Cytogenet Cell Genet 90:106–114 (2000)

Cytogenetics and Cell Genetics

Isolation and characterization of a new FHL1 variant (FHL1C) from porcine skeletal muscle

A. Krempler,[a] S. Kollers,[b] R. Fries[b] and B. Brenig[a]

[a] Institute of Veterinary Medicine, University of Göttingen;
[b] Institute of Animal Breeding and Genetics, Technical University of Munich (Germany)

Abstract. Four and a half LIM domain protein 1 (FHL1) was initially described as an abundant skeletal muscle protein with four LIM domains and a GATA like zinc finger. FHL1 was shown to be expressed in skeletal muscle as well as in a variety of other tissues. Recently, alternatively spliced FHL1 mRNAs were identified coding for C-terminal truncated proteins. The tissue distribution of these variants is more restricted and their functional properties seem to be different. We have isolated and characterized a new variant of FHL1 from porcine skeletal muscle (FHL1C). FHL1C is characterized by a newly identified start codon resulting in a 16 amino acids longer N-terminal region. We have isolated and characterized the porcine FHL1C gene spanning approximately 14 kb and harboring six exons. Using primer extension analysis, the transcription start site of FHL1C was mapped, indicating that FHL1C is regulated by an alternative promoter. The tissue distribution of FHL1C expression was studied by RT-PCR. The porcine FHL1C gene was assigned to the distal part of the long arm of the X chromosome by fluorescence in situ hybridization and screening of a somatic porcine/rodent cell hybrid panel.

Muscle cell differentiation and function is mainly triggered by the action of the myogenic basic helix-loop-helix-proteins, namely MyoD (Davis et al., 1987), Myf-5 (Braun et al., 1989), Myogenin (Edmondson and Olson, 1989; Wright et al., 1989) and MRF-4 (Rhodes and Konieczny 1989; Minor and Wold, 1990). But cell function and especially transcriptional control requires the action of many factors working together in a big machinery. In this respect, proteins from the large family of zinc finger proteins gained more and more attention (Kempler and Brenig, 1999).

To identify zinc finger proteins that are important in the function of porcine skeletal muscle cells, we screened a cDNA library by a PCR approach. The most abundant product was a cDNA sequence homologous to a recently described group of LIM-domain proteins, termed four and a half LIM domain proteins (FHLs). This group is characterized by four repeats of a LIM domain and a single zinc finger at its N-terminal end which is related to the GATA1-type zinc fingers (Morgan and Madgwick, 1996). The LIM motif is a protein-binding interface found in a diverse group of proteins that includes LIM kinases and LIM-homeodomain proteins. The motif is frequently found in proteins involved in cell differentiation and cell fate determination (Sanchez-Garcia and Rabbitts, 1994).

Up to now four different isoforms of four and a half LIM domain proteins were described (FHL1-4) in various species (Morgan et al., 1996; Chan et al., 1998; Morgan and Madgwick, 1999). FHL1 and FHL3 are mainly expressed in skeletal muscle where they show different expression patterns in skeletal muscle cell lines and seem to be reciprocally related to each other (Lee et al., 1998; Morgan and Madgwick, 1999). FHL2 is the main isoform in cardiac tissue (Chan et al., 1998) and is identical to the protein DRAL on the amino acid level (Genini et al., 1997). FHL4 was found to be expressed exclusively in testis, supposing a function in spermatogenesis (Morgan et al., 1999).

The chromosome locations of FHL1, FHL2 and FHL3 are already known in humans. FHL1 is located on chromosome Xq27.2 near the fragile X allele, the myotubularin and emirin

Received 5 April 2000; manuscript accepted 22 May 2000.

Request reprints from Dr. Bertram Brenig, Institute of Veterinary Medicine, Groner Landstrasse 2, 3 D–37073 Göttingen (Germany); telephone: +49-551-393383; fax: +49-551-393392; e-mail bbrenig@gwdg.de

Table 1. Primer sequences and application

Primer	Sequence (5'–3')	Application
ZF-deg	ACK CAY ACW GGR GAG AAG CCS	screen of cDNA library
pBK-CMVT7	GTA ATA CGA CTC ACT ATA GGG C	screen of cDNA library
F1	GGG TTT GGA GAC TTG CAT G	isolation of single phages screen of genomic phage library
F2	AGA CTC ACA AGA GCA GCC G	isolation of single phages screen of genomic phage library
F3	AAG AGA TAA AGC CCC CAC	isolation of single phages
F4	GCC ACA AAT CCA AAT CC	isolation of single phages
pBK-CMVT3	AAT TAA CCC TCA CTA AAG GG	isolation of single phages
F1gen	ATC AAA AGA GAG AGG GAC G	screen of PAC library
F2gen	ATC ACA GGG ACA GAA GG	screen of PAC library
F/R	GTA ACA CAC ACA AAG CAC TCG GC	5'- RACE
F/NR	CCG TCC TTC TGC ACG TAC TTC TTT CCC	5'- RACE
F/NR2	TTT AAG GGG AAG CGG TGA ACA CG	5'- RACE
Ftrans-rev	TTC TGC ACG TAC TTC TTT CC	RT - PCR
F-RTEx6	AAG AAG GAC AGG AGG AGC	RT - PCR
Ftrans-fwd1	TTT CTC TGC GGC CTC TC	RT - PCR
Ftrans-fwd2	TTT TTA CTG GGT CCT ATT CTT C	RT - PCR
Ftrans-fwd3	TCT CGG TTA TTT TCG TGT TC	RT - PCR
F-RTEx4	GAA CGT GGA GTA CAA GGG	RT - PCR
Fprex1	CCC TGC AGT AGT GGC AGT CAA ACT TCT CGG	primer extension
Fprex4	CCC CAG AGA TCG CAA CCT TAA TCC C	primer extension
Fg31	AAC AGC CTC ACT GAC TGA C	screen of cell hybrid panel
Fg32	AAA CCC AAC CTC CTT CCT C	screen of cell hybrid panel

genes (Lee et al., 1998; Greene et al., 1999). FHL2 was assigned to chromosome 2q12→q13 with no obvious correlation to a cardiac disorder (Chan et al., 1998; Genini et al., 1997). FHL3 was positioned to chromosome 1p34.2→p32.3 in close proximity to congenital ptosis allele (PTOS1) and the Schwartz-Jampel Syndrome allele (SJS). Both phenotypes can also include myopathic disorders (Morgan et al., 1999; Lee et al., 1998).

Different splicing variants have been isolated for FHL1 (FHL1B/SLIMMER and KyoT2). These splicing variants are devoid of one or two LIM domains in the C-terminal part. The shortened forms are interesting since a frameshift, resulting from the splicing event, leads to a different C-terminal amino acid composition (Taniguchi et al., 1998; Brown et al., 1999; Lee et al., 1999). In KyoT2 this region was shown to bind to the transcription factor RBP-J and negatively regulate its transcriptional activity (Taniguchi et al., 1998). FHL1 is the only member of the FHL family where the genomic structure is partly known, being organized in six exons plus one alternatively spliced exon. The reported sequence spans a genomic region of around 4 kb (Greene et al., 1999; Taniguchi et al., 1998).

Materials and methods

Construction of a porcine skeletal muscle cDNA library and preparation of plate pools

Total RNA was isolated from skeletal muscle (1 mg tissue) of a 4 w old piglet (German Landrace) using 25 ml Trizol reagent (Gibco/BRL). mRNA was purified using the Poly(A) Quick mRNA Isolation Kit (Stratagene). For the construction of the library the ZAP Express cDNA synthesis Kit (Stratagene) was used. The final titer of the porcine skeletal muscle cDNA library was 3.4×10^8 pfu/ml. For the plate pools 5×10^5 pfu were spread on 10 NZY agar plates (5×10^4 pfu/plate) and grown overnight at 37 °C. Each plate was overlaid with 10 ml SM buffer (100 mM NaCl, 10 mM $MgSO_4$, 50 mM Tris-HCl [pH 7.5], 0.01 % gelatin) and after incubation at 4 °C with gentle shaking the supernatants were removed and the plates washed with an additional 10 ml of SM buffer. Both supernatants were pooled, treated with 1 ml chloroform and centrifuged (5,000 rpm, 10 min) at room temperature. Aliquots (1 ml) were taken from the supernatant and stored either at 4 °C with the addition of 10 % chloroform or at –80 °C with the addition of 7 % DMSO. The remainder of the supernatants was used to prepare DNA following standard protocols (Ausubel et al., 1990). The isolated DNA was dissolved in an appropriate volume of TE (10 mM Tris-HCl [pH 7.5], 1 mM EDTA) and stored at 4 °C.

Screening of plate pools and isolation of porcine FHL1 cDNA phages

The isolated DNA from the phage pools was used to screen the library by PCR. A degenerated forward primer ZF-deg was used in combination with a reverse primer sequence (T7 promoter region) included in the phage arm (pBK-CMVT7). Touch-down PCR was performed in a 100 µl standard reaction (1 µM of each primer, 2 pM dNTPs (Roth), 5 units Taq polymerase (Boehringer)) with the following cycling conditions in a Techne Genius thermal cycler (thermo-DUX): initial denaturation for 3 min at 95 °C; 30 cycles with 1 min of denaturation at 95 °C, followed by 1 min at the annealing temperature and 1 min of extension at 72 °C. The annealing temperature was decreased by 0.9 °C per cycle starting with 70 °C. Products were cloned into the *Sma*I site of the pGEM-4Z vector (Promega) and transformed into the XL1-blue *E. coli* (Stratagene). Single clones were sequenced and compared to entries in the GenBank (NCBI) with the BLAST algorithm (Altschul et al., 1990).

For the isolation of FHL1 cDNAs, PCR products were used as a hybridization probe in filter lifts of the corresponding phage pools. 5×10^4 pfu/pool were plated on NZY agar plates and double filter lifts were made. For hybridization and detection the ECL nucleic acids labeling kit (Amersham-Pharmacia) was used following the manufacturers instructions. After the isolation of a single positive phage the pBK-CMV phagmid was excised with the ExAssist helper phage (Stratagene) and transformed into the XLOLR *E. coli* (Stratagene).

Isolation of additional, overlapping cDNA clones was performed mainly as described above. Primers F1, F2, F3, and F4 were designed from the 5′-end of the longest available DNA sequence. Positive pools were used for a second PCR where the forward primer of the first screen was replaced by a primer that originated from the T3 promoter sequence included in the phage (pBK-CMVT3). All primers used in the experiments are listed in Table 1.

Isolation of genomic fragments from a phage library and a PAC library

Primers F1 and F2 were used to screen 20 plate pools of a porcine genomic library (λFIX II, Stratagene) and a single phage was isolated by hybridization with the PCR product as a probe. Phage-DNA was prepared following a protocol supplied by Qiagen. Isolated phage DNA was cut with different restriction enzymes and the products were cloned into pGEM-4Z for sequencing.

The isolation of the genomic region of FHL1 from a porcine PAC library (Al-Bayati et al., 1999) was performed with primers from an intron region (F1gen and F2gen). The screening procedure was as described by Al-Bayati et al. (1999). PAC-DNA was isolated following a protocol from Qiagen. The DNA was digested with different enzymes, blotted onto Nytran plus membrane (Schleicher and Schuell) and hybridized with the 5′-RACE product. Appropriate fragments were used for subcloning and sequencing.

DNA sequencing

Sequencing reactions were performed with the Thermosequenase Sequencing Kit (Amersham, Pharmacia) and the products were applied to a 6 % Sequa-Gel (Biozym) and run in a LI-COR 4000 Sequencer. Analysis of the sequences was done with the program SEQUENCHER 3.1. The DNA sequences have been deposited with the EMBL DNA Sequence database under accession numbers AJ275967 (FHL1 gene) and AJ275968 (FHL1C cDNA).

5′-RACE PCR, RT-PCR, and primer extension analysis

Total RNA was isolated from 1 mg of porcine skeletal muscle (German Landrace) with Trizol reagent (Gibco/BRL). mRNA was obtained by purification with Oligotex (Qiagen) following the suppliers instructions. The eluted mRNA was subjected to cDNA synthesis using the Marathon cDNA Amplification Kit (Clontech). Subsequent PCR was performed with the Advantage 2 PCR Enzyme System (Clontech). The first and second amplifications were carried out using the supplied primers AP1 and AP2 and the gene specific primers F/R, F/NR, and F/NR2. Cycling conditions were as described in the manufacturers manual.

For RT-PCR 1 μg RNA was subjected to reverse transcription with SuperScriptII reverse transcriptase (Gibco/BRL). For primers we used gene specific oligonucleotides (Ftrans-rev and F-RTEx6) to obtain only specific templates for the following PCR: 2 μl of the reverse transcription reaction were used in a 50 μl standard PCR reaction, where the reverse primer was the same as for the RT reaction (forward primers: Ftrans-fwd1, Ftrans-fwd2, Ftrans-fwd3, and F-RTEx4). Specific products were separated on a 2 % agarose gel, cut, cloned and sequenced.

Primer extension was performed as described (Altermann et al., 1999). The primers (Fprex1 and Fprex4) originated from the first and second exon coding for the FHL1C variant. Sequencing reactions on plasmids harboring the sequence of the corresponding extension primer were used as size standards.

Chromosome assignment

Fluorescence in situ hybridization (FISH) was performed as described before (Toldo et al., 1993; Solinas-Toldo et al., 1995). Briefly, swine metaphase spreads were prepared from peripheral lymphocytes obtained from a normal, healthy boar. The QFQ-banded spreads were photographed prior to in situ hybridization using a cooled CCD camera. The hybridization signal was detected and amplified by incubations with avidin-FITC and biotinylated anti-avidin. The chromosomes were DAPI-counterstained. The probe was assigned to a specific chromosome band by measuring the distance of the signal position to the telomere and the length of the entire chromosome and subsequent calculation of the fractional length (FLpter).

A somatic cell hybrid panel consisting of DNAs from 27 rodent/porcine hybrid cell lines (Yerle et al., 1996) was screened by PCR. The primers (Fg31 and Fg32) originated from an intron region of the porcine FHL1 gene and were used in a 50 μl standard PCR reaction together with 25 ng of template DNA. Evaluation of the PCR results was performed by use of the interactive PCR data interpreting web-page at INRA (http://www.toulouse.inra.fr/lgc/pig/hybrid.htm).

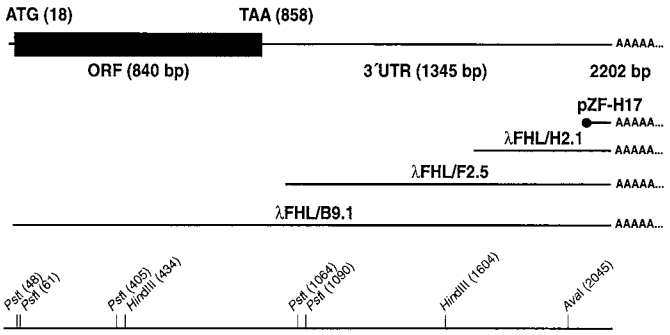

Fig. 1. Isolation and analysis of the porcine FHL1 cDNA. The porcine FHL1 cDNA was characterized by three overlapping cDNA phages (λFHL/H2.1, λFHL/F2.5 and λFHL/B9.1). λFHL/H2.1 was isolated by hybridization with the PCR product pZF-H17. The open reading frame is indicated by a black box, the positions of the start and stop codon are indicated. A partial restriction map is shown.

Results

Isolation and characterization of the porcine FHL1 cDNA

In an attempt to isolate zinc finger proteins expressed in porcine skeletal muscle, we screened a cDNA library for zinc finger motifs. The sequence of the forward primer ZF-deg resembled parts of the highly conserved inter-finger region, the H-C link, in classical zinc finger proteins. Several fragments were obtained, cloned and sequenced. One of the most abundant products was a 180-bp DNA sequence (pZF-H17) with homology to the 3′-untranslated region of the human LIM domain protein SLIM1/FHL1 (Morgan et al., 1996; Lee et al., 1998). We therefore isolated the corresponding phage, designated λFHL/H2.1, containing a 518-bp insert and two additional phages, λFHL/F2.1 and λFHL/B9.1. Together these phages harbored the complete 3′-untranslated and coding region of the porcine FHL1 cDNA. The cDNA had a length of 2,202 bp, with a 3′-UTR of 1,345 bp and an ORF of 840 bp. The translation start site was very close to the 5′-end (18 nt) and resembled the consensus initiation sequence (RCCAUGG) (Kozak, 1996). A canonical polyadenylation signal (AATAAA) was present at position 2184, followed by a polyA-tail 14 bp downstream from this motif (Fig. 1).

Translation of the ORF results in a 280-amino-acid protein. The protein contains four LIM domains (C2HC/C4) and a single N-terminal GATA like zinc finger. Each LIM domain is separated from the following domain by eight amino acids. Analysis of the amino acid composition revealed a high content of lysines (12.5 %) and cysteines (11.5 %). The porcine FHL1 has a calculated molecular mass of 33.6 kDa with a pI of 8.7.

Identification of FHL1C by 5′-RACE PCR

To further analyze the 5′-region of the porcine FHL1 cDNA, we performed 5′-RACE PCR. A 175-bp fragment was obtained that extended the 5′-UTR by 102 nt. Within this sequence an additional inframe ATG 48 bp upstream of the common FHL1 start codon was present. Translation from this ATG elongated the FHL1 protein by 16 amino acids (Fig. 2A). No homologies

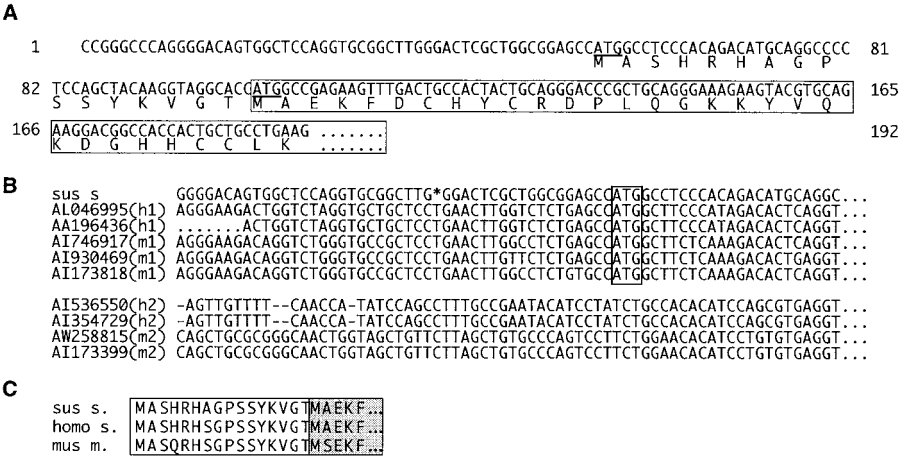

Fig. 2. DNA and amino acid sequence comparisons of FHL1 and FHL1C. (**A**) Nucleotide sequence and deduced amino acid sequence of the porcine FHL1C 5′-RACE clone. The newly identified ATG and the common FHL1 start codon are underlined. The box indicates the region also present in FHL1, FHL1B/SLIMMER and KyoT2. (**B**) Multiple alignment of various EST entries with homologies to FHL1 and FHL1C. The upper panel shows human (h1: AL046995, AA196436) and mouse (m2: AI746917, AI930469, AI173818) sequences that correspond to the porcine FHL1C isolate. The lower panel consists of database entries from man (h2: AI536550, AI354729) and mouse (m2: AW258815, AI173399) that are homologous to the published FHL1 5′-regions. The conserved start codon in FHL1C is boxed. (**C**) Deduced amino acid sequence of FHL1C of pig (sus s.), man (homo s.) and mouse (mus m.). The open box indicates the 16 additional amino acids at the N-terminus of FHL1C. The grey box represents the common FHL1, FHL1B/SLIMMER and KyoT2 sequence.

Table 2. Exon/intron boundaries of the porcine FHL1/FHL1C gene

Exon	bp	Splice acceptor	Splice donor	Intron	bp
1	364		AGACATGCAG**gt**aaacgcgg	1	6714
2	182	ccccccgc**ag**GCCCCTCCAG	TGACTCCAAG**gt**agcgcccg	2	212
3	175	cggttccc**ag**GAGGTGCACT	ATCGTGGCAG**gt**accggcca	3	614
4	170	ccccccgc**ag**GCGATCAGAA	GTGCAACAAG**gt**acgttgtc	4	438
5a	187	cctggtct**ag**GCCATCACAT	CCCATCACTG**gt**aggctaaa	5a	575
5b	200	ccatcctc**ag**GGAAAAGGAC	TCGAGGCCCG**gt**aagtgcac	5b	352
6	1496	tcctccgc**ag**GGTTTGGTAA			

to known proteins were found when comparing the 16 amino acids to the SWISSPROT database.

To investigate whether different 5′-regions account for diverse gene products of FHL1, we compared the first 50 bp of the human FHL1 coding region with the dbEST database. The analysis revealed the existence of two different 5′-regions in FHL1 cDNAs isolated from man and mouse. Several ESTs (AL046995, AA196436, AI746917, AI930469, AI73818) matched the porcine 5′-UTR sequence (Fig. 2B). In all cases the newly identified start codon upstream of the common FHL1 ATG was present. The deduced amino acid sequences were homologous to the porcine N-terminal region indicating that there exists an N-terminally elongated variant of FHL1, named FHL1C (Fig. 2C).

Genomic structure of the porcine FHL1C gene

Primers F1 and F2, flanking a 204-bp region of the porcine 3′-UTR, were used to screen a genomic phage library by PCR. A single phage, designated λFHLgen9.2.2, was isolated. The fragment carried five exons coding for the porcine FHL1C variant and a previously described, alternatively spliced exon found in the human FHL1 variant FHL1B/SLIMMER. The exons were within a genomic region of about 4.6 kb. The first exon included the common FHL1 start codon and the sixth exon carried the stop codon together with the 3′-UTR. The 3′-part of the RACE clone was found on the first exon, the 5′-part of the RACE product was not present on the insert. The homology ended at a potential splice site with an AG dinucleotide and a poly-pyrimidine stretch 1.5 kb downstream from the first base of the genomic phage insert.

To obtain the remaining upstream sequences, we isolated a PAC clone, designated PACFHL-O062-F4, using primers F1gen and F2gen. A 13-kb *Eco*RI restriction fragment that hybridized with the 5′-RACE product was isolated. This fragment overlapped with the 5′-end of λFHLgen9.2.2 and harbored the remaining DNA sequences of the 5′-RACE product. This 5′-part is separated from the following exon by a 6.7-kb intron. Hence, the new FHL1C variant is encoded by six exons spanning a genomic region of approximately 14 kb (Fig. 3). All exon/intron junctions followed the GT/AG rule of splice donor/acceptor sites (Table 2).

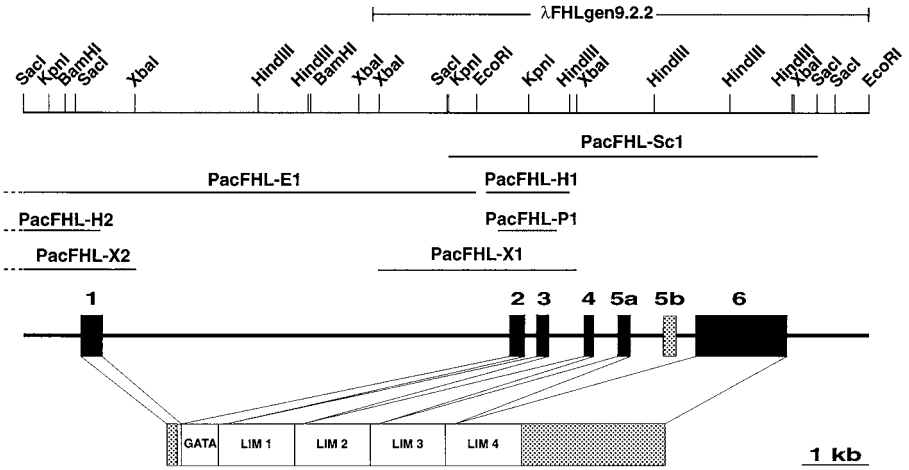

Fig. 3. Genomic organization of porcine FHL1C. Schematic representation of the genomic organization of FHL1C. In the upper panel the positions of the phage insert and the isolated PAC fragments are indicated. The lower panel illustrates the genomic organization of FHL1C. Black boxes represent exons and the cross-hatched box indicates the alternatively spliced exon used in FHL1B/SLIMMER.

Analysis of the 5′-region of FHL1C

To map the transcription start site of the FHL1C gene, we performed primer extension analysis. Elongation of primer Fprex1 was terminated at an AC dinucleotide 396 bp upstream of its binding site in exon 2. Extension of a second primer (Fprex4), binding within exon 1, resulted in the same extension product as well as an additional one further upstream. This longer product ended at the sequence CCACGGC, 290 bp from the last known base of the 5′-RACE clone (Fig. 4A). 36–41 bp upstream a TATAA-box and at position –79 to –85 a CAAT-box was identified. Comparison of this region with the TRANSFAC database (Heinemeyer et al., 1999) revealed an AP1 binding site in negative orientation, the core sequence for four Sp1 binding sites and two core motives of E-boxes (Fig. 4B). To confirm the transcription start site we performed RT-PCR with different primers between the transcription start and the exon-intron boundary of exon 1. Three forward (Ftrans-fwd1, Ftrans-fwd2, Ftrans-fwd3) and one reverse primer binding in exon 2 (Ftrans-rev) were used. Ftrans-fwd2 and Ftrans-fwd3 resulted in products of 335 bp and 236 bp and were confirmed by sequencing. No product was obtained using Ftrans-fwd1 (data not shown). Therefore the region 247 bp upstream the exon-intron boundary is part of exon 1 of FHL1C.

Expression of FHL1 and FHL1C

To analyze the tissue distribution of FHL1 and FHL1C expression, RNA was isolated from skeletal muscle, brain, heart, lung, liver, spleen, kidney, blood vessel and testis. RT-PCRs were performed with two different primer pairs. F-RTEx6 and F-RTEx4 were used to amplify FHL1 transcripts containing at least exon 6 and exon 4. This allowed us to detect FHL1 and FHL1C transcripts, the splice variant FHL1B/SLIMMER including the alternatively spliced exon 5b as well as KyoT2 that lacks exon 5a and 5b. Three different transcripts were detected with 318 bp, 436 bp and 637 bp. The fragments were isolated and sequenced. Only the 637-bp band contained the expected FHL1 and/or FHL1C transcript, whereas the other products contained unknown sequences. Transcripts for the splice variants FHL1B/SLIMMER and KyoT2 were not detected. The highest intensities of the 637-bp PCR product were found in skeletal muscle and in the heart. Significant amounts were also seen in lung and the bloodvessel, weak signals were present in brain, spleen, kidney and testis. No product could be detected in liver (Fig. 5, upper panel).

To analyze the expression pattern of specific FHL1C transcripts we performed anchored PCR. In the first round we used the primers Ftrans-fwd2 and Ftrans-rev. The second amplification was performed with the primers Ftrans-fwd3 and Ftrans-rev. The products (239 bp) showed a high tissue specificity. The most prominent signal was seen in skeletal muscle. Lung and blood vessel showed significant amounts of amplification products. Weak signals were detected in heart. All other tissues showed no specific signal (Fig. 5, lower panel).

Chromosomal assignment of the FHL1/FHL1C gene

The phage clone λFHLgen9.2.2 was used to locate the porcine FHL1 gene. Four preparations of metaphase spreads were analyzed and all of them showed a single hybridization signal on the long arm of the X chromosome. Calculation of the fractional length (see Materials and methods) assigned the porcine FHL1 gene to chromosome Xq22→q25 (Fig. 6).

To refine the chromosome location we screened a rodent/porcine cell hybrid panel (Yerle et al., 1996) by PCR. Evaluation of the PCR-results positioned the porcine FHL1 gene on chromosome Xq23→q24 (probability 0.7992; maximum correlation 0.87; error risk lower than 0.1%) (Chevalet et al., 1997).

Discussion

Using a touch-down PCR strategy, we were able to isolate cDNA clones of the porcine FHL1 gene. Like the human FHL1 ortholog (Morgan et al., 1996), the porcine FHL1 cDNA contains an ORF of 840 bp, coding for 280 amino acids. The porcine and human DNA sequences are highly conserved with 94.7% identity. The TAA stop codon is followed by a 1,331-bp

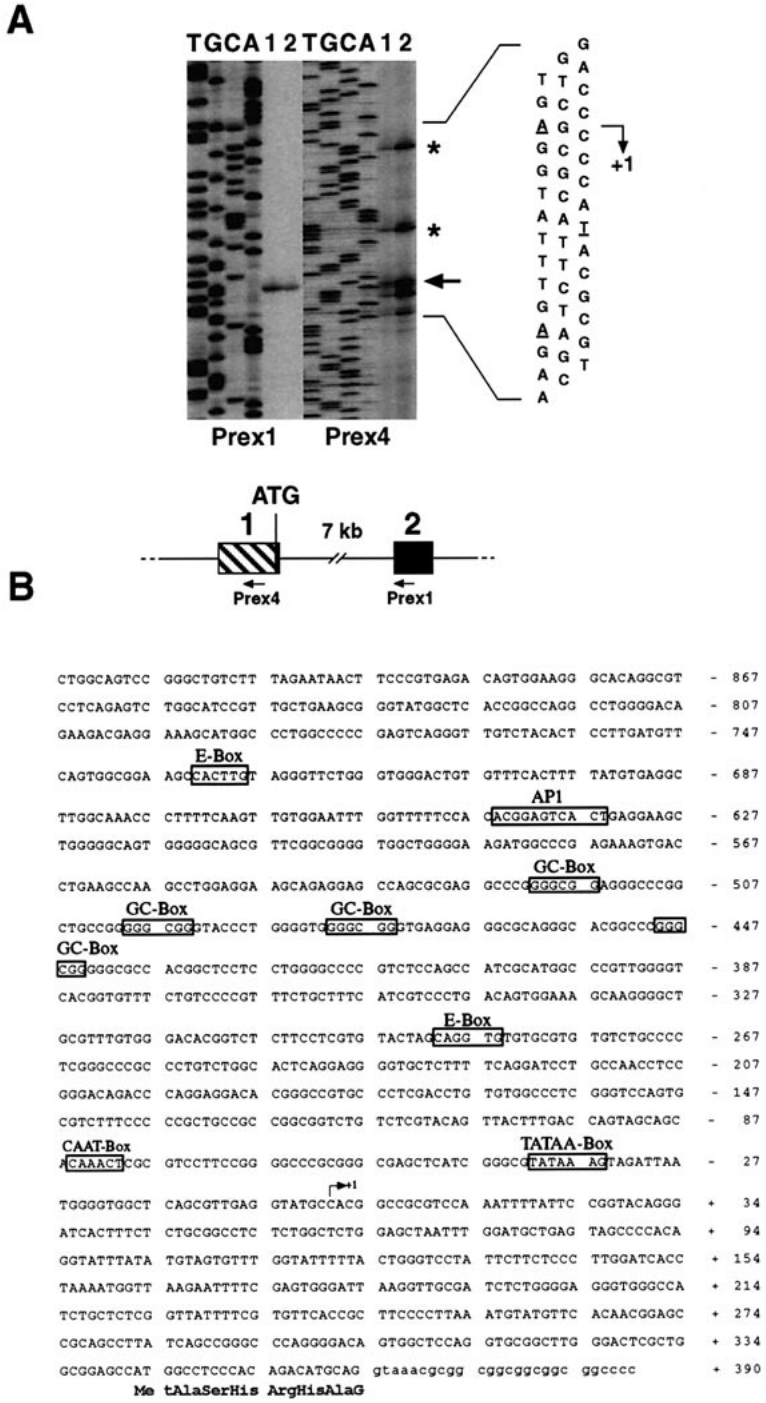

Fig. 4. Primer extension analysis of the FHL1C promoter. (**A**) A detail of the primer extension analysis using Fprex1 (exon 2) and Fprex4 (exon 1) is shown. Both primers resulted in the same elongation product (arrow). Two further products were obtained with Fprex4 indicated by asterisks. The reference DNA sequence was generated using both primers, respectively. The DNA sequence flanking the transcription start site (+1) is shown on the right side. The diagram illustrates the location of the primers and the genomic organization of the region. (**B**). Nucleotide sequence of the porcine FHL1C 5′-flanking region. The transcription start site is indicated by an arrow (+1). Putative promoter elements are indicated.

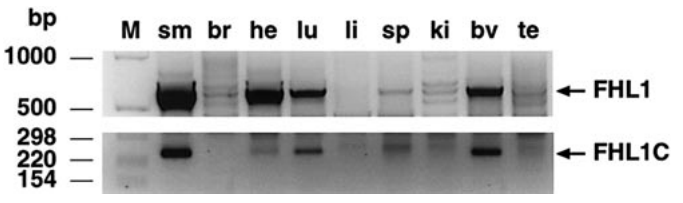

Fig. 5. Detection of FHL1 and FHL1C transcripts in different tissues. Amplicons of FHL1 transcripts in various porcine tissues are shown in the upper panel. RT-PCR was done as described using primers F-RTEx6 and F-RTEx4. The arrow indicates the FHL1 specific fragment. The lower panel shows amplicons of FHL1C transcripts. RT-PCR reactions were performed with primers Ftrans-rev, Ftrans-rev/Ftrans-fwd2 and Ftrans-rev/Ftrans-fwd3 as described. FHL1C specific products are indicated by an arrow. Sm: skeletal muscle; br: brain; he: heart; lu: lung; li: liver; sp: spleen; ki: kidney; bv: blood vessel; te: testis.

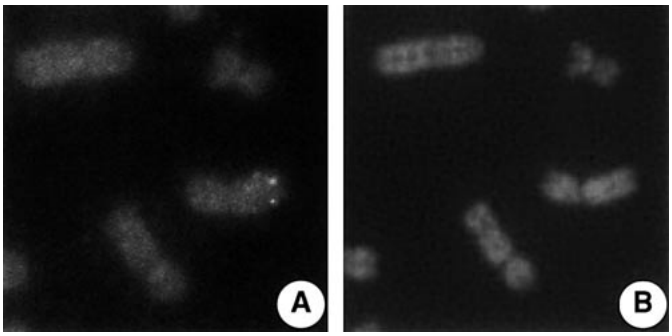

Fig. 6. Chromosome assignment of the FHL1 gene by FISH. Metaphase spreads were prepared and hybridized as described in Materials and methods. (**A**) Signals were detected on both chromatids at Xq22→q25. (**B**) Q-banding of the same metaphase spread.

```
                                                                CACGGCCGCGTCCAAATTTTATTCCGGTACAGGGAT
CACTTTCTCTGCGGCCTCTCTGGCTCTGGAGCTAATTTGGATGCTGAGTAGCCCCACAGGTATTTATATGTAGTGTTTGGTATT
TTTACTGGGTCCTATTCTTCTCCCTTGGATCACCTAAAATGGTTAAGAATTTTCGAGTGGGATTAAGGTTGCGATCTCTGGGGA
GGGTGGGCCATCTGCTCTCGGTTATTTTCGTGTTCACCGCTTCCCCTTAAATGTATGTTCACAACGGAGCCGCAGCCTTATCAG
CAGCCGGGCCCAGGGGACAGTGGCTCCAGGTGCGGCTTGGGACTCGCTGGCGGAGCCATGGCCTCCCACAGACATGCAGGCCCC
                                                                M  A  S  H  R  H  A  G  P

TCCAGCTACAAGGTAGGCACCATGGCCGAGAAGTTTGACTGCCACTACTGCAGGGACCCGCTGCAGGGAAAGAAGTACGTGCAG
S  S  Y  K  V  G  T  M  A  E  K  F  D  C  H  Y  C  R  D  P  L  Q  G  K  K  Y  V  Q

AAGGACGGCCACCACTGCTGCCTGAAGTGCTTCGACAAGTTCTGCGCCAACACCTGCGTGGAGTGCCGCAAACCCATCGGCGCT
K  D  G  H  H  C  C  L  K  C  F  D  K  F  C  A  N  T  C  V  E  C  R  K  P  I  G  A

GACTCCAAGGAGGTGCACTACAAGAACCGCTACTGGCACGACACCTGCTTCCGCTGCTCCAAGTGCCTTCACCCCTTGGCCAAT
D  S  K  E  V  H  Y  K  N  R  Y  W  H  D  T  C  F  R  C  S  K  C  L  H  P  L  A  N

GAGACCTTTATGGCCAAGGACAACAAGATCCTGTGCAACAAGTGCACCACTCGCGAGGATTCCCCCAAGTGCAAGGGGTGCTTC
E  T  F  M  A  K  D  N  K  I  L  C  N  K  C  T  T  R  E  D  S  P  K  C  K  G  C  F

AAGCCCATCGTGGCAGGCGATCAGAACGTGGAGTACAAGGGCACCGTCTGGCACAAAGACTGCTTCACCTGCAGCAACTGCAAG
K  P  I  V  A  G  D  Q  N  V  E  Y  K  G  T  V  W  H  K  D  C  F  T  C  S  N  C  K

CAAGTCATTGGGACAGGAAGCTTCTTCCCCAAAGGGGAGGACTTCTACTGCGTGACTTGCCATGAGACCAAGTTTGCCAAGCAC
Q  V  I  G  T  G  S  F  F  P  K  G  E  D  F  Y  C  V  T  C  H  E  T  K  F  A  K  H

TGCGTGAAGTGCAACAAGGCCATCACATCTGGAGGAATCACTTACCAGGATCAGCCCTGGCATGCCGAGTGCTTTGTGTGTGTT
C  V  K  C  N  K  A  I  T  S  G  G  I  T  Y  Q  D  Q  P  W  H  A  E  C  F  V  C  V

ACCTGCTCTAAGAAGCTGGCTGGGCAGCGTTTCACCGCTGTGGAGGACCAGTATTACTGCGTGGATTGCTACAAGAACTTTGTG
T  C  S  K  K  L  A  G  Q  R  F  T  A  V  E  D  Q  Y  Y  C  V  D  C  Y  K  N  F  V

GCCAAGAAGTGTGCTGGATGCAAGAACCCCATCACTGGGTTTGGTAAAGGCTCCAGTGTGGTGGCCTATGAAGGACAATCCTGG
A  K  K  C  A  G  C  K  N  P  I  T  G  F  G  K  G  S  S  V  V  A  Y  E  G  Q  S  W

CACGACTACTGCTTCCACTGCAAAAAATGCTCCGTGAATCTGGCCAACAAGCGCTTTGTTTTCCATCAGGAGCAAGTGTATTGC
H  D  Y  C  F  H  C  K  K  C  S  V  N  L  A  N  K  R  F  V  F  H  Q  E  Q  V  Y  C

CCCGACTGTGCCAAAAAGCTGTAAAGTGACAGGGGCTCCTGTCCTGTAAAATGGAATTTGAGTCTTGTTCCTTGTGTCCTCGCC
P  D  C  A  K  K  L

CTCTGCCCCGTACCACCCCTGGGGCCAGAGAGGCCTTTTCACCTCTCCGAGTTGCTCCTCCTGTCCTTCTTCCCATTTTACAGT
TACGCCTCAAGTACGGGCACGCAGTGATCATGTTAGGATTTAGCGAAAAGCAACCCTGCAGCAACGTTACTGTCCTGATAGCTG
CAGTTAGCAAACGAACACGTAGCTAGATTGACTCTTCTGCATGTTTATCCTAGAGCAGAAAAGCGCCGACCATTTAGCTGCTTA
GTGATGTAAGCAAGAAGCCTAAGAGATAAAGCCCCCACTGAGATGCCTCTCGTGGCTCAGCTGGGACCCACCGTGTCCGTCACA
CGACACACGCGAGAGTTGCAGCGGCTGCTCCAACTCGGCTGCTCACCCTGTCCTGTGAGCAGCAGGAACTTACCGAGATGCATG
GTTTAACTTCCCCATCAACACCTTCCCTCTGTTCTTTTGTGCTTTCACACCACTAATACGGATTTCCAGAAAACTAACATTTGA
ATTTAGCTGTAATTCTAACCGACCTTTCCCCTGTACTAACGTTGGATTTGGATTTGTGGCGTGCTTTCGGAGCGTTCCTCCTGC
CAAGCATGGAACACGCAAGGCGATTGGGAGAGTGTAGGCAGGTCTGAGAAAACGAGCCTGTGTCACCAGAATGTCGTCACAGCG
AATACTTCTGGAAGCTTAACAAAACTAACCCAGCTGTCCTTTATCTTTTTTTAATGAATAATGTTTTGTTTTAATGAATAGCAG
CATAGTTTATGGGTTTGGAGACTTGCATGAGAATGTTTTCGCCCCCTCAGACGTTCCTGCGGTTTTGCAATTTACCCTCCAGAA
GTAACCATAAAACCTGAGAGGGGCAGGTGAGCAGGTGCCAGGACTGTCACTGACATGGACATGCCGTTTCCAGACATGGCAGAG
TGGGCTTCCCCCTAACCGTCGGCGGCGGCTGCTCTTGTGAGTCTTAGGACCGTTCCTCAAGAGCTTACAGCTGGATCGCTGGGG
GAGCGGGGAGGCCACGGTCCCGGCAGGCAGAACCGGATGTCCACGTAGCGCTTACCTGAGACGCAGGAGCACCTCCTTACTGTA
TTCTAGCTACACTGTTATATTACATAGCATATCCCGAGACGACTTCAAGAGTACACACGTAATGACAGCGCGTATTAACATTCT
CGTAGGAGTGGTTAGAGAAGCCGAAGCCTCATTGCTACATTTTGTCATTAGCTTTGATCATCTAACGTTTCAGTGTATCCTTAC
GGAAATAAAGCAGCATATAAAG
```

Fig. 7. Nucleotide and amino acid sequence of porcine FHL1C. The conserved cyteines (C) and histidines (H) of the four LIM domains and the GATA like zinc finger are circled. The start and the stop codon are boxed, the polyadenylation signal is underlined.

3′-UTR in man and a slightly longer 3′-UTR in pig with 1,345 bp. Interestingly, the porcine and human 3′-UTRs show a high degree of homology which might indicate the presence of elements important for transcriptional regulation and mRNA stability (Decker and Parker, 1995). The longest murine FHL1 cDNA (Morgan et al., 1999) has an ORF of 840 bp and shows 93% identity with the porcine cDNA sequence. The 3′-UTR has a length of 1,402 bp and is therefore longer than in man and pig.

When we analyzed the 5′-end of the FHL1 gene using RACE PCR, we were able to identify a new variant, designated FHL1C. The FHL1C variant is characterized by an additional inframe start codon 48 bp upstream of the common FHL1 transcription start site leading to the addition of 16 amino acids at the N-terminus (Fig. 7). This new start codon was confirmed by sequencing the corresponding genomic region and was also found in murine and human EST database entries. Although the 16 amino acids do not contain any known regulatory pro-

tein motifs, it is interesting to note that FHL1, which is mainly detected in the cytoplasm, was also found in the nucleus of some cells (Brown et al., 1999; Lee et al., 1999). This could be an indication of different functional properties of the two variants. A similar disagreement in the 5′-UTR was reported for another member of the FHL-family. The amino acid sequences of FHL2 and the LIM protein DRAL are almost identical, however their N-terminal regions differ significantly (Chan et al., 1998). Up to now it is believed that both proteins are encoded by two different genes. Based on the results in this study, the presence of two different genes should be reconsidered.

To characterize the genomic organization of the FHL1 gene, we isolated a genomic phage and a PAC clone. 14,370 bp were sequenced harboring 6 exons of the FHL1C cDNA. Additionally, the alternatively spliced exon 5b of FHL1B/SLIMMER was identified (Brown et al., 1999; Lee et al., 1999). The analyzed porcine genomic region does not contain any sequences that code for the 5′-parts of all other FHL1 variants. However, a recently deposited human BAC sequence (RP11-535K18, Acc. No. AL078638) spans the entire chromosome region Xq26.2→q27.1. Within this sequence we could identify all known exons of FHL1, including the presumptive first exon of the human FHL1C variant. Additionally, two exons further upstream (26 and 52 kb, respectively) code for the 5′UTRs of other FHL1 variants. Therefore the human FHL1 gene spans a genomic region of at least ~50 kb, and FHL1C is a splicing variant of the FHL1 gene. Further evidence that FHL1C is transcribed from the FHL1 locus and not from another gene nearby comes from Southern blot analysis of the isolated porcine PAC clone. Hybridization with a FHL1 specific probe showed only one specific band when the whole PAC insert (>150 kb) was cut with several restriction enzymes (data not shown). Additionally, in FISH experiments only a single signal on chromosome X could be identified.

The overall structure of the human and porcine gene seems to be conserved. Especially well conserved are the intron regions flanking exons 4, 5a, and 5b. These exons are involved in alternative splicing and the conserved parts might play a functional role. However, there are also some areas of intron 1 that share significant homology and therefore might contain important regulatory elements for transcriptional control (Liska et al., 1992; Gaunitz et al., 1997).

The transcription start site of FHL1C was mapped to a CA dinucleotide by primer extension analysis. The transcription start site (CCACGGC) matches the consensus sequence (YCYNYYY) in five out of seven base pairs. Analysis of upstream sequences identified a TATAA-box and a putative CAAT-box within distances that are typical for eukaryotic promoters (Penotti, 1990). We analyzed 947 bp upstream of the transcription start site and found binding sites and core motifs for general and tissue specific transcription factors. In the corresponding human region we also identified a TATAA-box at a similar position, however the CAAT-box was not conserved (CAAAAT). Core sequences for E-boxes were also found. Hence, this region shows all characteristics of a promoter and may regulate the expression of FHL1C. However, so far we cannot exclude the possibility that there is a promoter further upstream. At least the gene variants FHL1, FHL1B/SLIMMER and KyoT2 must be regulated by a promoter upstream from their first two exons. In our RACE experiments we never observed products that contained both 5′-UTR sequences and all dbEST entries either show the FHL1C 5′-UTR or the one used for all other known variants. It is therefore likely that FHL1C is regulated from an alternative promoter.

Alternative promoters are found in several genes that show tissue specific expression (O'Mahoney et al., 1994; Aizencang et al., 2000). Our RT-PCR experiments using FHL1C specific primers demonstrated a tissue specific expression pattern in skeletal muscle and tissues from blood vessels. Transcripts were also identified in heart and lung.

Except for 16 residues at its N-terminus the amino acid sequence of FHL1C is the same as in FHL1. The pig, man and mouse FHL1Cs consist of 296 amino acids and show only a few amino acid exchanges when compared to each other. The porcine FHL1C is 98% identical to the human, and 93% identical to the mouse protein. Most of the amino acid exchanges occur in the first LIM domain and the linker region between the first and second LIM domain. In the second LIM domain an alanine in man and mouse is changed to a proline in pig. The third LIM domain is entirely conserved between pig and mouse and the fourth LIM domain between pig and man. The high degree of conservation might be an indication that FHL1 and FHL1C are phylogenetic important proteins. They only consist of four LIM domains, a GATA like zinc finger and the linker between two adjacent LIM domains composed of eight amino acids. So far the function of both proteins is unknown and it is not clear whether all zinc fingers are functionally important, but if this is the case already subtle changes in protein sequence might not be tolerated.

In conclusion, we have isolated a new gene variant of FHL1 from a porcine skeletal muscle cDNA library. FHL1C differs from FHL1 in 16 amino acids at the N-terminus. These 16 amino acids, together with the 5′-UTR, are encoded by a newly identified exon in the FHL1 gene. The sequence of the FHL1C cDNA is located within a 14-kb genomic region and is arranged in six exons. From primer extension analysis there is evidence that transcription of the FHL1C variant is regulated by an alternative promoter. This data is in agreement with the tissue specific expression pattern of FHL1C.

Acknowledgements

The authors are grateful to A. Deppe for expert technical assistance.

References

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K: Current protocols in molecular biology. (John Wiley & Sons, New York 1990).

Aizencang GI, Bishop DF, Forrest D, Astrin KH, Desnick RJ: Uroporphyrinogen III synthase. An alternative promoter controls erythroid-specific expression in the murine gene. J biol Chem 275:2295–2304 (2000).

Al-Bayati HK, Duscher S, Kollers S, Rettenberger G, Fries R, Brenig B: Construction and characterization of a porcine P1-derived artificial chromosome (PAC) library covering 3.2 genome equivalents and cytogenetical assignment of six type I and type II loci. Mammal Genome 10:569–572 (1999).

Altermann E, Klein JR, Henrich B: Synthesis and automated detection of fluorescently labeled primer extension products. Biotechniques 26:96–98, 101 (1999).

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: Basic local alignment search tool. J molec Biol 215:403–410 (1990).

Braun T, Buschhausen-Denker G, Bober E, Tannich E, Arnold HH: A novel human muscle factor related to but distinct from MyoD1 induces myogenic conversion in 10T1/2 fibroblasts. EMBO J 8:701–709 (1989).

Brown S, McGrath MJ, Ooms LM, Gurung R, Maimone MM, Mitchell CA: Characterization of two isoforms of the skeletal muscle LIM protein 1, SLIM1. Localization of SLIM1 at focal adhesions and the isoform slimmer in the nucleus of myoblasts and cytoplasm of myotubes suggests distinct roles in the cytoskeleton and in nuclear-cytoplasmic communication. J biol Chem 274:27083–27091 (1999).

Chan KK, Tsui SK, Lee SM, Luk SC, Liew CC, Fung KP, Waye MM, Lee CY: Molecular cloning and characterization of FHL2, a novel LIM domain protein preferentially expressed in human heart. Gene 210:345–350 (1998).

Chevalet C, Gouzy J, SanCristobal-Gaudy M: Regional assignment of genetic markers using a somatic cell hybrid panel: a WWW interactive program available for the pig genome. Comput Appl Biosci 13:69–73 (1997).

Davis RL, Weintraub H, Lassar AB: Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell 51:987–1000 (1987).

Decker CJ, Parker R: Diversity of cytoplasmic functions for the 3′ untranslated region of eukaryotic transcripts. Curr Opin Cell Biol 7:386–392 (1995).

Edmondson DG, Olson EN: A gene with homology to the myc similarity region of MyoD1 is expressed during myogenesis and is sufficient to activate the muscle differentiation program. Genes Dev 3:628–640 (1989).

Gaunitz F, Gaunitz C, Papke M, Gebhardt R: Cis-regulatory sequences from the first intron of the rat glutamine synthetase gene are involved in hepatocyte specific expression of the enzyme. Biol Chem 378:11–18 (1997).

Genini M, Schwalbe P, Scholl FA, Remppis A, Mattei MG, Schafer BW: Subtractive cloning and characterization of DRAL, a novel LIM-domain protein down-regulated in rhabdomyosarcoma. DNA Cell Biol 16:433–442 (1997).

Greene WK, Baker E, Rabbitts TH, Kees UR: Genomic structure, tissue expression and chromosomal location of the LIM-only gene, SLIM1. Gene 232:203–207 (1999).

Heinemeyer T, Chen X, Karas H, Kel AE, Kel OV, Liebich I, Meinhardt T, Reuter I, Schacherer F, Wingender E: Expanding the TRANSFAC database towards an expert system of regulatory molecular mechanisms. Nucl Acids Res 27:318–322 (1999).

Kozak M: Interpreting cDNA sequences: some insights from studies on translation. Mammal Genome 7:563–574 (1996).

Krempler A, Brenig B: Zinc finger proteins: watchdogs in muscle development. Mol Gen Genet 261:209–215 (1999).

Lee SM, Li HY, Ng EK, Or SM, Chan KK, Kotaka M, Chim SS, Tsui SK, Waye MM, Fung KP, Lee CY: Characterization of a brain-specific nuclear LIM domain protein (FHL1B) which is an alternatively spliced variant of FHL1. Gene 237:253–263 (1999).

Lee SM, Tsui SK, Chan KK, Garcia-Barcelo M, Waye MM, Fung KP, Liew CC, Lee CY: Chromosomal mapping, tissue distribution and cDNA sequence of four-and-a-half LIM domain protein 1 (FHL1). Gene 216:163–170 (1998).

Lee SM, Tsui SK, Chan KK, Kotaka M, Li HY, Chim SS, Waye MM, Fung KP, Lee CY: Chromosomal mapping of a skeletal muscle specific LIM-only protein FHL3 to the distal end of the short arm of human chromosome 1. Somat Cell Mol Genet 24:197–202 (1998).

Liska DJ, Robinson VR, Bornstein P: Elements in the first intron of the alpha 1(I) collagen gene interact with Sp1 to regulate gene expression. Gene Expr 2:379–389 (1992).

Miner JH, Wold B: Herculin, a fourth member of the MyoD family of myogenic regulatory genes. Proc natl Acad Sci, USA 87:1089–1093 (1990).

Morgan MJ, Madgwick AJ: Slim defines a novel family of LIM-proteins expressed in skeletal muscle. Biochem biophys Res Commun 225:632–638 (1996).

Morgan MJ, Madgwick AJ: The fourth member of the FHL family of LIM proteins is expressed exclusively in the testis. Biochem biophys Res Commun 255:251–255 (1999).

Morgan MJ, Madgwick AJ: The LIM proteins FHL1 and FHL3 are expressed differently in skeletal muscle. Biochem biophys Res Commun 255:245–250 (1999).

O'Mahoney JV, Brandon MR, Adams TE: Identification of a liver-specific promoter for the ovine growth hormone receptor. Mol Cell Endocrinol 101:129–139 (1994).

Penotti FE: Human DNA TATA boxes and transcription initiation sites. A statistical study. J molec Biol 213:37–52 (1990).

Rhodes SJ, Konieczny SF: Identification of MRF4: a new member of the muscle regulatory factor gene family. Genes Dev 3:2050–2061 (1989).

Sanchez-Garcia I, Rabbitts TH: The LIM domain: a new structural motif found in zinc-finger-like proteins. Trends Genet 10:315–320 (1994).

Solinas-Toldo S, Mezzelani A, Hawkins GA, Bishop MD, Olsaker I, Mackinlay A, Ferretti L, Fries R: Combined Q-banding and fluorescence in situ hybridization for the identification of bovine chromosomes 1 to 7. Cytogenet Cell Genet 69:1–6 (1995).

Taniguchi Y, Furukawa T, Tun T, Han H, Honjo T: LIM protein KyoT2 negatively regulates transcription by association with the RBP-J DNA-binding protein. Mol Cell Biol 18:644–654 (1998).

Toldo SS, Fries R, Steffen P, Neibergs HL, Barendse W, Womack JE, Hetzel DJ, Stranzinger G: Physically mapped, cosmid-derived microsatellite markers as anchor loci on bovine chromosomes. Mammal Genome 4:720–727 (1993).

Wright WE, Sassoon DA, Lin VK: Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD. Cell 56:607–617 (1989).

Yerle M, Echard G, Robic A, Mairal A, Dubut-Fontana C, Riquet J, Pinton P, Milan D, Lahbib-Mansais Y, Gellin J: A somatic cell hybrid panel for pig regional gene mapping characterized by molecular cytogenetics. Cytogenet Cell Genet 73:194–202 (1996).