bradsher J, et al. BLM helicase facilitates MUS81 endonuclease activity in human cells. Cancer Res 2005;65:2526 31.