Assignment to chromosome 11 of mouse p68 RNA helicase gene (*Hlr1*) and pseudogene (*Hlr1-ps1*)

Patrick Petry[a], Kenneth R. Johnson[b], Lydia Lemaire[a,**], Uwe A.O. Heinlein[a,*]

[a]*Institut für Genetik, Heinrich-Heine-Universität, Universitätsstraße 1, D-40225 Düsseldorf, Germany*
[b]*The Jackson Laboratory, Bar Harbor, ME 04609, USA*

Received 16 February 1995

Abstract The gene encoding murine p68 RNA helicase (*Hlr1*) was mapped to the distal portion of mouse chromosome 11 by linkage analysis of DNA restriction length polymorphisms using an interspecific genetic backcross between (C57BL/6J × SPRET/Ei) F1 hybrids and SPRET/Ei mice. A closely related gene (*Hlr1-ps1*) was identified, isolated, and mapped to the proximal part of the same chromosome. Sequence analysis and PCR results suggest that *Hlr1-ps1* is a pseudogene, flanked by DNA stretches similar to mouse insertion element IE118.

Key words: Genetic backcross; IE118; DEAD box protein

1. Introduction

The 'DEAD box family' of proteins includes mainly RNA-dependent ATPases involved in several cellular functions. Examples are the translation initiation factor eIF-4A, *Dea*D (*E. coli*), splicing factors PRP5 and PRP28 (yeast), An3 (*Xenopus laevis*), vasa (*Drosophila*), and p68 RNA helicase (see [1] for a review). The latter is highly conserved between man and mouse [2], but apart from its ATP-dependent capability of unwinding RNA, its biological function is unknown. The mouse gene is preferantially expressed in male germ cells from diplotene stages on [2], suggesting that p68 RNA helicase activity is required for proper spermatogenesis. The striking sequence homology between mouse and man (>95%) implies that the cellular function of p68 is also conserved and may represent a key component of the spermatogenic process. To obtain more information on the degree of similarity, we analyzed genomic organization features of the mouse p68 RNA helicase gene, *Hlr1*. In contrast to the human, the mouse genome contains two genomic loci hybridizing with the respective cDNA probes. This report presents data on the chromosomal assignment of these two loci, and evidence that one of these two represents a pseudogene (*Hlr1-ps1*) flanked by mouse insertion element sequences.

2. Experimental

2.1. Cloning procedures

All cloning, sequencing and computer analysis was performed according to standard procedures [3] or as described [2,4]. Genomic phage containing *Hlr1-ps1* sequences were isolated by probing an EMBL3 library of C57BL/6J genomic mouse DNA [5] with a 0.6 kb *Xho*I/*Eco*RI fragment of mouse p68 RNA helicase cDNA clone TNZ2 [2]. A partial phage library was generated by *Eco*RI digestion of genomic C57BL/6J DNA, isolation of fragments between 3.8 and 5.0 kb by preparative gel electrophoresis, and subsequent cloning of the isolated fragments in λgt10 phage vector.

2.2. PCR analysis

PCR amplification of gene and pseudogene sequences was performed essentially as described [6], using a synthetic DNA primer pair, 5′-GTGGAATCTTGATGAGCTGCCC-3′ (sense) and 5′-GCAAATAGGCCCATCACCTC-3′ (antisense), corresponding to nucleotides 295–316 and 670–651, respectively, of p68 RNA helicase cDNA [2].

2.3. Genetic mapping

To genetically map the *Hlr-1* and *Hlr1-ps1* genes, we used a previously described panel of DNAs from an interspecific backcross [7]. Female F1 hybrids between an inbred strain of *Mus spretus* (SPRET/Ei) and the standard laboratory strain C57BL/6J were backcrossed to male SPRET/Ei mice. This backcross panel has already been typed for more than 250 polymorphic loci distributed over all 19 autosomes and the X chromosome (K.R.J., unpublished). Using a 700 bp fragment of the p68 RNA helicase cDNA [2] as a probe, we detected *Eco*RI RFLPs between C57BL/6J and SPRET/Ei DNA; these RFLPs were used to map the *Hlr1* and *Hlr1-ps1* loci (Figs. 1 and 2). The computer program Map Manager [8] was used to perform linkage and haplotype analysis. Genetic linkage was determined by segregation analysis of the backcross progeny; the distributions of *Hlr1* and *Hlr1-ps1* genotypes were compared to genotype distributions of more than 250 prepositioned reference loci dispersed throughout the mouse genome. Gene order on a chromosome was determined by minimizing the number of double crossover events required to explain observed haplotype distributions.

3. Results and discussion

3.1. Genetic mapping

We detected linkage between p68 RNA helicase-hybridizing DNA fragments and markers on chromosome 11. Furthermore, the two diagnostic C57BL/6J-specific *Eco*RI fragments (9.0 kb and 4.5 kb, Fig. 1) did not co-segregate among the interspecific backcross mice and therefore represent separate loci. Both loci mapped to Chr 11; however, the 9.0 kb fragment mapped between proximal markers *Hba* and *Il-4*, whereas the 4.5 kb fragment mapped between distal markers *Xmv-42* and *Es-3* (Fig. 2). This distal region of mouse Chr 11 shows conserved linkage relationships with the syntenic human region on Chr 17, namely at 17q23-q25, to where the human p68 RNA helicase gene (*HLR1*) has been assigned [9]. It was therefore most likely that the 4.5 kb fragment represented the functional mouse gene, *Hlr1*. This assumption was proven by Southern blot and PCR data on genomic phage clones which showed that (i) the diagnostic 4.5 kb fragment contained a normal exon/intron structure in contrast to DNA with pseudogene sequences (Fig. 3); and (ii) pseudogene phage contained the diagnostic 9.0 kb fragment and an additional 1.5 kb fragment. The sequence of this 1.5 kb fragment was similar to the p68 RNA

*Corresponding author. Fax: (49)(211) 311 2279.

***Present address:* Max-Planck-Institut für Biophysikalische Chemie, D-37077 Göttingen, Germany.

SSDI 0014-5793(95)00270-7

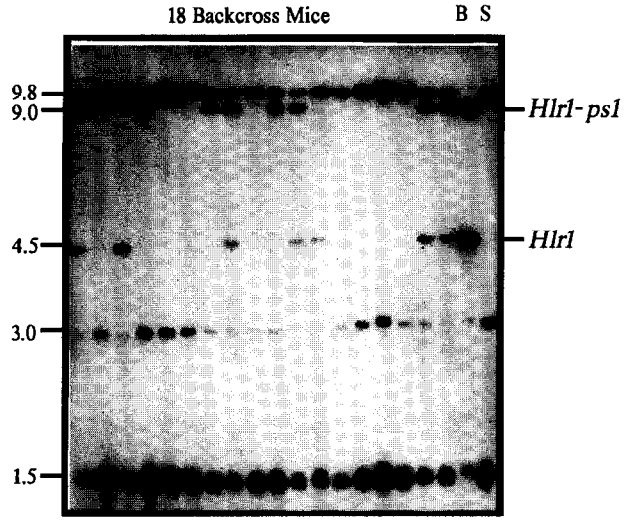

Fig. 1. Segregation among backcross progeny of *Eco*RI restriction fragments containing *Hlrl*-related sequences. The *Eco*RI RFLPs of the parental strains, C57BL/6J (B) and SPRET/Ei (S) are shown in the last two lanes. DNA fragment sizes (in kb) are shown on the left. Genotypes for *Hlrl* and *Hlrl-psl* were scored by the presence or absence of the C57BL/6J-derived fragments indicated on the right. These Southern blots were hybridized with a 0.7 kb 5′-portion of the murine p68 RNA helicase cDNA, *Eco*RI to *Xho*I [2].

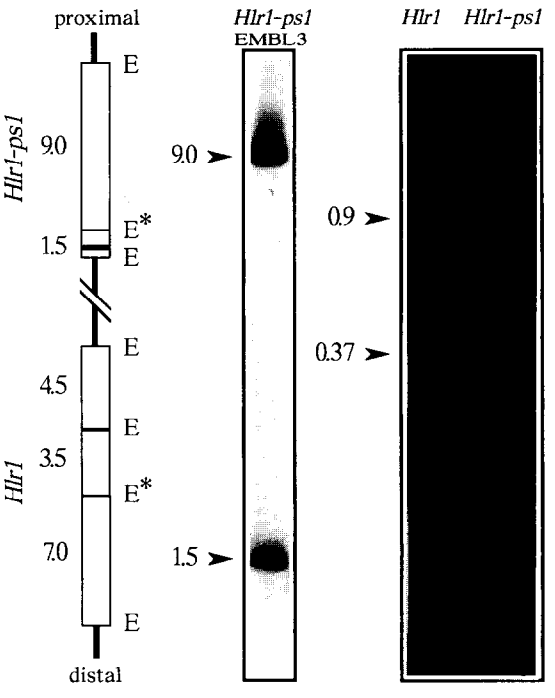

Fig. 3. *Left*: genomic arrangement of p68 RNA helicase gene and pseudogene on Chr 11. *Eco*RI restriction fragment lengths are given in kb. The black horizontal bar within the 1.5 kb pseudogene fragment indicates the position of IE118 insertion element sequences (see Fig. 4). Asterisks denote two homologous *Eco*RI sites corresponding to the *Eco*RI site in p68 RNA helicase cDNA at position 1318 (see [2] for details). *Middle*: Southern blot analysis of genomic EMBL3 phage carrying the *Hlrl-psl* pseudogene, indicating that the diagnostic 9.0 kb band used for the backcross analysis (see Fig. 1) corresponds to the pseudogene. *Right*: comparison of PCR fragments obtained from phage containing *Hlrl* and *Hlrl-psl*, respectively. Primers are located in the 5′-region of the gene (see section 2). A fragment size of 376 bp corresponds to the intron-free, cDNA-like arrangement (lane *Hlrl-psl*). The active gene contains three introns (363 bp, 87 bp, and 93 bp; P.P., U.A.O.H., unpubl.) between the two PCR primers, resulting in an increased fragment size of 909 bp (lane *Hlrl*).

helicase cDNA sequence, but contained frame shift mutations, thus being indicative for an inactive pseudogene (see Fig. 4 and below). Fluorescent in situ hybridization (FISH) analysis (not shown) corroborated the genetic mapping results by placing *Hlrl-psl* to cytogenetic band 11B3 and *Hlrl* to 11E2 (see Fig. 5). Until now, we have not been able to separate the loci for *Hlrl* and *TAZ4* (*Tex2*) [6], neither by genetic mapping nor by FISH analysis. It remains to be shown by 'extended fibre FISH' and pulse field gel electrophoresis how close these two testis-expressed genes are on the distal portion of Chr 11.

3.2. Hlrl-psl *pseudogene analysis*

Using the entire 2.3 kb mouse p68 RNA helicase cDNA [2] as a probe on genomic Southern blots with C57BL/6J-DNA, we detected five bands representing two pairs of restriction fragments from two genomic loci (see Fig. 3). These two loci include one homologous *Eco*RI site, which is present in the p68 RNA helicase cDNA at nucleotide 1316 [2]. Southern blot analysis with a 3′-cDNA probe revealed that two fragments of 7.0 kb and 1.5 kb contained the respective 3′-portions of the related genes. To obtain sequence information, we isolated

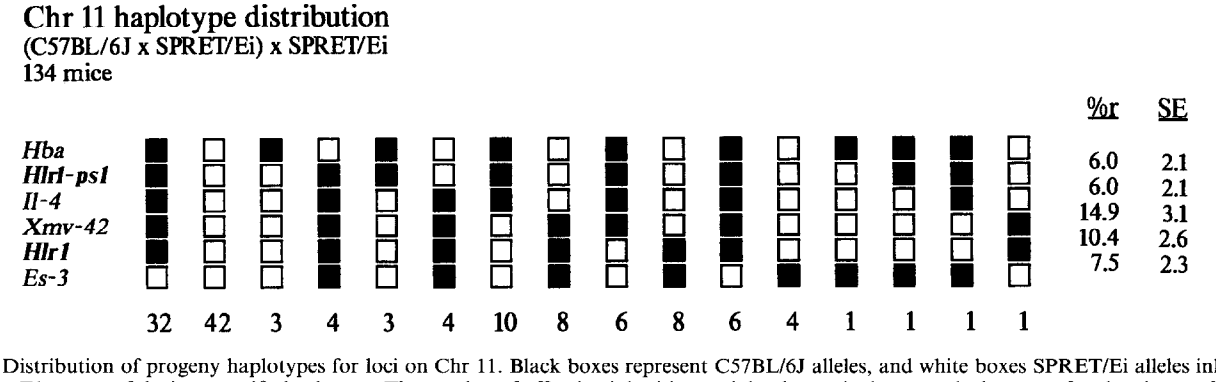

Fig. 2. Distribution of progeny haplotypes for loci on Chr 11. Black boxes represent C57BL/6J alleles, and white boxes SPRET/Ei alleles inherited from the F1 parent of the interspecific backcross. The number of offspring inheriting each haplotype is shown at the bottom of each column of boxes. Loci are listed from centromere to distal telomere. Percent recombination (%r) and standard error (SE) values between pairs of adjacent loci are shown on the right.

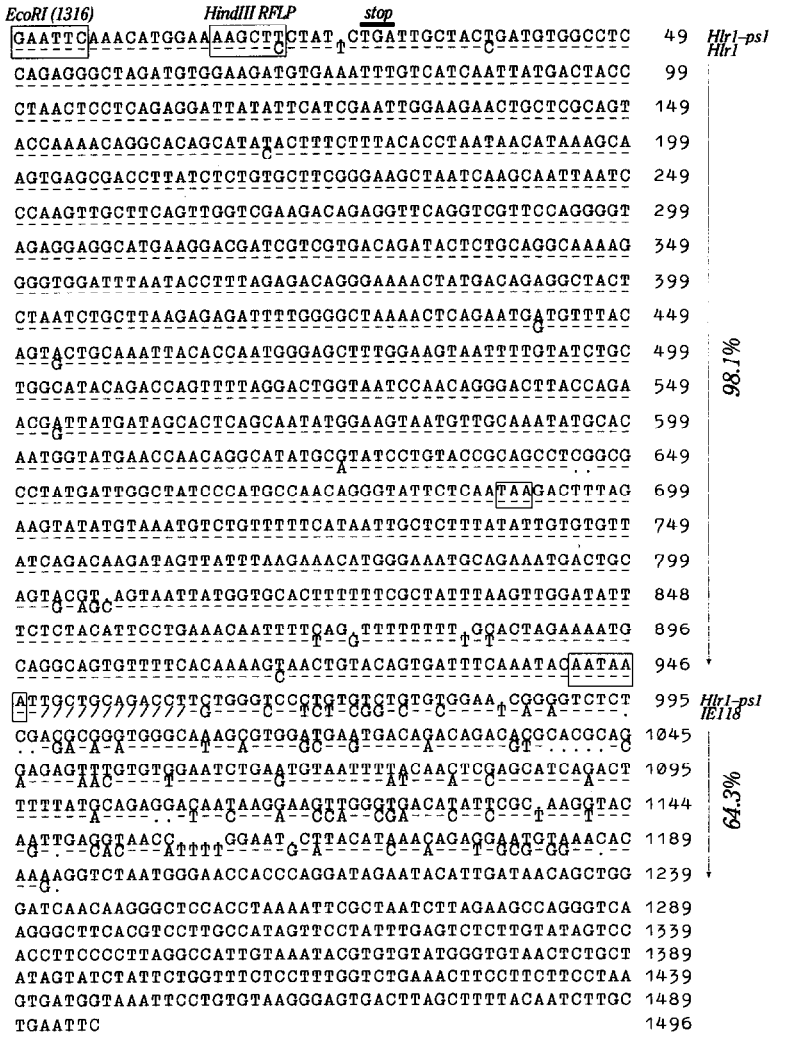

Fig. 4. Sequence analysis of the 1.5 kb fragment containing the 3′-portion of *Hlr1-ps1* pseudogene. The upper lines represent the sequence of the 1.5 kb pseudogene *Eco*RI fragment (see Fig. 3). The lower lines represent the p68 RNA helicase cDNA sequence [2] and the IE118 insertion element sequence [10], which are aligned to the pseudogene sequence (cDNA: nt 1–947; IE118: nt 960–1190). Identical nucleotides are indicated by a hyphens, differing nucleotides are printed. Hatched region indicates 12 nucleotides not assignable to either sequence. Deletions/gaps are marked by full stops. The similarity between these sequences is given on the right. The *Eco*RI site present at nt 1,316 of p68 helicase cDNA [2], a *Hin*dIII RFLP, which was used for further corroboration (not shown), as well as the stop codon and poly(A) signal within the *Hlr1*-derived cDNA [2] are indicated by boxes. The premature TGA stop codon introduced by a frame shift mutation in *Hlr1-ps1* is marked by a horizontal bar. The EMBL data library Accession Number of this pseudogene fragment sequence is X84259.

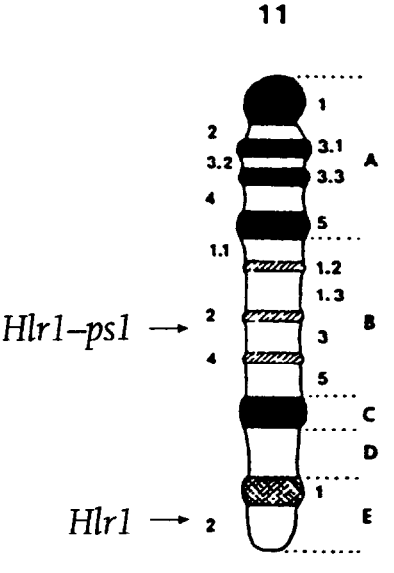

Fig. 5. Cytogenetic map position of *Hlr1* and *Hlr1-ps1* on a Chr 11 GIEMSA idiogram [11]. Positions were determined by fluorescent-in situ-hybridization (FISH) (not shown) as described [6].

genomic phage from a C57BL/6J library in λEMBL3 [5], and used those containing an internal, positive 1.5 kb fragment for subcloning into pBluescript KS. Sequence analysis (Fig. 4) showed that the sequences between the functional cDNA and the pseudogene differ in 17 nucleotides between the *Eco*RI site and the polyadenylation signal, corresponding to a similarity of 98.1%. Adjacent to the AATAAA motif, 239 nucleotides were found to be 64.3% similar to a retrovirus-LTR-like mouse

insertion element, IE118 [10]. Pseudogene and insertion element sequences were connected by 12 nucleotides which could not be assigned to either of these blocks. These data support the idea that the p68 RNA helicase-related sequence on the central region of Chr 11 represents a processed and inactive pseudogene, based on the following considerations: Deletion of a T at nt 27 introduces a stop codon at nt 29, so that production of a complete protein is not possible; there is no intron between nt 60 and 61 where a conserved intron exists in human and yeast [9]; there is no obvious exon/intron border after the AATAAA, excluding the possibility that the insertion element and following sequences are introning rather than flanking genomic DNA. To date, there are no hints whether the integration event leading to the pseudogene was aided by the insertion element, or whether it was hit by chance.

Acknowledgements: We are indebted to Rita Lange for partial sequence analysis. This work was supported by grants from the Deutsche Forschungsgemeinschaft (He1581/1–3 and 4–1 to U.A.O.H.) and the National Institutes of Health (GM46697 to K.R.J.). L.L. received a travel grant from the Deutscher Akademischer Austauschdienst (DAAD). We wish to thank Frank H. Ruddle and David C. Ward for their hospitality and generosity.

References

[1] Fuller-Pace, F.V. (1994) Trends Cell Biol. 4, 271–274.
[2] Lemaire, L. and Heinlein, U.A.O. (1993) Life Sci. 52, 917–926.
[3] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
[4] Dommers, S. and Heinlein, U.A.O. (1994) Biochem. Biophys. Res. Commun. 204, 1346–1351.
[5] Heinlein, U.A.O., Lange-Sablitzky, R., Schaal, H. and Wille, W. (1986) Nucl. Acids Res. 14, 6403–6416.
[6] Lemaire, L., Johnson, K.R., Bammer, S., Petry, P., Ruddle, F.H. and Heinlein, U.A.O. (1994) Genomics 21, 409–414.
[7] Johnson, K.R., Cook, S.A. and Davisson, M.T. (1992) Genomics 12, 503–509.
[8] Manly, K.F. (1993) Mammalian Genome 4, 303–313.
[9] Iggo, R., Gough, A., Xu, W., Lane, D.P. and Spurr, N.K. (1989) Proc. Natl. Acad. Sci. USA 86, 6211–6214.
[10] Man, Y.M., Delius, H. and Leader, D.P. (1987) Nucl. Acids Res. 15, 3291–3304.
[11] Evans, E.P. (1989) In: Genetic Variants and Strains of the Laboratory Mouse (Lyon, M.F. and Searle, A.G., Eds.), 2nd Edn., pp. 576–577, Oxford University Press, Oxford, UK.