A Strategic Shot-Gun Approach to Cloning of Isoform-Specific Myosin cDNAs

Kin-Chow Chang[(1)] and Kenneth Fernandes

Division of Development and Reproduction, Roslin Institute (Edinburgh), Roslin, Midlothian, U.K.

Abstract

One key feature of myosins, whether muscle (smooth and striated) or non-muscle (I and II) myosin, is their ability to hydrolyse ATP in an actin-dependent fashion. Based on the ATP-binding site, which is highly conserved in all known myosins, we described here a strategic approach that allows the shot-gun isolation of a comprehensive range of 5' end myosin cDNAs from virtually any myosin-containing tissue. This proposed strategy has a number of distinct advantages over other existing PCR cloning approaches. By cloning the 5' untranslated region, specific myosin isoforms can readily be distinguished and used as isoform-specific probes. Additionally, this approach facilitates locating the transcriptional start site on the corresponding genomic clone, allows for the isolation of overlapping full length cDNA clones and has the potential of finding novel ATP-binding cDNAs. Using this strategy, we showed here the isolation of 3 sarcomeric porcine myosin heavy chain (MyHC) cDNA isoforms from skeletal muscles, one of which was a novel developmentally-regulated skeletal muscle isoform, not previously reported in the pig.

Key words: myosin, heavy chain, shot-gun, cloning, cDNAs, PCR, library.

BAM 6 (5): 351-356, 1996

Myosins are a superfamily of actin filament-based motors, thought to be involved in a variety of motile processes. Myosin II is a two-headed filament forming structure with a coiled-coil tail. It is predominantly present in striated and smooth muscles as muscle myosin, and in non-muscles as conventional myosin. Myosin I or unconventional myosins, on the other hand, have structures that resemble the head of myosin II but lack the coiled-coil tail. Based on sequence alignment of this head motor region, nine classes of unconventional myosins have been recently identified [2, 5, 6, 8, 11]. Each class of unconventional myosin has a distinctive tail region, which is thought to confer specific motility function such as cell motility, movement of organelles and membrane recycling [e.g. 1, 7, 12]. One key feature of myosins, whether muscle or non-muscle myosin, is their ability to hydrolyse ATP in an actin-dependent fashion. The corresponding amino acid sequence spanning the ATP-binding site of all known myosins is highly conserved and is located in the head region of the molecule. We described here a strategic approach that allows the shot-gun isolation of a comprehensive range of 5' end myosin cDNAs from virtually any myosin-containing tissue. This approach has the advantages of being able to readily distinguish between different myosin isoforms, based on their 5' untranslated regions, and to predict the likely location of the transcriptional start site on the corresponding gene. Using this strategy, we showed here the isolation of 3 sarcomeric porcine myosin heavy chain (MyHC) cDNA isoforms from skeletal muscles, one of which was a novel isoform not previously reported in the pig.

Materials and Methods

λ *cDNA library synthesis*

Messenger RNA was directly isolated from the *gastrocnemius* of a 2-day-old Large white pig (Quick-Prep mRNA purification kit, Pharmacia). Randomly primed cDNA synthesis was performed on the isolated mRNA (cDNA synthesis kit, Stratagene) followed by bi-directional cloning into the λ vector λExCell (Pharmacia). In this way, a randomly-primed muscle-specific porcine λ cDNA library was made. It was amplified once [13] and total library DNA was isolated (Lambda kit, Qiagen); the quality of which was assessed by PCR using sets of β-actin and porcine slow/β MyHC primers [3].

Polymerase chain reaction (PCR)

The shot-gun approach to the cloning of myosin cDNAs relied on the highly conserved ATP-binding site (Figure 1). Two degenerate primers, based on the GESGAGKT and EAFGNAKT regions which are conserved in all known vertebrate myosins [2], were synthesized in the antisense direction (GTYTTNCCNG CNCCNGAYTC NCC and GTYTTNGCRA ANGCYTC, respectively). Primary PCR was performed on total library DNA, using the λ vector T7- and EAFGNAKT-primers. One µg of total λ DNA was used as template per reaction in a total reaction volume of 50 µl, containing 100 pMoles of each primer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs and 5 U Taq polymerase. PCR conditions used were 35 cycles of denaturation (55 s at 94°C), annealing (120 s at 53°C) and polymerisation (90 s at 70°C). DNA products, ranging from 0.4 to 0.7 kb, were isolated from a 0.8% agarose gel and purified using the glassbeads method (Geneclean, Bio101). A second PCR was performed on the purified first round PCR products using the same T7 primer and the internal GESGAGKT primer. Second round PCR products in the 0.4 to 0.7 kb range were again isolated from an agarose gel and cloned into a plasmid vector (TA cloning kit, Invitrogen).

Southern-, Northern hybridisation and sequencing of PCR products

Southern hybridisation was performed on the cloned PCR products using two sarcomeric-specific MyHC oligonucleotides, collectively, as probes, which are highly conserved in sarcomeric myosins (5'- CCCCACATCT TCTCCATCTC CG-3' and 5'-GACAGAGAAA ACCAGTCCAT CC-3') and are located 5' to the ATP-binding site sequence. Sarcomeric-specific MyHC positive clones, whose inserts were larger than 0.5 kb, were sequenced by the dideoxy-chain termination method [14],

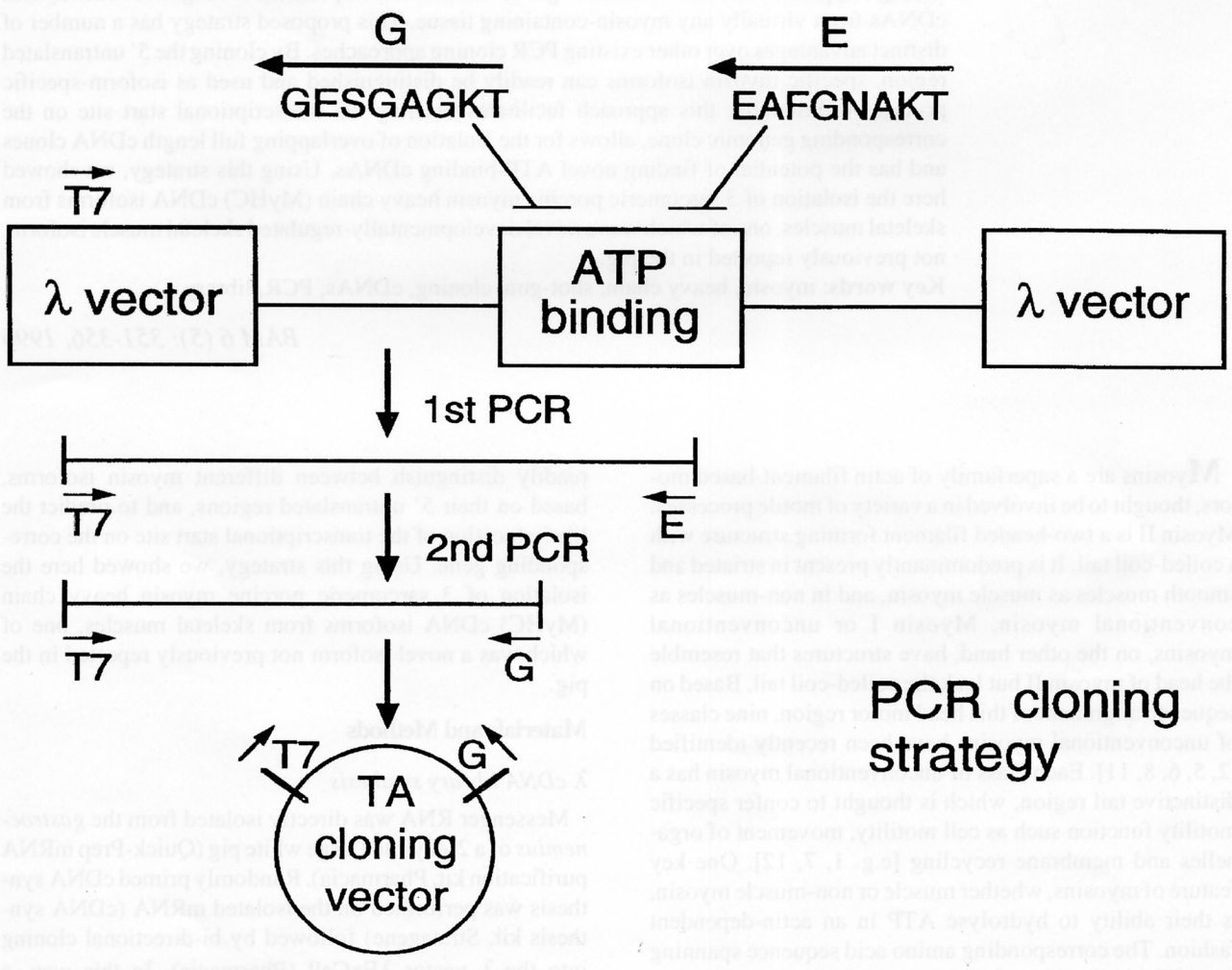

Figure 1. Schematic outline of the shot-gun cloning strategy in the isolation of 5' end myosin cDNAs. Total library DNA was used as template for PCR, using a T7-vector specific primer and primer G. T7-vector specific primer was the sense, upstream primer whereas G and E were degenerate, antisense primers derived from the highly conserved ATP-binding region. After the second round PCR with T7-vector specific primer and primer E, PCR products in the range of 0.4 to 0.7 kb were cloned into a plasmid vector (TA-cloning kit, Invitrogen) for subsequent analysis.

using a plasmid sequencing kit (Amersham-USB). For genomic Southern blotting, genomic pig DNA was isolated as described in [13] and for Northern blotting, total RNA was isolated from various porcine tissues with a commercial kit (RNeasy Total RNA Kit, Qiagen). Both blots were hybridised with a 78 bp probe, derived from the 5' untranslated region of a novel MyHC cDNA isoform (Figure 4). All hybridisations were performed on charged nylon membranes (Hybond-N+, Amersham) using Rapid-hyb (Amersham) hybridisation buffer.

Results and Discussion

Shotgun cloning of 5' end MyHC cDNAs from 2-day-old pig skeletal muscle

A 2-day-old pig skeletal muscle (*gastrocnemius*) λEx-Cell-cDNA library was first generated. Total library DNA was extracted and its quality assessed by performing PCR using sets of β-actin and porcine slow/β primers. The predicted PCR fragment sizes were correctly obtained (data not shown), suggesting that a representative library was produced. Figure 1 shows a schematic outline of the 5' end MyHC cloning steps on total muscle library DNA (see also Materials and Methods). After the second round PCR, DNA fragments in the range of 0.4 to 0.7 kb were cloned into the plasmid vector (TA cloning kit, Invitrogen). About 80% of the plasmid clones were found to carry an insert. Southern hybridisation using sarcomeric-specific oligonucleotides as probes revealed that 60% of the inserts were of the sarcomeric myosin-type (Figure 2). It is estimated that, based on published data [9, 10], the distance from the ATP-binding site to the transcriptional start site is approximately 0.7 kb. Based on this assumption, about half the sarcomeric positive inserts seemed to possess the most 5' ends of their corresponding messages.

Classification of myosin heavy chain cDNA clones

Over 30 sarcomeric MyHC clones were sequenced from both ends. Through sequence homology, 3 types of MyHC isoforms were identified. About 10% of the clones were found to belong to the slow/β isoform [3], 20% corresponded to the fast IIA isoform [4] and the remaining 70% belonged to a novel isoform, not previously reported in the pig. Figure 3 shows the DNA sequence homology of part of the 5' translated regions of the 3 PCR-derived MyHC isoforms. The novel isoform is more homologous to the MyHC IIA than to the slow/β isoform. Their corresponding 5' untranslated regions are least homologous to each other (data not shown) and are candidates for isoform-specific MyHC probes [3, 4].

Demonstration of a novel porcine skeletal MyHC isoform

The longest PCR-derived insert, identified as the novel MyHC isoform, had a 78 bp untranslated region 5' to the putative translational start site. This 78 bp region was amplified by PCR and cloned into a plasmid vector (TA cloning kit, Invitrogen). To determine the potential use of this region as an isoform-specific probe, genomic Southern

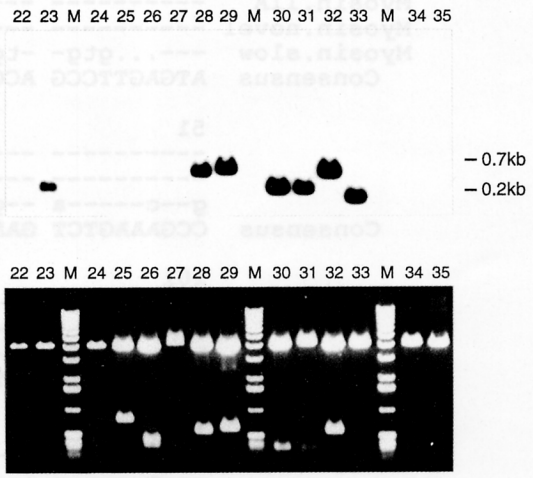

Figure 2. Screening for myosin heavy chain (MyHC)-specific clones. Cloned second round PCR products were excised from its vector by endonuclease restriction with EcoRI and separated in a 0.8% agarose gel (Lower panel). About 80% of clones were found to contain an insert. Southern hybridisation was performed on the same gel, using two MyHC-specific oligonucleotides, collectively, as probes (Top panel). About 60% of the inserts were MyHC-specific.

hybridisation was performed (Figure 4). Only a single banding pattern was found for each genomic digest, suggesting that this 78 bp fragment is likely to hybridise specifically only to its own isoform. Next, a Northern hybridisation was performed on various porcine tissues, using the same probe (Figure 5). The results showed that this novel MyHC isoform was expressed in the skeletal muscles (*biceps femoris*) of a neonatal pig, but no expression was found in non-skeletal muscle tissues of the same animal. Hence this novel isoform, unlike the slow/β isoform, is a skeletal muscle MyHC (skMyHC) isoform. In addition, no detectable expression of this skMyHC isoform was found in the back muscles (*longissimus dorsi*) of a 50-day-old foetus, suggesting that this developmentally-regulated skMyHC isoform is first switched on either perinatally or postnatally. Further *in vivo* expression studies, involving the use of *in situ* hybridisation, are needed to define the isoform type of this skMyHC.

Other PCR-derived non-MyHC clones

About 40% of PCR-derived clones (Figure 1) were not sarcomeric myosins. A more limited sequencing exercise was conducted on these clones. In 15 clones examined, no non-muscle myosin was found. However, based on sequence homology, a number of known ATP-binding clones were identified, which included genes encoding mitochondria ATPase synthase and valosin-containing protein. A partial porcine cDNA homologue of the nebulin gene was also found. At this stage, it is not clear whether

```
                1                                                   50
Myosin.IIA      ---------- ---------- ---------- ---------- ----------
Myosin.novel    ---------- ---------- ---t--t--- ---------- -t--------
Myosin.slow     ---...gtg- -tgc---g-- ----gc---- --------c- ----c-----
  Consensus     ATGAGTTCCG ACCAGGAAAT GGCCATATTT GGGGAGGCTG CCCCTTACCT

                51                                                  100
                ---------- ---------- ---------- ---------- --c-------
                ---------- ---------- ---------- --------aa ----------
                g--c-----a --g------- -gc-g--a-- -----cc--- -----c--cc
  Consensus     CCGAAAGTCT GAAAAGGAGC GCATTGAGGC CCAGAATAGG CCTTTTGATG

                101                                                 150
                -------a-- ---------- --a------- ---------- ----------
                ----a----- ---------- --g------- ---------- ----------
                t-----a-ga t----a---t c-t--tga-- -a--ggag-- c-----g-cc
  Consensus     CCAAGACGTC AGTCTTTGTG GC-GAGCCCA AGGAATCCTT TGTCAAAGGG

                151                                                 200
                ---------- ---------- ---------- ---------- -------a--
                ---g------ ---------- ---------- ---------- -------c--
                -ag--ttt-t ctc----g-- t--c-----c --c-ctg--- -c--gcat--
  Consensus     ACTATCCAGA GCAGAGAAGG AGGGAAAGTG ACAGTGAAGA CGGAAGC-GG

                201                                                 250
                ---------- ---------- ---------- c--------- ----------
                ---------- ---------- ---------- ---------- ----------
                caa---gg-- --t-----g- -g-------- -c-g-agca- --t--a----
  Consensus     GGCGACTTTG ACAGTGAAAG AAGACCAGGT GTTCCCCATG AACCCTCCCA

                251                                                 300
                ---------- ---------- ---------- ---------- ---------g
                ---------- ---------- ---------- ---------- ----------
                -g--c----- ---------- ---------c ------tt--- ----------
  Consensus     AATTTGACAA GATCGAGGAC ATGGCCATGA TGACCCACCT GCACGAGCCC

                301                                                 350
                -gc------- ---------- ---------- ---------- ----------
                --t------- ---------- ---------- ---------- ----------
                --g-----c- ---------- g-----c--- --ct------ -------t--
  Consensus     GC-GTGCTGT ACAACCTCAA AGAGCGTTAC GCAGCCTGGA TGATCTACAC

                351                                                 400
                ---------- ---------- ---------- ---------- ----------
                ---------- ---------- ---------- ---------- ----------
                ------t--g ---------- ----aa---- ---------- ----------
  Consensus     CTACTCGGGC CTCTTCTGTG TCACCGTCAA CCCCTACAAG TGGCTGCCGG

                401                                                 450
                --------c- c--------- ---------- ---------- ----------
                ---------- a--------- ---------- ---------- ----------
                ---------- t-----c--- g-t------c -g-----g-- -a-gagc---
  Consensus     TGTACAACGC -GAGGTGGTG ACGGCCTACA GAGGCAAAAA GCGCCAGGAG
```

Figure 3. Sequence homology of the 5' translated regions of the 3 PCR-derived MyHC isoforms. The novel isoform is highly homologous to MyHC IIA [4]. The 5' untranslated region was most divergent in all 3 isoforms (data not shown).

this is an incidental finding or that nebulin has a consensus ATP-binding site. Additionally, about half the clones sequenced did not show significant homology with any sequence in the GenBank/EMBL database. Further work is needed to evaluate the functional importance of this group of unknown cDNAs.

Feasibility and versatility of the proposed shot-gun approach

The success of the proposed strategy in the shot-gun PCR cloning of a range of 5' end myosin cDNA isoforms depends on two main requirements. The first is the availability of total λ-cDNA-library DNA from the tissue of interest, in this case a 2-day-old porcine skeletal muscle λExCell-cDNA was generated. The second requirement is the use of the highly conserved ATP-binding site, from which two degenerate antisense primers (GESGAGKT and EAFGNAKT) are derived as downstream PCR primers. A randomly-primed cDNA library is necessary to

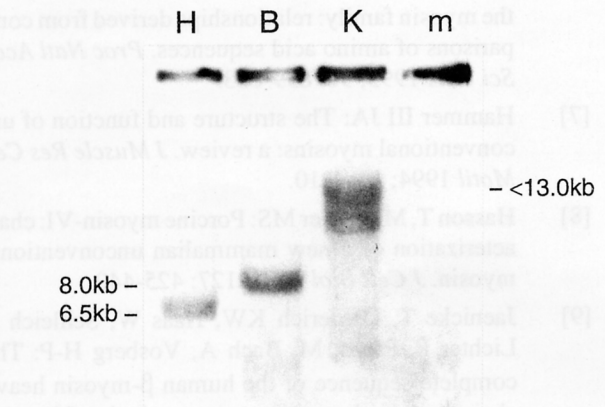

Figure 4. Evaluating the isoform-specificity of the 5' untranslated region of the novel MyHC isoform. A 78 bp fragment from the 5' untranslated region was used as a probe in Southern hybridisation on porcine genomic DNA (H = Hin dIII; B = Bam HI; K = Kpn I). Note only a single band pattern was found in each lane.

ensure that the 5' ends of myosins are properly represented. Oligo(dT) primed cDNAs are considered less suitable because myosin cDNAs are too long (about 6 kb) for representative full length cDNA synthesis. In our porcine skeletal muscle λExCell-cDNA library, cDNA inserts were cloned into the λ vector in either orientation. Though not a prerequisite, unidirectional cloning into the λ vector may improve on the PCR cloning efficiency (Figure 1), in that twice as much cDNAs are in the correct orientation with respect to the predetermined λ vector upstream primer. This proposed strategy has a number of distinct advantages over other existing PCR cloning approaches. Firstly, by cloning the 5' untranslated region, specific myosin isoforms can readily be distinguished. In our case 3 MyHC isoforms had been identified (Figure 3). Secondly, the 5' untranslated regions are useful for the generation of isoform-specific probes for *in vitro* analysis (Figures 4, 5) and for *in vivo* expression studies [3, 4]. Thirdly, having identified the most 5' untranslated sequence, locating the transcriptional start site on the corresponding genomic clone will be greatly facilitated. Fourthly, this approach has the potential for isolating clones coding for novel ATP-binding proteins. Finally, this approach is amenable to the isolation of full length, overlapping cDNA clones by subsequent library screening.

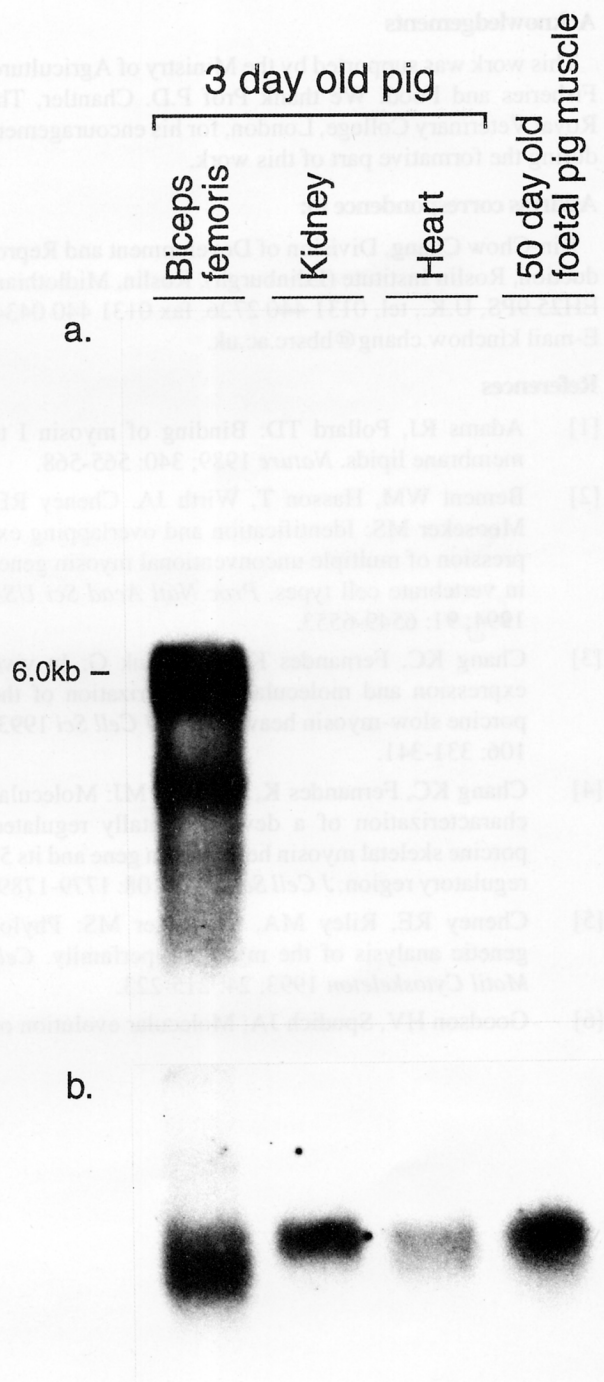

Figure 5. The novel porcine MyHC isoform is developmentally expressed skeletal muscles. Northern hybridisation was performed on total RNA from various porcine tissues using the 78 bp fragment (Figure 4) as probe. Lane 1 = biceps femoris of neonatal Meishan pig; lane 2 = kidney of neonatal Meishan pig; lane 3 = heart of neonatal Meishan pig; lane 4 = back muscle of a 50 day Meishan-Large white foetus. Results showed that the novel isoform is a skeletal muscle-specific isoform not expressed in early foetal development (Top panel). Lower panel shows the same blot hybridised with a β-actin probe to demonstrate RNA integrity in all lanes.

Acknowledgements

This work was supported by the Ministry of Agriculture, Fisheries and Food. We thank Prof P.D. Chantler, The Royal Veterinary College, London, for his encouragement during the formative part of this work.

Address correspondence to:

Kin-Chow Chang, Division of Development and Reproduction, Roslin Institute (Edinburgh), Roslin, Midlothian, EH25 9PS, U.K., tel. 0131 440 2726, fax 0131 440 0434, E-mail kinchow.chang@bbsrc.ac.uk.

References

[1] Adams RJ, Pollard TD: Binding of myosin I to membrane lipids. *Nature* 1989; 340: 565-568.

[2] Bement WM, Hasson T, Wirth JA, Cheney RE, Mooseker MS: Identification and overlapping expression of multiple unconventional myosin genes in vertebrate cell types. *Proc Natl Acad Sci USA* 1994; 91: 6549-6553.

[3] Chang KC, Fernandes K, Goldspink G: In vivo expression and molecular characterization of the porcine slow-myosin heavy chain. *J Cell Sci* 1993; 106: 331-341.

[4] Chang KC, Fernandes K, Dauncey MJ: Molecular characterization of a developmentally regulated porcine skeletal myosin heavy chain gene and its 5' regulatory region. *J Cell Sci* 1995; 108: 1779-1789.

[5] Cheney RE, Riley MA, Mooseker MS: Phylogenetic analysis of the myosin superfamily. *Cell Motil Cytoskeleton* 1993; 24: 215-223.

[6] Goodson HV, Spudich JA: Molecular evolution of the myosin family: relationships derived from comparisons of amino acid sequences. *Proc Natl Acad Sci USA* 1993; 90: 659-663.

[7] Hammer III JA: The structure and function of unconventional myosins: a review. *J Muscle Res Cell Motil* 1994; 15: 1-10.

[8] Hasson T, Mooseker MS: Porcine myosin-VI: characterization of a new mammalian unconventional myosin. *J Cell Biol* 1994; 127: 425-440.

[9] Jaenicke T, Diederich KW, Haas W, Schleich J, Lichter P, Pfordt M, Bach A, Vosberg H-P: The complete sequence of the human β-myosin heavy chain gene and a comparative analysis of its product. *Genomics* 1990; 8: 194-206.

[10] Karsch-Mizzachi I, Feghali R, Shows TB, Leinwand LA: Generation of a full-length human perinatal myosin heavy chain-encoding cDNA. *Gene* 1990; 89: 289-294.

[11] Knight AE and Kendrick-Jones J: A myosin-like protein from a higher plant. *J Mol Biol* 1993;231: 148-154.

[12] Li DQ, Miller M, Chantler PD: Association of a cellular myosin II with anionic phospholipids and the neuronal plasma membrane. *Proc Natl Acad Sci USA* 1994; 91: 853-857.

[13] Sambrook J, Fritsch EF, Maniais T: Molecular cloning: a laboratory manual, Cold Spring Habor, NY, 1989.

[14] Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA* 1977; 74: 5463-5467.