Vol. 203, No. 1, 1994 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
August 30, 1994 Pages 594-601

STRUCTURE AND CHROMOSOMAL LOCUS OF THE MOUSE GENE ENCODING A CEREBELLAR PURKINJE CELL-SPECIFIC HELIX-LOOP-HELIX FACTOR HES-3

Teiji Sakagami[1], Kazuhiro Sakurada[1+], Yuka Sakai[2], Tomomasa Watanabe[2*], Shigetada Nakanishi[1], and Ryoichiro Kageyama[1#]

[1]Institute for Immunology, Kyoto University Faculty of Medicine, Sakyo-ku, Kyoto 606, Japan

[2]Department of Biochemistry, Institute for Developmental Research, Aichi Prefecture Colony, Kasugai, Aichi 480-03, Japan

Received July 15, 1994

Summary: HES-3 is a cerebellar Purkinje cell-specific helix-loop-helix factor structurally related to the products of the *Drosophila hairy* and *Enhancer of split* genes. Here, we report the nucleotide sequence and chromosomal locus of the mouse *Hes-3* gene. This gene consists of four exons and the exon-intron boundaries are well conserved when compared with those of the mouse *Hes-1* and *Drosophila hairy* genes. Southern blot and interspecies backcross analyses show that the mouse *Hes-3* gene is a single-copy gene and is located around position 80 on chromosome 4. Further analysis indicates that this locus is close to the *Hes-5* locus, which is different from the *Hes-1* locus (position 26 on chromosome 16). These results suggest that the *Hes-3* and *Hes-5* genes may be clustered on chromosome 4 while the *Hes-1* gene is not.

Mammalian HES family consists of at least five helix-loop-helix (HLH) factors structurally related to the products of the *Drosophila hairy* and *Enhancer of split* [*E(spl)*] genes (1-5), which regulate neurogenesis (6-10). Among the members of the HES family, HES-3 has a unique feature in transcriptional activity. Unlike the other members of the family, HES-3 does not seem to have a complete basic region at the amino-terminal end of the HLH domain (1). As a result, HES-3 does not directly interact with the DNA template, but it can prevent other basic HLH (bHLH) factors from binding to the DNA (the E box) by forming nonfunctional heterodimers, like Id proteins (1,11-13). Furthermore, HES-3 shows a unique expression patterns. HES-3 expression occurs specifically in cerebellar Purkinje cells, large arborized neurons providing the only known output from the cerebellum (1). HES-3 mRNA appears by postnatal day 14, gradually increases, and reaches the maximal level by 1-2 months after birth, coinciding

+Present address: Tokyo Research Laboratories, Kyowa Hakko Kogyo Co., Machida 194, Japan.
*Present address: Faculty of Veterinary Medicine, Hokkaido University, Sapporo 060, Japan.
#To whom corresponding should be addressed. Fax: 81-75-753-4404.

0006-291X/94 $5.00

with maturation of Purkinje cells (1,14). Thus, HES-3 may be involved in the maturation and functional maintenance of Purkinje cells rather than their fate determination.

In this study, to extend the molecular analysis of HES-3 further, we characterized the structure and chromosomal locus of the mouse *Hes-3* gene. We found that structural features, such as the exon-intron boundaries, of the *Hes-3* gene are very similar to those of *Hes-1* gene (4). However, interspecies backcross analysis suggests that *Hes-3* and *Hes-1* genes are located on different chromosomes.

EXPERIMENTAL PROCEDURES

Isolation of the mouse HES-3 gene - The mouse genomic library (Stratagene) was screened by hybridization with the ~1-kilobase (kb) EcoRI fragment of the rat HES-3 cDNA as a probe (1). Four clones were isolated from 1 x 10^6 plaques. The fragments hybridized positively were subcloned into pBluescript and subjected to sequence analysis.

Southern blot analysis - Mouse liver DNA digested by restriction enzymes was electrophoresed on 0.7% agarose gel and transferred to a nylon membrane filter. The ^{32}P-labeled ~0.7-kb EcoRI-PstI fragment of the rat cDNA was hybridized to the DNA at 65°C in a solution containing 0.75 M NaCl, 0.075 M sodium citrate, and 0.5% SDS.

Interspecies backcross analysis - NOD, BALB/c, and Japanese wild mice (MOL-MIT) were used for interspecies backcross analyses as described previously (15). More than 30 markers were analyzed to search for the locus closely linked to the *Hes-3* and *Hes-5* loci. The ~1-kb rat HES-3 and ~1.3-kb rat HES-5 cDNAs were used as a probe.

RESULTS

Structural Organization of Mouse Hes-3 Gene - Four genomic clones containing the mouse *Hes-3* gene were isolated by screening 1 x 10^6 plaques of a mouse genomic library with the rat HES-3 cDNA probe, and the region hybridized by the probe was sequenced. As shown in Fig. 1, the mouse *Hes-3* gene consisted of four exons, and the exon-intron boundaries possessed the consensus splicing signal conforming to the GT-AG rule, except that the donor site of the first intron was CT instead of GT. The deduced amino acid sequence of mouse HES-3 was very similar to that of rat HES-3. They showed 96% identity in the bHLH domain and 86% identity in the whole coding region.

Comparison between the mouse *Hes-3* and *Hes-1* and *Drosophila hairy* genes showed that they had similar genomic organization, except that the *Drosophila hairy* gene had only three exons (Fig. 2A). All of the introns of the *Hes-3*, *Hes-1*, and *Drosophila hairy* genes interrupted the equivalent positions of the protein-coding regions (Fig. 2B). These results suggest that these genes originated from the same ancestral gene.

On Southern blot analysis, only a single band hybridized with the HES-3 cDNA was detected when digested by EcoRI, HindIII, or BamHI, suggesting that the *Hes-3* gene is a single-copy gene (Fig. 3).

Interspecies Backcross Analysis of Assignment of Chromosomal Locus - To determine the chromosomal locus of the mouse *Hes-3* gene, we carried out interspecies backcross analysis using two mating combinations among three mice strains, NOD and MOL-MIT,

Fig. 1. Nucleotide sequence of mouse *Hes-3* gene. The nucleotide sequence of the coding strand together with the deduced amino acid sequence of HES-3 are indicated. The amino acid sequence deduced from the region between the possible non-ATG translation codon ATT and the first ATG codon are shown in *italics* because it is not known whether this region is translated. The *upper-* and *lower-case letters* represent the exon sequence and the flanking and intron sequences, respectively. The *arrow* (at nucleotide residue 1) corresponds to the 5'-end of the several longest rat HES-3 cDNAs, and this position is probably a transcription initiation site. The canonical TATA box (tttaat) is *underlined*. The bHLH region is *overlined*. The stop codon is depicted by *asterisks*.

and BALB/c and MOL-MIT. Southern blot analysis showed that the three strains exhibited a restriction endonuclease fragment length variation (RFLV) of the *Hes-3* gene when digested by PstI; NOD and BALB/c mice gave a 2.3-kb band, and MOL-MIT mice

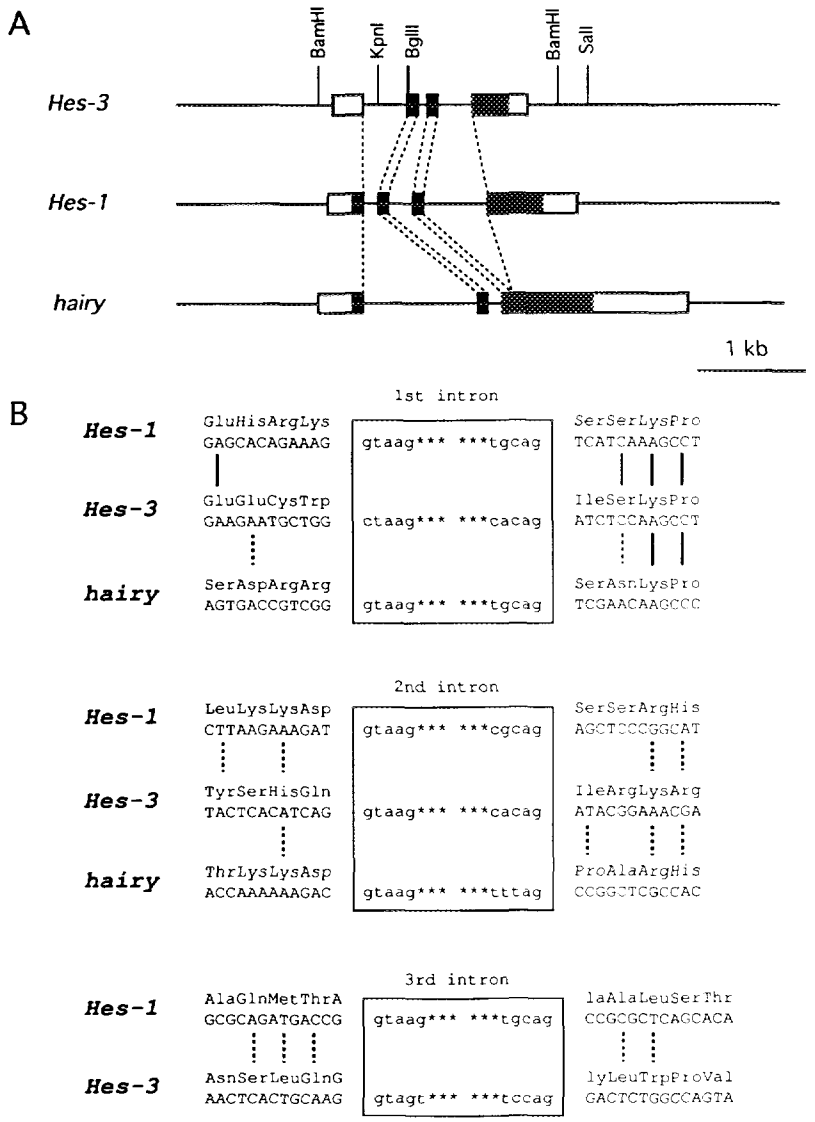

Fig. 2. Comparison of the mouse *Hes-3* and *Hes-1* and *Drosophila hairy* genes. (A) Schematic structures of the mouse *Hes-3* and *Hes-1* and *Drosophila hairy* genes. *Closed* and *open boxes* represent the coding and noncoding regions, respectively. *Thin lines* indicate the flanking and intron regions. All of the introns interrupted the equivalent positions of each transcribed region, which are connected by *dotted lines*. Restriction sites of the mouse *Hes-3* gene are shown. (B) Comparison of the exon-intron boundaries of the mouse *Hes-3* and *Hes-1* and *Drosophila hairy* genes. Amino acid residues conserved or changed conservatively among the three genes are connected by *solid* and *dotted lines*, respectively. It is not yet known whether the region of the *Hes-3* gene interrupted by the first intron shown here is translated (see *Discussion*).

a 1.8-kb band (data not shown). Therefore, high molecular weight DNAs isolated from a total of 77 backcross progenies, (NOD x MOL-MIT)F_1 x NOD and (BALB/c x MOL-MIT)F_1 x BALB/c, were subjected to linkage analysis by Southern blot experiments. As shown in Tables I and II, the *Hes-3* gene was closely linked to the *lymphocyte-specific*

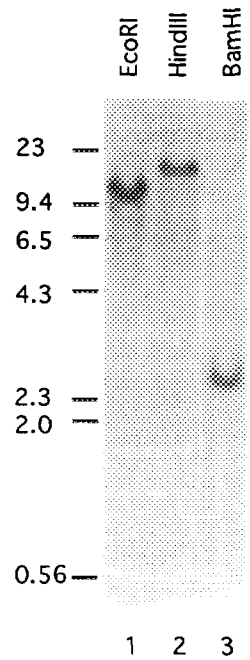

Fig. 3. Southern blot analysis of the mouse *Hes-3* gene. Mouse liver DNA (5 μg) was digested by a restriction enzyme, as indicated above each lane, and transferred to a nylon membrane filter. ^{32}P-labeled rat HES-3 cDNA was hybridized. The marker sizes (in kb) are indicated on the left.

Table I. Genotypes of Gametes Transmitted by (NOD x MOL-MIT)F_1 to Progeny of the Backcross[ab]

Chr.4	Gene locus							Number		
	38 *Orm-1*		57 *Lck*		(80) *Hes-3*					
F1 gametes	A		A		A			12	}	19
	B		B		B			7		
	A		A	*	B			3	}	6
	B		B	*	A			3		
	A	*	B		B			11	}	14
	B	*	A		A			3		
	A	*	B	*	A			1	}	4
	B	*	A	*	B			3		

Orm-1 - Lck, 18/43 = 41.9+/-7.5cM
Lck - Hes-3, 10/43 = 23.3+/-6.4cM

[a]A and B indicate the restriction types of NOD and MOL-MIT, respectively.
[b]The *asterisks* show recombination.

Table II. Genotypes of Gametes Transmitted by (BALB/c x MOL-MIT)F_1 to Progeny of the Backcross[ab]

Chr.4	Gene locus			Number	
	29 *Mup-1*	67 *Gpd-1*	(80) *Hes-3*		
F1 gametes	A	A	A	6	} 16
	B	B	B	10	
	A *	B	B	8	} 14
	B *	A	A	6	
	A *	B *	A	2	} 4
	B *	A *	B	2	

Mup-1 - *Gpd-1*, 18/34 = 52.9+/-8.6cM
Gpd-1 - *Hes-3*, 4/34 = 11.8+/-5.5cM

[a]A and B indicate the restriction types of BALB/c and MOL-MIT, respectively.
[b]The *asterisks* show recombination.

protein tyrosine kinase locus (*Lck*) and the *glucose phosphate dehydrogenase-1* locus (*Gpd-1*) on chromosome 4. *Lck* and *Gpd-1* are located 57 cM and 67 cM apart from the centromere, respectively (16). The three-point cross-test using the *Lck* and *Gpd-1* loci showed that, out of 43 progenies, 10 had recombination between the *Hes-3* and *Lck* loci

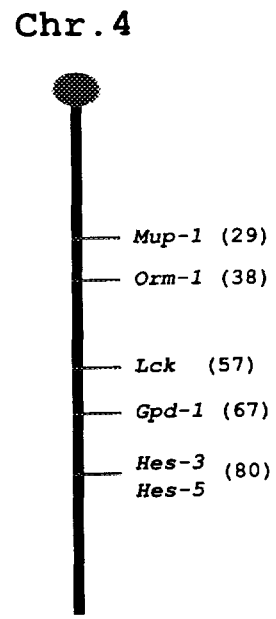

Fig. 4. Assignment of chromosomal loci of the *Hes-3* and *Hes-5* genes based on the results of Tables I and II. The *Mup-1* (position 29), *Orm-1* (position 38), *Lck* (position 57), and *Gpd-1* (position 67) loci and the calculated locus of the *Hes-3* gene (position 80) on chromosome 4 are shown on the right. *Hes-5* locus is closely linked to *Hes-3* locus (see *Text*), and therefore *Hes-5* locus is assigned to the same position as *Hes-3* locus.

(23%)(Table I) and that, out of 34 progenies, 4 had recombination between the *Hes-3* and *Gpd-1* loci (12%)(Table II). These results suggest that the mouse *Hes-3* gene is located ~80 cM apart from the centromere on chromosome 4 (Fig. 4).

We previously showed that the mouse *Hes-1* gene is located ~26 cM apart from the centromere on chromosome 16 (4), thus indicating that the *Hes-3* and *Hes-1* genes are not clustered. We next examined whether another member of the HES gene family, *Hes-5*, forms a cluster with either *Hes-3* or *Hes-1*. Southern blot analysis showed that the NOD and MOL-MIT strains exhibited RFLV of the *Hes-5* gene when digested by AatI (data not shown). Linkage analysis demonstrated that the *Hes-5* locus was closely linked to the *Hes-3* locus; out of 33 progenies examined, no one exhibited recombination between the two loci (data not shown, see Fig. 4). These results suggest that the *Hes-3* and *Hes-5* genes may be clustered on chromosome 4.

DISCUSSION

In this study, we report the whole nucleotide sequence and chromosomal locus of the mouse *Hes-3* gene. The *Hes-3* gene consists of four exons and its features of the genomic organization, such as the exon-intron boundaries, are very similar to those of the *Hes-1* and *Drosophila hairy* genes. Thus, the *Hes* genes may be more related to *hairy* than to the *E(spl)* genes, which do not have any introns within the protein-coding regions (7). Interspecies backcross analyses show that the *Hes-3* and *Hes-5* loci are very close to each other. These data suggest that some of the *Hes* genes may be clustered on chromosome 4, although it remains to be determined whether the *Hes* genes are located next to each other, like *E(spl)* gene complex.

Because HES-3 is a cerebellar Purkinje cell-specific HLH factor, it is possible that HES-3 mutation could result in immaturation or degeneration of Purkinje cells and cause cerebellar ataxia. Several hereditary diseases of cerebellar dysfunctions have been characterized and their chromosomal loci have been determined in mice (16). However, none of them are mapped to the *Hes-3* locus, suggesting that HES-3 is not responsible for these cerebellar abnormalities. Thus, further experiments such as a gene disruption will be necessary to address the exact function of HES-3 in Purkinje cell development.

Our previous data showed that the first methionine codon of rat HES-3 is located five amino acid residues upstream of the HLH domain. When this methionine codon is utilized as a translation initiation site, HES-3 has a short incomplete basic region. However, there is a possible non-ATG initiation sequence, ATT (17), at 207 nucleotides upstream of the methionine codon (1). Although our previous analysis showed that the in vitro translation starts from the first methionine codon but not from the ATT, we do not know yet where the in vivo translation starts (1). Our present study showed that the upstream possible non-ATG initiation codon ATT is also present in the mouse *Hes-3* gene exactly at the same position as that of rat *Hes-3* gene. Furthermore, the deduced amino acid sequences of the regions between the ATT and the first ATG codons of mouse and rat HES-3 are also well conserved (87% identity). Thus, it is possible that

translation may start from the ATT codon in vivo to some extent. If this is the case, the higher molecular weight form of HES-3 would have a complete basic region that enables HES-3 to directly bind to the DNA template. Further studies will be required to determine whether the two types of HES-3 exist in vivo.

Our studies now provide the basis to examine the problems discussed a bove, and further analyses will help understand the molecular mechanisms of Purkinje cell development.

ACKNOWLEDGMENTS

This work was supported by research grants from the Ministry of Education, Science and Culture of Japan, the Sankyo Foundation, and the Inamori Foundation.

REFERENCES

1. Sasai, Y., Kageyama, R., Tagawa, Y., Shigemoto, R., and Nakanishi, S. (1992) *Genes Dev.* **6,** 2620-2634.
2. Akazawa, C., Sasai, Y., Nakanishi, S., and Kageyama, R. (1992) *J. Biol. Chem.* **267,** 21879-21885.
3. Ishibashi, M., Sasai, Y., Nakanishi, S., and Kageyama, R. (1993) *Eur. J. Biochem.* **215,** 645-652.
4. Takebayashi, K., Sasai, Y., Sakai, Y., Watanabe, T., Nakanishi, S., and Kageyama, R. (1994) *J. Biol. Chem.* **269,** 5150-5156.
5. Ishibashi, M., Moriyoshi, K., Sasai, Y., Shiota, K., Nakanishi, S., and Kageyama, R. (1994) *EMBO J.* **13,** 1799-1805.
6. Moscoso del Prado, J., and Garcia-Bellido, A. (1984) *Roux's Arch. Dev. Biol.* **193,** 242-245.
7. Klämt, C., Knust, E., Tietze, K., and Campos-Ortega, J.A. (1989) *EMBO J.* **8,** 203-210.
8. Rushlow, C.A., Hogan, A., Pinchin, S.M., Howe, K.M., Lardelli, M., and Ish-Horowicz, D. (1989) *EMBO J.* **8,** 3095-3103.
9. Skeath, J.B., and Carroll, S.B. (1991) *Genes Dev.* **5,** 984-995.
10. Campuzano, S., and Modolell, J. (1992) *Trends Genet.* **8,** 202-208.
11. Benezra, R., Davis, R.L., Lockshon, D., Turner, D.L., and Weintraub, H. (1990) *Cell* **61,** 49-59.
12. Sun, X.H., Copeland, N.G., Jenkins, N.A., and Baltimore, D. (1991) *Mol. Cell. Biol.* **11,** 5603-5611.
13. Christy, B.A., Sanders, L.K., Lau, L.F., Copeland, N.G., Jenkins, N.A., and Nathans, D. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 1815-1819.
14. Ito, M. (1984) in *The Cerebellum and Neural Control*, pp. 24-25, Raven Press, New York.
15. Watanabe, T., Shimizu, A., Ohno, K., Masaki, S., and Kondo, K. (1989) *Biochem. Genet.* **27,** 689-697.
16. Green, M.C. (1989) in *Genetic variants and Strains of the Laboratory Mouse* (Lyon, M., and Searle, A.G., eds) 2nd Ed., pp. 12-403, Oxford University Press, Oxford.
17. Xiao, J.H., Davidson, I., Matthes, H., Garnier, J.-M., and Chambon, P. (1991) *Cell* **65,** 551-568.