

Role of Human Disease Genes for the Maintenance of Genome Stability in Plants

H Puchta*, D Kobbe, K Wanieck, A Knoll, S Suer, M Focke & F Hartung

Abstract

In the plant genome a row of homologs of human genes can be found that, if mutated, are correlated with a high incidence of cancer in humans. Here we describe our recent results on homologs of the breast cancer genes BRCA1/BARD1 and RecQ helicase homologs in the model plant *Arabidopsis thaliana*. HsBRCA1 and HsBARD1 are tumor suppressor proteins that are involved in many cellular processes, such as DNA repair. Loss of one or the other protein results in early embryonic lethality and chromosomal instability. The *Arabidopsis* genome harbors one BRCA1 homolog, and we were able to identify a BARD1 homolog as well. AtBRCA1 and AtBARD1 are able to interact with each other as indicated by *in vitro* and *in planta* experiments. Our analyses of T-DNA insertion mutants for both genes, revealed that in plants, in contrast to animals, these genes are dispensable for development or meiosis. Nevertheless, we could show that AtBARD1 plays a prominent role in the regulation of homologous DNA repair in somatic cells. RecQ helicases are known as mediators of genome stability. The loss of RecQ function is often accompanied by hyperrecombination due to a lack of crossover suppression. *Arabidopsis thaliana* possesses seven different RecQ genes. We could show that two of them (AtRECQ4A and 4B) arose as a result of a recent duplication and are still 70% identical on protein level. Disruption of these genes, surprisingly, leads to antagonistic phenotypes: the *AtRECQ4A* mutant shows sensitivity to DNA damaging agents, enhanced homologous recombination and lethality in an *Atmus81* background. Moreover, mutation of *AtRECQ4A* partially suppresses the lethal phenotype of an *AtTOP3a* mutant. In contrast, the *AtRECQ4B* mutant shows a reduced level of HR and none of the other phenotypes described above. Finally, we have started to characterize the different RecQ proteins of *Arabidopsis* by biochemical means and present here the results on AtRECQ2.

Breast Cancer Genes: BRCA1, BRCA2 and BARD1

Genome stability is a crucial aspect of any living organism. It is essential to cope with DNA damages and to repair them properly. Many human diseases, especially cancer, are linked to different kinds of defects in DNA repair mechanisms which might lead to chromosome abnormalities or uncontrolled proliferation. Defects in DNA repair or recombination genes often predispose humans to cancer development, as it is seen for instance in patients of Bloom and Werner syndrome or breast cancer. Those diseases are linked to mutations in genes whose encoded proteins function as DNA repair proteins. Usually, cancer is not due to a single mutation, but a certain mutation might lead to a higher susceptibility for developing cancer. As breast cancer is the most prevalent cancer worldwide many studies have been made to identify and to understand the role of certain breast cancer factors. In 1994, the first gene which seemed to be genetically linked to the development of hereditary breast cancer was

identified [1]. This gene was called *Breast Cancer Susceptibility Gene 1* (BRCA1). Besides BRCA1, another breast cancer gene was identified the following year and named BRCA2 [2]. Later on, it could be shown that both proteins are important factors for homologous recombination (HR) during DNA repair.

As proteins are often part of different complexes and their function depends on certain interactions, there are several known proteins, which are important for the function of BRCA1. The most prominent interaction partner of BRCA1 is the *Breast Cancer Associated Ring Domain Protein* BARD1. The name implies its interaction with BRCA1 over their RING domains. **Figure 1A** shows the general structure of the human breast cancer proteins. Homologs of these genes were identified in several other organisms as well, and interestingly also in the model plant *A. thaliana* [3-5].

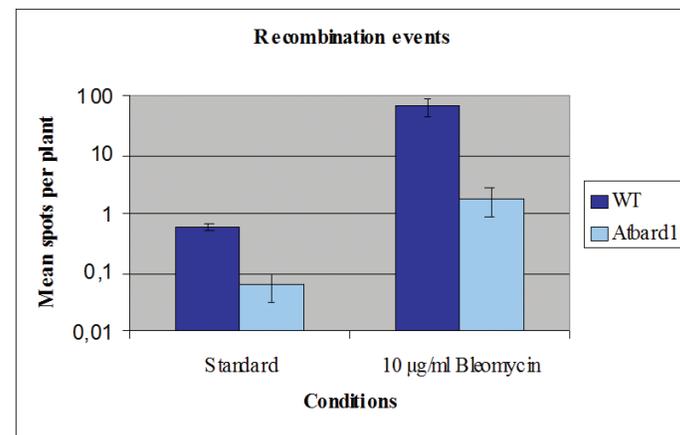
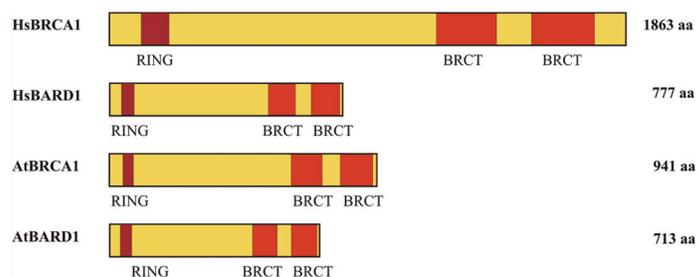


Figure 1 Structure of human and plant breast cancer proteins and the AtBARD1 function during HR. **(A)** The characteristic features of breast cancer proteins are the N-terminal RING and the C-terminal BRCT domains. Both domains are important for protein-protein interactions and the proper functions of the proteins. RING = Really Interesting New Gene; BRCT = BRCA1 C-terminal; Hs = Homo sapiens; At = *Arabidopsis thaliana*. **(B)** The *Atbard1* mutant plants show a reduction of HR under standard conditions in comparison to the wild type (WT). After induction of DSB with Bleomycin, the wild type shows a clear induction of HR. The HR in *Atbard1* mutant plants after induction with Bleomycin is less induced, indicating a role for AtBARD1 in HR. The diagram shows the mean number of blue spots per plant (bars show the SD) from three different experiments with a logarithmic scale for a better presentation.

University of Karlsruhe, Department of Botany II, Karlsruhe, Germany

* Corresponding author. E-mail: holger.puchta@bio.uka.de

Structure of plant homologs

As mentioned above, in *A. thaliana* and other plants, gene homologs of BRCA1, BRCA2 and BARD1 were identified. AtBRCA1 and AtBARD1 are single copy genes, whereas AtBRCA2 is a duplicated gene. The homology between the breast cancer genes and their plant orthologs is mainly conserved in the region of their functional domains. Both BRCA1 and BARD1 possess an N-terminal RING and two C-terminal BRCT-domains. The RING domain of AtBRCA1 has 34% identity and 61% similarity to the human ortholog, whereas the homology of the BRCT region has 28% identity and 61% similarity [4]. The alignment of AtBARD1 with HsBARD1 shows 22% identity and 38% similarity [5]. The RING domain is a protein interaction domain, which is necessary for the E3-Ubiquitin Ligase function of BRCA1 and BARD1 alone as well as for the respective heterodimer. The function of the BRCT domains is based on interactions with phosphopeptides and might therefore be important for regulatory processes. Both BRCA1/BARD1 and BRCA2 play essential roles during mitotic and/or meiotic DNA repair and recombination events. Their essential role in mammals is shown in early embryonic lethality of homozygous BRCA1, BARD1 and BRCA2 mutant mice [6-8]. Interestingly, in *A. thaliana* homozygous single mutant plants of the breast cancer genes are, in contrast to their mammalian counterparts, not lethal. This offers a unique system for studying putative, conserved functions of these breast cancer genes in a higher eukaryotic organism.

Functional Analysis in Plants

BRCA1 and BARD1 are important factors of homologous recombination in somatic cells. Studies in different organisms led to the conclusion that this function seems to be conserved. Via a recombination assay system for *A. thaliana*, it was shown that the *Atbard1* single mutants display a defect in homologous recombination and that they are sensitive to MMC, which leads to DNA cross links [5]. **Figure 1B** shows the clear reduction of HR in *Atbard1* mutant plants under standard as well as under genotoxic conditions indicating a role for AtBARD1 in HR. AtBRCA2 is also important for HR and plays an essential role during meiosis [3]. As the genome of *A. thaliana* harbors two almost identical homologs, those proteins might have a redundant function.

Research on the function of breast cancer genes in plants might give insight into conserved functions of those genes, as well as reveal plant specific functions. Our present studies concentrate on other proteins that interact with breast cancer proteins. One such family is the RecQ family, whose proteins are also linked to several human diseases.

RecQ helicases as genomic caretakers

Out of the five human homologs of *E. coli* RecQ, mutations in three of them have been shown to result in severe autosomal recessive hereditary diseases. Bloom Syndrome (BS), Werner Syndrome (WS) and Rothmund-Thomson Syndrome (RTS) result from biallelic loss-of-function mutations in the genes BLM, WRN or HsRECQ4, respectively. Patients diagnosed with the disorders RAPADILINO Syndrome and Baller-Gerold Syndrome (BGS) have recently been shown to carry mutations in the HsRECQ4 gene as well.

All of these syndromes result in a set of common characteristics, for example genomic instability and a predisposition to cancer malignancies [9]. However, there are also syndrome-specific features and unique cellular and genetic defects that suggest non-redundant cellular functions for these RecQ helicases. BS patients present with a proportional growth deficiency, skin abnormalities, such as sun sensitivity, hypo- and hyperpigmentation, fertility defects and changes in fat and sugar metabolism [10]. Notably, there is an increased predisposition to all types of cancer with high incidence. In BS fibroblasts, the hallmark characteristic is the elevated rate of sister chromatid exchanges due to increased homologous recombination.

Shortly after the connection of BLM with BS, it was shown that mutations in the WRN gene are causative for the recessive disorder Werner Syndrome [11,12]. WS is a progeroid disease resulting in premature aging that develops during the second decade of life and is associated with age-related disorders like greying and loss of hair, skin atrophy, atherosclerosis, osteoporosis, bilateral cataracts and type II diabetes mellitus. WS is also associated with a high incidence of cancer, but contrary to BS patients, individuals with WS show a predisposition primarily to sarcomas. On the cellular level, WS manifests in genomic instability due to chromosome breaks, reciprocal chromosomal translocations and genomic deletions.

The third and only other human RecQ helicase that has been associated with a disease is HsRECQ4, mutated in about 60% of all persons diagnosed with Rothmund-Thomson Syndrome [13]. RTS manifests in skeletal abnormalities, poikiloderma (skin atrophy and dyspigmentation), cataracts, hypogonadism, early greying and loss of hair. The cancer predisposition typical for RecQ-related disorders is seen in RTS as well, but it is restricted mainly to osteosarcomas.

RecQ helicases in plants

In total, there are seven different RecQ like genes present in the model plant *A. thaliana* (**Fig. 2**). As mentioned before, mutations in the human gene coding for the BLM protein lead to a severe genetic disorder called Bloom syndrome [12]. On the sequence level, two of them, *AtRECQ4A* and *AtRECQ4B*, can be considered putative HsBLM homologs. Regarding the conserved domains within the seven helicase motifs, both proteins share a sequence identity of approximately 53% with the HsBLM protein, and incidentally about 46% with the yeast RecQ homolog SGS1.

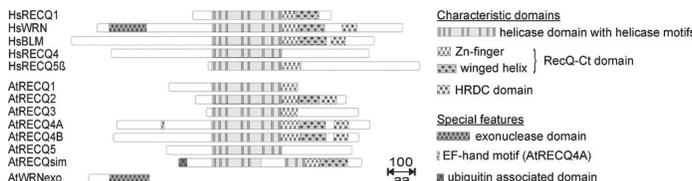


Figure 2 Human and Arabidopsis RecQ homologs. The typical structure of RecQ helicases is found in virtually all members of this family. The RecQ helicase domain, the RecQ-Ct and the HRDC domain (Helicase and RNase D C-terminal) are most common, whereas the exonuclease domain is only found in the WRN protein, and the EF-hand motif only in AtRECQ4A. In plants and insects, the exonuclease domain of the HsWRN protein is coded by a small gene, which seems to be fused in HsWRN.

AtRECQ4A and *AtRECQ4B* exhibit an identity of about 70% regarding their DNA and protein sequences, and have therefore arisen from a recent duplication event [14]. Amazingly, in contrast to their high sequence similarity, mutations in *AtRECQ4A* and *4B* lead to oppositional phenotypes respectively. Whereas *Atrecq4A* reflects the “RecQ typical” phenotype showing sensitivity towards genotoxic agents, such as MMS or cisplatin, *Atrecq4B* plants are not more affected than the wild type control plants [15,16]. Furthermore, we could show hyperrecombination for *Atrecq4A*, as it has been shown for other RecQ mutants, such as human *blm* or yeast *sgs1* [15-18]. In contrast to the expected results obtained for *Atrecq4A* as typical RecQ mutant, in *Atrecq4B* plants somatic HR is strongly reduced compared to the wild type [15]. This hyporecombination phenotype has not been described for any eukaryotic RecQ mutant so far, and points to a positive involvement of the AtRECQ4B protein in the recombination process.

Another property that seems to be conserved in AtRECQ4A is the interaction with a type 1A topoisomerase. This interaction has directly been shown for yeast SGS1 with TOP3, as well as for the human BLM protein with HsTOP3a [19]. Due to the severity of the phenotypes, it is problematic to study eukaryotic TOP3a functions with the help of

the respective mutants. Nevertheless, we could show very recently for *Atrecq4A* that this mutation is able to rescue the lethal phenotype of *top3a* mutants in *Arabidopsis*, resulting in sterile but viable plants [15].

Finally, double mutants of *Atmus81* and *Atrecq4A* develop poorly and die within about two weeks. This is in line with the lethality found in double mutants of the structure-specific endonuclease MUS81 and the respective RecQ homologs of budding and fission yeast, SGS1 and RQH1, which are synthetically lethal. This result points to a high level of conservation of the somatic RECQ and MUS81 functions, such as the involvement of both proteins in two parallel pathways working on stalled replication forks [16].

In both cases, the results obtained for the RecQ double mutants with *mus81* or *top3a*, respectively, are restricted to the *recq4A* mutation of *A. thaliana*, whereas for *Atrecq4B* no effect on the *Attop3a* and *Atmus81* phenotypes could be observed [15,16]. Therefore it can be concluded that in *Arabidopsis* RECQ4A, and not RECQ4B, is in most aspects the functional homolog of BLM and SGS1. Nevertheless, the recombination promoting function of RECQ4B might also have originated from a common BLM-like ancestor protein.

Biochemical characterization of AtRECQ2

Whereas the T-DNA insertion mutants enable us to analyze the effect of missing or truncated proteins, biochemical analysis reveals what reactions the proteins are possibly able to catalyze. The two approaches are complementary, with the biochemical analysis a reaction is being tested for feasibility, for example, if a participation in a specific pathway is proposed. However, biochemical analysis does not state whether a specific reaction is really taking place *in vivo*, since it may be influenced, for instance, by post-translational modifications of the protein or by protein interaction partners. On the other hand, with the help of specific information on the types of substrates the enzyme might process, a more focused *in vivo* analysis can be performed.

The method used is briefly introduced here: DNA sequences are designed with complementary and non-complementary stretches in such a way that a specific DNA-structure will form. The DNA structure is built by heating one labeled oligonucleotide (32 P) together with the other constituents and cooling them down slowly with subsequent purification. It is incubated with the enzyme under defined conditions. Then, the reaction products are analyzed via native polyacrylamide gel electrophoresis (PAGE), in which the original structure is separated from the product(s). The structures containing a 32 P label can be analyzed and quantified allowing the calculation of the percentage of unwinding.

Here, some data of our biochemical characterization of *A. thaliana* RECQ2 is presented [20]. AtRECQ2, together with AtRECQ4A and AtRECQ4B, possess the complete set of (uninterrupted) domains, characteristic for RecQ-helicases, as do HsWRN and HsBLM (Fig. 2). Additionally, AtRECQ2 interacts with AtWRNexo [21]. AtWRNexo is homologous to the exonuclease domain of HsWRN (Fig. 2) and biochemical analysis of AtWRNexo has revealed conserved properties [22]. We analyzed the biochemical functions of AtRECQ2 in order to classify it either as orthologous to the HsWRN-helicase or as a potentially plant-specific protein with its own set of functions.

We expressed AtRECQ2 in *E. coli* and successfully purified it with the help of an N-terminal calmodulin binding peptide tag and a C-terminal hexahistidine tag. In order to be able to judge the purity of our AtRECQ2 preparation, we also cloned, expressed and purified AtRECQ2-K117M in an identical fashion. For AtRECQ2-K117M, the substitution of lysine by methionine in the Walker A motif leads to an abolishment of ATPase and therefore helicase activity. Thus, an activity observed in assays with AtRECQ2 that is missing with AtRECQ2-K117M is due to our enzyme of interest.

We were able to show that AtRECQ2 is a (d)NTP dependent 3' to 5' DNA helicase (Fig. 3A and B). This can not be taken for granted as for

example, no helicase activity was shown for HsRECQ4 [23]. The ability of AtRECQ2 to use all nucleotide cofactors to catalyze unwinding is not common for RecQ helicases. These properties are only similar to those of HsWRN helicase, for which ATP and dATP are best as well, followed by dCTP and CTP. Strand unwinding by HsWRN can also be measured with GTP, dGTP, UTP and dTTP but it is not efficient [24]. All other RecQ-homologs analyzed for the usage of different (d)NTPs are more restricted.

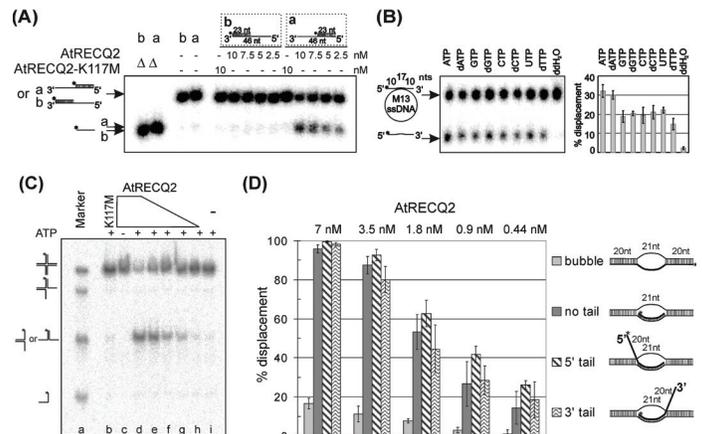


Figure 3 Biochemical analysis of AtRECQ2 [20]. **(A)** AtRECQ2 is a 3'-5' DNA helicase. The substrates are schematically drawn in the dotted boxes; * marks the 32 P-label. The arrows indicate the positions of migration of the substrates (top) or the 23 nt – unwinding products of substrate a (unwound by 3' 5' helicases) and b (unwound by 5' 3' helicases), respectively (bottom). The reactions with AtRECQ2-K117M show no unwinding, thus the AtRECQ2 preparation is devoid of helicase contamination. : heat-denatured substrate. **(B)** AtRECQ2 can use different nucleotide-triphosphates for unwinding. Helicase data are presented as autoradiogram and as mean with SD. For the respective reactions ATP was substituted by dATP, GTP, dGTP, CTP, dCTP, UTP, dTTP and water. **(C)** AtRECQ2 can branch migrate Holliday Junctions. Autoradiography of the reaction products. In reaction c ATP was omitted. It is clearly visible that the main product of AtRECQ2 is splayed arm, indicative for branch migration. **(D)** Quantification of helicase data (mean and SD) of reactions with a bubble and different D-Loop-structures, schematically drawn on the right. The invading strands of the D-Loop structures were unwound. No significant preference for a special D-Loop structure can be observed with different AtRECQ2 concentrations used.

The reaction of AtRECQ2 was also analyzed on DNA substrates that mimic recombination intermediates: a partially mobile Holliday Junction and different D-Loops (Fig. 3C and 3D). The analysis of the reaction of AtRECQ2 on Holliday Junctions reveals splayed arm products, characteristic for branch migration. As it was shown before by others, splayed arm is the main product of many RecQ-helicases - also for HsWRN [25]. D-Loops are formed in an early step of homologous recombination, when an ssDNA invades the homologous dsDNA and pairs with the complementary strand of the duplex. AtRECQ2 can displace invading strands of D-Loops, regardless if there is a protruding ssDNA or not, and irrespectively of the protruding ssDNA's directionality. The analyzed bubble was partly unwound. Melting of productive D-Loops in which the 3' end of ssDNA is invading can be considered as anti-recombinogenic. Also, unwinding of the unproductive D-Loops with 5' invasions and 3' protruding tails by RecQ-helicases may be important. The data of other RecQ helicases on their action on D-Loops are published. Whereas some RecQ proteins show a preference for 3' tailed D-Loops, the properties of HsWRN [26] are similar to those of AtRECQ2: the invading strands of all three D-Loop substrates are similarly well displaced.

To sum up, the biochemical properties of AtRECQ2 are closest to those of HsWRN. Therefore, the hypothesis that those two are functionally homologous is reinforced and it will be highly interesting to see whether the future analysis of the respective T-DNA insertion mutants will sustain this conclusion.

In general, we can summarize that a reasonable number of genes involved in genome stability and cancer predisposition in animals are well-conserved in plants. The function of these genes is often similar on a general level, such as the preservation of genome stability, as well as regarding their biochemical properties. Nevertheless, besides helping to understand basic mechanisms of genome stability in eukaryotes, our research also has a strong biotechnological potential. A better understanding of homologous recombination in plants might help us to set up new approaches in green gene technology, such as gene targeting or improved breeding.

ACKNOWLEDGEMENTS

We want to thank S. Blanck, K. Demand, W. Reidt, R. Wurz-Wildersinn, S. Hettinger, C. Moock, S. Buss and S. Zeiler for their contribution to this article. We want to apologize that, due to space limitation, we were not able to cite all the relevant literature. This work was supported by DFG grants HA5055/1-1 and PU137/10-1.

BIBLIOGRAPHY

- Miki, Y. *et al.* A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* **266**, 66–71 (1994).
- Wooster, R. *et al.* Identification of the breast cancer susceptibility gene BRCA2. *Nature* **378**, 789–92 (1995).
- Siaud, N. *et al.* *Brca2* is involved in meiosis in *Arabidopsis thaliana* as suggested by its interaction with *Dmc1*. *Embo J.* **23**, 1392–1401 (2004).
- Lafarge, S., And Montane, M.H. Characterization of *Arabidopsis thaliana* ortholog of the human breast cancer susceptibility gene 1: *AtBRCA1*, strongly induced by gamma rays. *Nucleic Acids Res* **31**, 1148–55 (2003).
- Reidt, W., Wurz, R., Wanieck, K., Chu, H.H., Puchta, H. A homologue of the breast cancer-associated gene BARD1 is involved in DNA repair in plants. *Embo J.* **25**, 4326–37 (2006).
- Sharan, S.K. *et al.* Embryonic lethality and radiation hypersensitivity mediated by Rad51 in mice lacking *Brca2*. *Nature* **386**, 804–810 (1997).
- Gowen, L.C. *et al.* *Brca1* deficiency results in early embryonic lethality characterized by neuroepithelial abnormalities. *Nature Genetics* **12**, 191–94 (1996).
- Mccarthy, E.E., Celebi, J.T., Baer, R., Ludwig, T. Loss of *Bard1*, the heterodimeric partner of the *Brca1* tumor suppressor, results in early embryonic lethality and chromosomal instability. *Mol Cell Biol* **23**, 5056–63 (2003).
- Ouyang, K.J., Woo, L.L., Ellis, N.A. Homologous recombination and maintenance of genome integrity: Cancer and aging through the prism of human *RecQ* helicases. *Mech Ageing Dev.* Mar 15 [Epub Ahead Of Print] (2008).
- German, J., Sanz, M.M., Ciocci, S., Ye, T.Z., Ellis, N.A. Syndrome-causing mutations of the BLM gene in persons in the Bloom's Syndrome Registry. *Hum. Mutat.* **28**, 743–53 (2007).
- Yu, C.E. *et al.* Positional cloning of the Werner's syndrome gene. *Science* **272**, 258–62 (1996).
- Ellis, N.A. *et al.* The Bloom's syndrome gene product is homologous to *RecQ* helicases. *Cell* **83**, 655–66 (1995).
- Kitao, S. *et al.* Mutations in RECQL4 cause a subset of cases of Rothmund-Thomson syndrome. *Nat Genet.* **22**, 82–84 (1999).
- Hartung, F., Puchta, H. The *RecQ* gene family in plants. *Journal Of Plant Physiology* **163**, 287–296 (2006).
- Hartung, F., Suer, S., Puchta, H. Two closely related *RecQ*-helicases have antagonistic roles in homologous recombination and DNA repair in *Arabidopsis thaliana*. *Proc Natl Acad Sci Usa* **104**, 18836–841 (2007).
- Hartung, F., Suer, S., Puchta, H. The Role of *AtMUS81* in DNA Repair and its Genetic Interaction with the Helicase *AtRecQ4A*. *Nucleic Acids Res.* **34**, 4438–48 (2006).
- Watt, P.M., Louis, E.J., Borts, R.H., Hickson, I.D. Sgs1: a eukaryotic homolog of *E. coli* *RecQ* that interacts with topoisomerase II *in vivo* and is required for faithful chromosome segregation. *Cell* **81**, 253–60 (1995).
- Langlois, R.G., Bigbee, W.L., Jensen, R.H., German, J. Evidence for increased *in vivo* mutation and somatic recombination in Bloom's syndrome. *Proc Natl Acad Sci Usa* **86**, 670–4 (1989).
- Mankouri, H.W. And Hickson, I.D. The *RecQ* helicase-topoisomerase III-Rmi1 complex: a DNA structure-specific 'dissolvosome'? *Trends Biochem Sci.* **32**, 538–46 (2007).
- Kobbe, D., Blanck, S., Demand, K., Focke, M., Puchta, H. *AtRECQ2*, a *RecQ*-helicase homologue from *Arabidopsis thaliana*, is able to disrupt different recombinogenic DNA-structures *in vitro*. *Plant J.* Apr 12 [Epub Ahead Of Print] (2008).
- Hartung, F., Plchova, H., Puchta, H. Molecular characterisation of *RecQ* homologues in *Arabidopsis thaliana*. *Nucleic Acids Res.* **28**, 4275–82 (2000).
- Plchova, H., Hartung, F., Puchta, H. Biochemical characterization of an exonuclease from *Arabidopsis thaliana* reveals similarities to the DNA exonuclease of the human Werner syndrome protein. *J. Biol. Chem.* **45**, 44128–138 (2003).
- Macris, M.A., Krejci, L., Bussen, W., Shimamoto, A., Sung, P. Biochemical characterization of the RECQ4 protein, mutated in Rothmund-Thomson syndrome. *DNA Repair (Amst).* **5**, 172–80 (2006).
- Shen, J.C., Gray, M.D., Oshima, J., Loeb, L.A. Characterization of Werner syndrome protein DNA helicase activity: directionality, substrate dependence and stimulation by replication protein A. *Nucleic Acids Res.* **26**, 2879–85 (1998).
- Mohaghegh, P. *et al.* The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases. *Nucleic Acids Res.* **34**, 2269–79 (2001).
- Orren, D.K., Theodore, S., Machwe, A. The Werner syndrome helicase/exonuclease (WRN) disrupts and degrades D-loops *in vitro*. *Biochemistry* **41**, 13483–8 (2002).