Mammalian Genome 8, 863–865 (1997).

Mammalian Genome

© Springer-Verlag New York Inc. 1997

Cloning, chromosomal mapping, and expression of the human eHAND gene

Mark W. Russell,[1] Priscilla Baker,[2] Seigo Izumo[3]

[1]Cardiovascular Research Center and Departments of Pediatrics and Communicable Diseases, F1310 MCHC/Box 0204, University of Michigan Medical Center, 1500 East Medical Center Dr., Ann Arbor, Michigan 48109-0204, USA
[2]National Center for Human Genome Research, National Institutes of Health, Bethesda, Maryland, USA
[3]Department of Internal Medicine, University of Michigan Medical Center, Ann Arbor, Michigan, USA

Received: 15 April 1997 / Accepted: 20 June 1997

Basic helix-loop-helix transcription factors (bHLH) have been demonstrated to be vital for skeletal muscle, neuronal, hematopoietic, craniofacial, and limb development (Jan and Jan 1993; Zhuang et al. 1994; Howard et al. 1996). Recently, two bHLH transcription factors, dHAND and eHAND, were determined to be required for normal cardiac development in both the mouse and chicken. In developing chick embryos, antisense oligonucleotides directed at both of these proteins arrested cardiac development prior to cardiac looping (Srivastava et al. 1995). Recent studies have suggested that at least one of these two bHLH proteins, eHAND, may be downstream of the Csx/Nkx2.5 homeodomain protein in the transcription factor cascade that directs cardiac morphogenesis and that it may be required for normal cardiac looping, one of the earliest phases of cardiac structural development (Olsen and Srivastava 1996).

In the mouse, these two bHLH proteins are expressed in both the mesodermal and neural crest elements of the developing heart (Cserjesi et al. 1995). They are expressed in the precardiac mesoderm prior to formation of a heart tube, and their level of expression increases as the heart tube forms and begins to loop. Later in cardiac development, eHAND expression is restricted to the left ventricle and conotruncus, while dHAND expression is restricted to the right ventricle (Srivastava and Olsen 1996). Therefore, these two bHLH transcription factors may have a role not only in cardiac looping but in ventricular specification and growth. Furthermore, prominent eHAND expression in the conotruncus suggests that it also may have a role in conotruncal development. Owing to its potential role in cardiac looping, ventricular maturation and conotruncal development, eHAND may directly or indirectly have a role in the genesis of some congenital cardiac defects in humans. In this study, we present the cloning, cDNA sequencing, and high-content radiation hybrid mapping of the human eHAND gene.

BLAST sequence similarity search with the murine eHAND cDNA sequence identified a human CpG island clone (Genbank accession: Z63841) with high homology to the murine eHAND gene. By use of the MacVector 4.5 (Eastman Kodak Co., New Haven, Conn.) program PCR primers (sense: 5′-ACCAGCTACATCGCCTACCTGATG-3′, antisense: 5′-TCCCTTTTCCGCTTGCTCTCAC-3′) were designed to amplify a fragment of this putative gene from human genomic DNA. The resulting 122-bp PCR fragment was gel purified, random prime labeled with α-^{32}P dCTP, and used to screen a human fetal (26-week gestation) cardiac cDNA library (Clontech Inc., Palo Alto, Calif.). This 122-bp eHAND probe hybridized to 28 out of 1×10^6 plated clones (based on duplicate filters). Two clones were plaque purified and isolated from plate lysates with the Qiagen Lambda Maxi Kit (Qiagen Inc., Chatsworth, Calif.). The cDNA inserts, one 0.4 kb and one 1.6 kb, were isolated by digestion with *Eco*RI, subcloned into a pBluescript vector, and sequenced on both strands by fluorescent automated sequencing (Model 377A, Applied Biosystems, Foster City, Calif.).

The resulting 1.62-kb human eHAND cDNA sequence (Genbank accession: pending) includes a 190-bp 5′ untranslated region, a 645-bp open reading frame, and a 794-bp 3′ untranslated sequence with a polyadenylation signal. Comparison with other, previously identified eHAND genes revealed a very high nucleotide homology between the human, sheep (Genbank accession: U43716), murine (U40041), and chicken (S79216) eHAND transcripts (data not shown). The predicted human eHAND protein is 215 amino acids (aa) in length, while the murine and chicken eHAND genes encode for a 216-aa protein and a 204-aa protein respectively (Fig. 1). The murine and human eHAND proteins have 93% aa identity overall and 96% aa identity in the 57-aa basic helix-loop-helix region, while the sheep and human eHAND genes have 100% aa identity in the bHLH region. This is strong evidence that this transcript encodes the human homolog of the eHAND gene.

Amino acid sequence analysis revealed several potentially important phosphorylation sites within the eHAND protein. Conserved across all four species were consensus sites for phosphorylation by protein kinase C (residue 135), tyrosine kinase (residue 34), and cAMP- and cGMP-dependent protein kinases (residues 95, 132, and 137). An additional cAMP-dependent protein kinase site at the carboxy terminus of the protein (K199) was conserved in all species but the sheep. Interestingly, the sheep eHAND protein is highly homologous to the other three but would be predicted to stop at the aa corresponding to I198 in the human. The sheep cDNA sequence for this region contains a polyadenine and three polyguanine tracts (sequence: **GGG**ATT**AAAAGGGG**CGCACA**GGG**). If each of these tracts were shortened by one nucleotide (**GG**ATT**AAAGGG**CGCACA**GG**), then the resulting aa sequence would be identical to the other three species. This 3′ sequence divergence indicates that either this carboxy terminal region is not vital to the function of the sheep eHAND protein or that an expansion of these tracts occurred during sequencing. Therefore, it is possible that the K199 cAMP-dependent protein kinase site is also conserved across the four species and is a physiologically important site.

The regions demonstrating the greatest variability between the different eHAND proteins are the polyhistidine and alanine-rich tracts which are N-terminal of the basic helix-loop-helix domain. Similar polyhistidine and polyalanine sequences have been noted to occur together in other, developmentally regulated genes, such as the homeodomain proteins *cut* and *Gsh2,* but their functions have not been well defined (Hseih-Li et al. 1995). The variability between species of the eHAND polyhistidine and alanine-rich seg-

Correspondence to: M.W. Russell

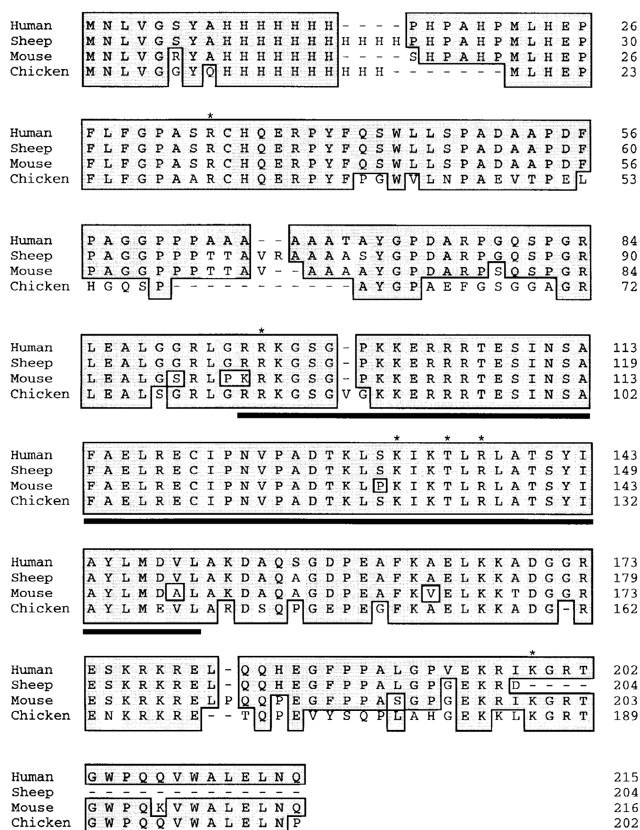

Fig. 1. Comparison of the predicted amino acid sequence of the human, sheep, mouse, and chicken eHAND genes. Identical amino acid residues are enclosed in shaded boxes. Conserved amino acid substitutions are contained within the box but are unshaded. Non-conserved amino acid substitutions are unshaded and not enclosed within a box. Gaps in the alignment are represented by dashes (-). The bHLH domain is underlined. Asterisks (*) denote those residues that are conserved consensus sites for phosphorylation by either cAMP or cGMP-dependent kinases (residues 95, 132, 137, and 199), tyrosine kinase (residue 34), or protein kinase C (residue 135).

ments has several possible explanations: 1) these regions are relatively less important to the functions of the eHAND protein, 2) the functions of these two regions vary slightly between the different species, or 3) although the presence of these two regions may be important to the functions of the eHAND protein, their precise length is less important.

To examine the expression pattern of eHAND in human adult tissues, a 340-bp *Pst*I fragment (including the entire bHLH coding sequence) was random prime labeled and hybridized overnight at 42°C to a multiple-tissue Northern blot (Clontech Inc.). The blot was washed in $2 \times$ SSC/0.1% SDS at room temperature for 30 min, then with $1 \times$ SSC/0.1% SDS at 50°C for 20 min. After exposure at –70°C for 96 h, two bands, 1.8 and 2.4 kb, were detected in the heart but not in brain, placenta, lung, liver, skeletal muscle, kidney, or pancreatic tissue (Fig. 2). Since the murine eHAND and dHAND cDNA sequences have 77% nucleotide identity over the corresponding 340-bp fragment, it is possible that this probe detects both transcripts. This possibility was examined by serially reprobing the blot with 3′ untranslated human eHAND and dHAND cDNA probes (data not shown). The 3′ eHAND probe detected a single band at 1.8 kb, while the dHAND probe detected a single band at 2.4 kb. Therefore, while the 2.4 kb band in Fig. 2 may represent an alternatively spliced eHAND transcript with a different 3′ untranslated region, it most likely represents cross-hybridization of the original probe to the dHAND transcript.

In order to determine the map location of the human eHAND gene, we developed an STS from the 3′ untranslated cDNA sequence (sense: 5′-ACGCAGGAAGATGAA-AGGCTG-3′, antisense: 5′-GATGGCAGGATGAACAAACAC-3′) and used it to map the gene on a high-content radiation hybrid panel (Stanford G3 RH panel; Brown 1996). The panel was scored twice, and a high LOD score of 7.1 to marker *D5S2640* on Chr 5q32 was obtained both times. This marker maps just telomeric to the Treacher Collins Syndrome locus and near the previously mapped annexin gene, Anx6 (Fig. 3). This map localization is consistent with the mapping of the mouse eHAND gene (Hand1) to mouse Chr 11 at 31.5 cM (Cross et al. 1995) just telomeric to the mouse *Anx6* gene on Chr 11 at 29 cM (Mouse Genome Database; http// www.informatics.jax.org/mgd.html).

The Treacher Collins syndrome (TCOF1), an autosomal dom-

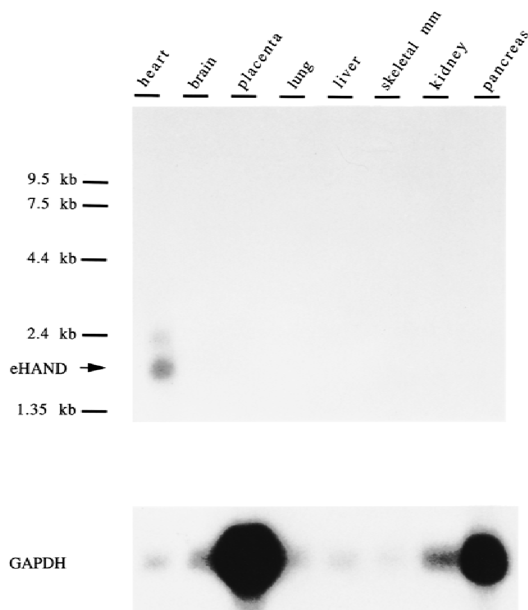

Fig. 2. Human multiple-tissue Northern blot analysis of eHAND expression with a 340-bp human eHAND cDNA *Pst*I fragment (includes the eHAND bHLH coding sequence). A GAPDH probe was used to demonstrate the differences in mRNA loading on the multiple-tissue Northern blot (Clontech Laboratories Inc.).

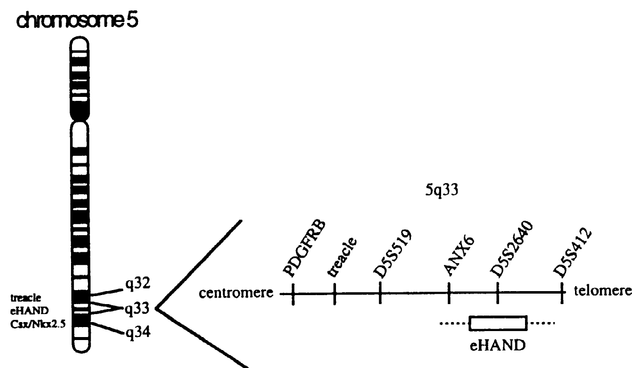

Fig. 3. Representation of human Chr 5 with approximate ordering of genes and markers on 5q33 [based on Stanford RH map (http://shgc.stanford.edu), HUGO Chr 5 genetic map (through Genome Database; http://gdbwww.gdb.org/), and TCOF1 physical map (The Treacher Collins Syndrome Collaborative Group 1996)]. Radiation hybrid mapping localizes eHAND to within 500 kb of D5S2640 (Brown 1996). The exact localization of eHAND compared with ANX6 cannot be precisely determined; approximate eHAND localization is represented by a box with dashed lines.

inant disorder characterized by abnormal pinnae, hypoplasia of the facial bones and cleft palate, is caused by mutations or deletions of the treacle gene, which would be predicted to terminate transcription prematurely (The Treacher Collins Syndrome Collaborative Group 1996). Patients with TCOF1 are at increased risk for congenital cardiovascular abnormalities, which can include cardiomyopathy (2% of patients) and ventricular septal defects (Vatré 1971). Given the vital role of eHAND in murine and avian cardiac development and the proximity of the human eHAND gene to treacle, it is possible that some TCOF1 patients with congenital cardiac abnormalities may have larger deletions involving the eHAND gene. Evidence for the presence of a gene or genes on Chr 5q involved in cardiac growth and development is demonstrated by patients with trisomy 5q, approximately 67% of whom have cardiac defects, ranging from ventricular septal defects to more complex abnormalities (Turbay et al. 1996).

Interestingly, while the mouse eHAND gene is highly expressed in the heart during development, it is not expressed at levels detectable by Northern hybridization or RNase protection in the adult heart (Cserjesi et al. 1995). The relatively high level of eHAND expression in the human adult heart suggests that, in the human, it may have cardiac functions beyond its roles in cardiac structural development. In skeletal muscle, there is evidence that bHLH factors such as myoD and myogenin are important in the expression and activity of immediate early response genes such as c-fos and c-jun (Rudnicki and Jaenisch 1995). MyoD, a skeletal muscle-specific bHLH protein, has been determined to associate with the immediate early protein c-jun and down-regulate the expression of c-fos (Rudnicki and Jaenisch 1995). In the heart, increases in the level of expression of these immediate early genes have been associated with ischemia and models of hypertrophy (Morgan and Baker 1991). Since MyoD is not expressed in cardiac myocytes, it is possible that eHAND is the bHLH protein that regulates the expression and activity of some immediate early response genes in the heart. Consequently, it may have a role in structural remodeling or hypertrophy of the heart in response to stimuli such as hypertension and heart failure.

The human eHAND gene, therefore, is a highly conserved bHLH transcription factor that is expressed in the prenatal and adult heart. It has been demonstrated to have a vital role in the morphologic development of the heart in other species, and its relatively high level of expression in the adult human heart suggests that it may have an important role in the transcriptional regulation of adult cardiocytes. The cloning of this gene will allow the evaluation of its potential roles in human cardiac structural development and in cardiac adaptation to acquired cardiovascular disorders.

Acknowledgments. We thank Drs. Lawrence Brody and Stephanie Burns Wechsler for their critical readings of this manuscript and for their technical advice. This research was supported in part by grants from the Child Health Research Center, University of Michigan (M.W. Russell) and the NIH (S. Izumo).

References

Brown DM (1996) High content mapping with radiation hybrids. J NIH Res 8, 57

Cross JC, Flannery ML, Blanar MA, Steingrimsson E, Jenkins NA, Copeland NG, Rutter WB, Werb Z (1995) Hxt encodes a basic helix-loop-helix transcription factor that regulates trophoblast cell development. Development 121, 2513–2523

Cserjesi P, Brown D, Lyons GE, Olson EN (1995) Expression of the novel helix-loop-helix gene eHAND in neural crest derivatives and extraembryonic membranes during mouse development. Dev Biol 170, 664–678

Howard TD, Paznekas WA, Green ED, Chiang LC, Ma N, De Luna RIO, Delgado CG, Gonzalez-Ramos M, Kline AD, Jabs EW (1996) Mutations in *TWIST,* a basic helix-loop-helix transcription factor, in Saethre-Chotzen syndrome. Nature Genet 15, 36–41

Hseih-Li HM, Witte DP, Szucsik JC, Weinstein M, Li H, Potter SS (1995) *Gsh-2,* a murine homeobox gene expressed in the developing brain. Mech Dev 50, 177–186

Jan YN, Jan LY (1993) HLH proteins, fly neurogenesis, and vertebrate myogenesis. Cell 75, 827–830

Morgan HE, Baker KM (1991) Cardiac hypertrophy. Mechanical, neural, and endocrine dependence. Circulation 83, 13–25

Olson EN, Srivastava D (1996) Molecular pathways controlling heart development. Science 272, 671–676

Rudnicki MA, Jaenisch R (1995) The MyoD family of transcription factors and skeletal myogenesis. Bioessays 17, 203–209

Srivastava D, Olsen EN (1996) Deletion of the cardiac basic helix-loop-helix transcription factor, dHAND, results in embryonic lethality. Circulation (Suppl.) 94, I–46

Srivastava D, Cserjesi P, Olson EN (1995) A subclass of bHLH proteins required for cardiac morphogenesis. Science 270, 1995–1999

The Treacher Collins Syndrome Collaborative Group (1996) Positional cloning of a gene involved in the pathogenesis of Treacher Collins Syndrome. Nature Genet 12, 130–136

Turbay D, Wechsler SB, Blanchard KM, Izumo S (1996) Molecular cloning, chromosomal mapping and characterization of the human cardiac-specific homeobox gene hCsx. Mol Med 2, 86–96

Vatré J-L (1971) Étude génétique et classification clinique de 154 cas de dysostose mandibulo-faciale (syndrome de Franschetti), avec de leurs associations malformatives. J Génét hum 19, 17–100

Zhuang Y, Soriano P, Weintraub H (1994) The helix-loop-helix gene E2A is required for B cell formation. Cell 79, 875–884