J Hum Genet (2003) 48:107–109

SHORT COMMUNICATION

Gaia Roversi · Alessandro Beghini · Giovanna Zambruno
Mauro Paradisi · Lidia Larizza

Identification of two novel *RECQL4* exonic SNPs and genomic characterization of the IVS12 minisatellite

Received: October 24, 2002 / Accepted: November 21, 2002

Abstract Rothmund-Thomson syndrome is a rare autosomal recessive disorder characterized by a widely heterogeneous clinical presentation. Only a subset of clinically diagnosed patients carry *RECQL4* gene mutations, probably because of their genetic heterogeneity and/or the complexity of molecular testing. We here describe the polymorphic sites of the *RECQL4* gene that detail its genomic structure and may be of interest as modulators of the splicing process and gene expression. We characterized two novel and one already described single-nucleotide polymorphism in the coding region of the *RECQL4* gene, which were shown by the exonic splicing enhancer (ESE) score matrix to fall into high-score motifs recognized by serine/arginine-rich proteins. We also describe the genomic structure of a G-C rich minisatellite flanking the 3′ splice site of IVS12 in the helicase domain of the *RECQL4* gene, which may enhance mutations such as those described at the IVS12 acceptor site. *RECQL4* polymorphic sites may be useful for identifying alleles associated with missplicing and, more generally, in cancer-susceptibility association studies.

Key words *RECQL4* gene · Rothmund-Thomson syndrome · SNPs · ESE · Minisatellite

Introduction

RECQL4 is a human DNA helicase gene belonging to the *RECQ* helicase gene family including the Werner (*RECQ3*) and Bloom (*RECQ2*) syndrome genes (Kitao et al. 1999a).

RECQL4 gene mutations have been identified in a subset of patients with Rothmund-Thomson syndrome (RTS, MIM#268400; Kitao et al. 1999b; Lindor et al. 2000; Wang et al. 2002; Balraj et al. 2002), a rare autosomal recessive genodermatosis, which, in addition to the poikilodermatous rash, has a variable clinical presentation and is associated with genomic instability and predisposition to malignancy (Wang et al. 2001).

The *RECQL4* gene has a peculiar genomic structure: most of the introns are less than 100 bp in length, and are thus predisposed to inefficient splicing (Berget 1995), even in the absence of classical splicing mutations (Wang et al. 2002; Beghini et al. 2003). A thorough analysis of all the elements detailing its gene stucture at both the genomic and the RNA level is therefore warranted to discover its RNA processing and to address genotype–phenotype correlations in RTS patients.

During *RECQL4* gene mutation screenings of patients with a clinical diagnosis of RTS, we identified in the *RECQL4* coding region two novel and one already described (Balraj et al. 2002) single-nucleotide polymorphism (SNP) falling into predicted functional exonic splicing enhancers (ESEs). ESEs are *cis*-acting elements that function as binding sites for serine/arginine-rich (SR) proteins, a family of essential splicing factors that are also involved in alternative splicing regulation (Cartegni et al. 2002).

We also characterized the genomic structure of an intronic G-C rich minisatellite (Kitao et al. 1999a) abutting the acceptor splice site.

Subjects and methods

Peripheral blood samples were obtained from 50 healthy Italian subjects and 2 Italian patients with a clinical diagnosis of RTS but no *RECQL4* gene mutations. The genomic DNAs were extracted according to standard protocols. Informed consent was obtained for all studies.

The SNPs were genotyped by means of sequencing polymerase chain reaction (PCR)-amplified DNA fragments. The positions of the used primers are shown in Fig. 1a. The fragments containing SNPs 1 and 2 were amplified using

G. Roversi · A. Beghini · L. Larizza (✉)
Department of Biology and Genetics for Medical Sciences, University of Milan, Via Viotti3/5, 20133 Milan, Italy
Tel. +39-02-503-15859; Fax +39-02-503-15864
e-mail: lidia.larizza@unimi.it

G. Zambruno · M. Paradisi
Istituto Dermopatico dell'Immacolata, IDI-IRCSS, Rome, Italy

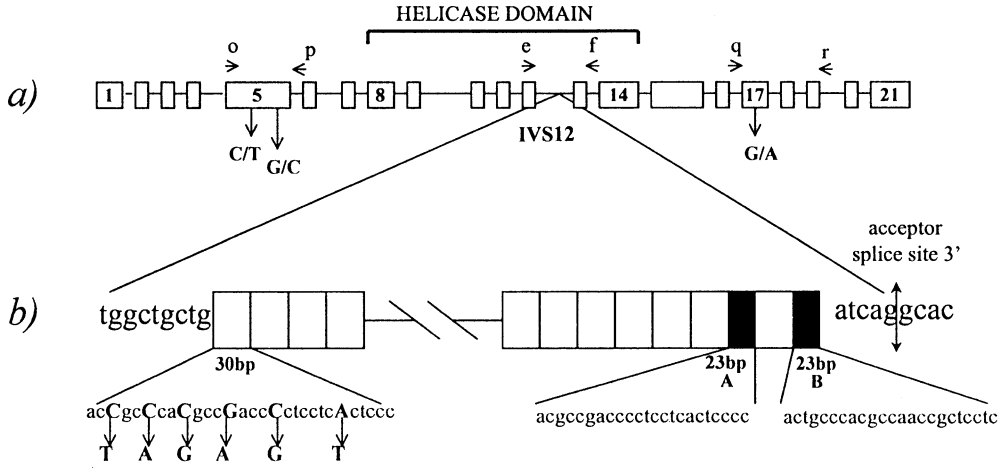

Fig. 1. a Schematic representation of *RECQL4* genomic structure: the positions of the identified single-nucleotide polymorphisms are indicated by *arrows*. Exons are indicated by *boxes* and introns by *thin lines*. **b** Organization of the GC-rich minisatellite in IVS12. The *empty* and *full boxes* indicate repeats of 30 and 23 bp (A and B type) and show their sequences and variations (*uppercase letters*). The *double-arrowed line* points to the IVS12 acceptor splice site, which is only 5 bp from the last minisatellite repeat

primers O (5′gcagaaaaagtcagtgatgagc3′) and P (5′tgggc gggaaatacgggagg3′) under the following conditions: one cycle at 94°C for 4 min; followed by 33 cycles at 94°C for 40 s, 62°C for 35 s, and 72°C for 40 s; and a final extension at 72°C for 3 min. For SNP 3, primers Q (5′cgaccacc tatacccattgc3′) and R (5′gccttactacactcactctg3′) were used under the following conditions: one cycle at 94°C for 4 min; followed by 33 cycles at 94°C for 40 s, 58°C for 35 s, and 72°C for 40 s; and a final extension at 72°C for 3 min. All the PCR products were purified and sequenced using the ABI Prism Big Dye Terminator Cycle Sequencing Ready Reaction Kit and an ABI Prism 3100 sequence analyzer (Applied Biosystems, Warrington, UK).

We analyzed the flanking sequences of each SNP using sequence-motif matrices that predict the functional ESEs recognized by the SR proteins SF2/ASF, SC35, SRp40, and SRp55 (http:/exon.cshl.edu/).

For the minisatellite analysis, we amplified the genomic region corresponding to IVS12 in three normal control subjects and two RTS patients using primers E (5′gcctcacaccac tgccgcctctgg3′) and F (5′gacaggcagatggtcagtgggatg3′) under the following conditions: one cycle at 94°C for 4 min; followed by 33 cycles at 94°C for 30 s, 68°C for 25 s, and 72°C for 30 s; and a final extension at 72°C for 7 min. The molecular weights of the PCR products were electrophoretically assessed on 2% agarose gel, and the fragments corresponding to each allele were extracted, purified, and sequenced as described earlier.

Results and discussion

RECQL4 exonic SNPs

Figure 1a shows the genomic structure of the *RECQL4* gene and the positions of the three SNPs identified in exons 5 and 17. The DNA and protein variation data are given in the upper part of Table 1. The allelic frequencies were obtained by genotyping a panel of 50 subjects from the Italian population. The heterozygosity index was 0.42 for SNPs 1 and 3, and 0.44 for SNP 2.

The lower part of Table 1 shows the results obtained by applying sequence-motif matrices predicting the functional ESEs recognized by the SR proteins SF2/ASF, SC35, SRp40, and SRp55 to the flanking sequences of each SNP (http:/exon.cshl.edu/), including the motif scores of each of the recognized SNP motifs.

As can be seen, both the SF2/ASF heptamer and the SRp55 hexamer high-score motifs are disrupted by the C738T variation (SNP 1), the G801C substitution at SNP 2 introduces new binding sites for SC35 and SRp40, and the G3014A of SNP 3 destroys the heptamer high-score motif for SF2/ASF.

All the SNP alleles were highly frequent in the population, thus predicting no likely haplotype linkage with the RTS phenotype. The genotyping of two RTS patients without *RECQL4* gene mutations revealed haplotypes as different as those observed in normal controls.

RECQL4 IVS12 minisatellite

Figure 1b shows the structure and sequences of the 30- and 23-bp monomers of the G-C rich minisatellite located in the 3′ end portion of IVS12. This minisatellite consists of imperfect tandem repeats of a 30-mer unit. The repeat copy number is polymorphic and, on the basis of molecular weight assessments in a sample of five subjects, should range from 17 to 20 units. The interspersion pattern of variant repeats differing in one or more base substitutions (uppercase letters in Fig. 1b) or in the loss of the first (23 A) or the last seven nucleotides (23 B) was also highly variable among alleles. Curiously, the first 5′ unit (30 mer) and the last three 3′ units (23A-30-23B mer), which respectively define the variable-number tandem repeats start and end were conserved in the ten analyzed chromosomes, whereas the internal region was variable in terms of base substitutions, but not in terms of unit length. There was no significant difference in the number of tandem repeats between RTS patients and controls.

Because the minisatellite flanks the 3′ splice site of intron 12, it may enhance mutation at the IVS12 3′ splice site: the finding of two different changes among the few

Table 1. *RECQL4* SNPs and allele variation in SR protein score motifs

	SNP 1		SNP 2		SNP 3	
Exon	5		5		17	
cDNA position	738		801		3014	
Sequence	TCAG(C/T)ATCCGTG		AGGA(G/C)CCCTGG		CGGC(G/A)GGCT	
aa variation	ser/ser (silent)		glu/asp (conservative)		arg/gln (not conservative)	
Frequencies[a]	C = 0.55	T = 0.45	G = 0.6	C = 0.4	G = 0.55	A = 0.45
ESE FINDER						
SF2/ASF (1.95[b])						
Motif	**C**atccgt	<1.95	<1.95	<1.95	cggcGgg	<1.95
Score	2.93				3.57	
SC35 (2.383[b])						
Motif	**C**atccgtg	**T**atccgtg	ga**G**ccctg	ga**C**ccctg gga**C**ccct	<2.383	<2.383
Score	2.5	2.5	3.78	6.09	2.5	
SRp40 (2.67[b])						
Motif	<2.67	<2.67	<2.67	**C**ccctgg	<2.67	<2.67
Score				3.32		
SRp55 (2.67[b])						
Motif	Ag**C**atc	<2.67	<2.67	<2.67	<2.67	<2.67
Score	3.64					

SNP, Single-nucleotide polymorphism; aa, amino acid; SR, serine/arginine-rich
[a] estimated from the genotyping of 100 chromosomes
[b] Threshold value for each SR protein

identified mutations in RTS patients (Kitao et al. 1999b; Beghini et al. 2003) is consistent with this hypothesis.

Putative role of *RECQL4* polymorphisms in missplicing

Both the minisatellite and the SNPs belonging to high-score motifs for SR proteins may play a role in *RECQL4* mRNA missplicing, which also seems to be common in wild-type subjects (Wang et al. 2002; Beghini et al. 2003).

Furthermore, because any cSNP important for splicing modulation may subtly influence the efficiency and tissue- or stage-specificity of alternative splicing (Cartegni et al. 2002), the identified *RECQL4* SNPs may contribute to the phenotype variability of RTS syndrome by cooperating with mutations located elsewhere in the *RECQL4* gene or in other loci.

In addition to their putative involvement in the splicing process, the three exonic SNPs could be useful for expression studies aimed at identifying the alleles associated with missplicing and/or the underlying sequence elements modulating splicing.

It is well known that DNA helicase deficiencies are associated with a predisposition for cancer, and that *RECQL4* helicases interact with the proteins required for chromosome maintenance (Mohaghegh and Hickson 2001). Intragenic *RECQL4* gene polymorphisms could be very useful in loss-of-heterozygosity studies of sporadic genomically unstable tumors for which defects in DNA synthesis, recombination, and repair can be hypothesized.

Acknowledgments This study was supported by the AIRC (Associazione Italiana per la Ricerca sul Cancro): 2001 grant to L.L.

References

Balraj P, Concannon P, Jamal R, Beghini A, Hoe TS, Khoo AS, Volpi L (2002) An unusual mutation in RECQ4 gene leading to Rothmund-Thomson syndrome. Mutat Res 508:99–105

Beghini A, Castorina P, Roversi G, Modiano P, Larizza L (2003) RNA processing defects of the helicase gene RECQ4 in a compound heterozygous Rothmund-Thomson patient. Am J Med Genet (in press)

Berget SM (1995) Exon recognition in vertebrate splicing. J Biol Chem 270:2411–2414

Cartegni L, Chew SL, Krainer AR (2002) Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat Rev Genet 3:285–298

Kitao S, Lindor NM, Shiratori M, Furuichi Y, Shimamoto A (1999a) Rothmund-Thomson syndrome responsible gene, *RECQL4*: genomic structure and products. Genomics 61:268–276

Kitao S, Shimamoto A, Makoto G, Miller RW, Smithson WA, Lindor NM, Furuichi Y (1999b) Mutations in *RECQL4* cause a subset of cases of Rothmund-Thomson syndrome. Nat Genet 22:82–84

Lindor NM, Furuichi Y, Kitao S, Shimamoto A, Arndt A, Jalal S (2000) Rothmund-Thomson syndrome due to *RECQ4* helicase mutations: report and clinical and molecular comparisons with Bloom syndrome and Werner syndrome. Am J Med Genet 90:223–228

Mohaghegh P, Hickson ID (2001) DNA helicase deficiencies associated with cancer predisposition and premature ageing disorders. Hum Mol Genet 10:741–746

Wang LL, Levy ML, Lewis RA, Chintagumpala MM, Lev D, Rogers M, Plon SE (2001) Clinical manifestations in a cohort of 41 Rothmund-Thomson syndrome patients. Am J Med Genet 102:11–17

Wang LL, Worley K, Gannavarapu A, Chintagumpala MM, Levy ML, Plon SE (2002) Intron-size constraint as mutational mechanism in Rothmund-Thomson syndrome. Am J Hum Genet 71:165–167