The Journal of Experimental Biology 215, 239-246

doi:10.1242/jeb.064055

RESEARCH ARTICLE

Muscular tissues of the squid *Doryteuthis pealeii* express identical myosin heavy chain isoforms: an alternative mechanism for tuning contractile speed

Justin F. Shaffer* and William M. Kier

Department of Biology, University of North Carolina, Chapel Hill, NC 27599, USA

*Author for correspondence (shafferj@unc.edu)

Accepted 17 October 2011

SUMMARY

The speed of muscle contraction is largely controlled at the sarcomere level by the ATPase activity of the motor protein myosin. Differences in amino acid sequence in catalytically important regions of myosin yield different myosin isoforms with varying ATPase activities and resulting differences in cross-bridge cycling rates and interfilamentary sliding velocities. Modulation of whole-muscle performance by changes in myosin isoform ATPase activity is regarded as a universal mechanism to tune contractile properties, especially in vertebrate muscles. Invertebrates such as squid, however, may exhibit an alternative mechanism to tune contractile properties that is based on differences in muscle ultrastructure, including variable myofilament and sarcomere lengths. To determine definitively whether contractile properties of squid muscles are regulated *via* different myosin isoforms (i.e. different ATPase activities), the nucleotide and amino acid sequences of the myosin heavy chain from the squid *Doryteuthis pealeii* were determined from the mantle, arm, tentacle, fin and funnel retractor musculature. We identified three myosin heavy chain isoforms in squid muscular tissues, with differences arising at surface loop 1 and the carboxy terminus. All three isoforms were detected in all five tissues studied. These results suggest that the muscular tissues of *D. pealeii* express identical myosin isoforms, and it is likely that differences in muscle ultrastructure, not myosin ATPase activity, represent the most important mechanism for tuning contractile speeds.

Supplementary material available online at http://jeb.biologists.org/cgi/content/full/215/2/239/DC1

Key words: myosin, squid, muscle, contraction, myofilament.

INTRODUCTION

Muscular performance, including contractile speed and power output, is largely dictated at the molecular level by the motor protein myosin. Composed of two heavy chains and four light chains, muscle myosin-II is a hexameric protein that composes the thick filaments in muscle sarcomeres. The myosin heavy chain is composed of three distinct regions: the globular subfragment-1 (S1) head at the N terminus, the flexible subfragment-2 (S2) neck or linker region, which gives flexibility to the protein, and the light meromyosin (LMM) rod at the C terminus. The myosin S1 heads (two per myosin protein) include binding sites for ATP and for the thin filament protein actin. Myosin is an enzyme that uses the energy from ATP hydrolysis to bind actin, undergo a conformational change (the power stroke) and induce filament sliding, as described in the sliding filament model of muscle contraction (Huxley, 1969). Variations in the arrangement of myosin proteins in the sarcomere or in the myosin protein sequence can lead to changes in myosin activity, thus affecting interfilament sliding velocity and muscle performance.

The contractile speed of striated muscles is correlated at the molecular level with the myosin isoform. Myosin isoforms with higher ATPase activities (i.e. higher ATP hydrolysis rates) are usually found in muscles that display fast shortening velocities. For example, mammalian fast-twitch muscle fibres IIa, IIb and IIx contain the fast myosin isoforms IIA, IIB and IIX, respectively (Bottinelli et al., 1991). These fast myosin isoforms possess higher ATPase activities compared with the slow myosin isoform found in slow-twitch muscle fibres (also found in cardiac muscle). The relationship between muscle shortening velocity and ATPase activity was elegantly demonstrated by Barany (Barany, 1967). Using muscles from invertebrates, lower vertebrates and higher vertebrates, Barany showed that actin- and Ca^{2+}-activated myosin ATPase activities were proportional to shortening velocities, from 0.1 to 24 muscle lengths s^{-1}. The myosin ATPase activities from these muscles were also inversely proportional to the muscle isometric twitch contraction time. Taken together, these results demonstrated that myosin ATPase activity plays a crucial role in determining the speed of muscle contraction.

The ATPase activity of myosin is dependent on the primary amino acid sequence of the myosin heavy chain (Spudich, 1994; Gauvry et al., 1997). Two regions of amino acids in the myosin S1 heads are thought to play an especially important role. Referred to as surface loops because they are thought to be flexible and did not crystallize in the S1 crystal structure shown by Rayment et al. (Rayment et al., 1993), these regions may contribute to regulating myosin ATPase activity by affecting ADP dissociation rates [surface loop 1 (Murphy and Spudich, 1998)] and/or by modulating interactions between myosin and actin [surface loop 2 (Uyeda et al., 1994)]. Perreault-Micale et al. (Perreault-Micale et al., 1996b) demonstrated that amino acid differences in surface loop 1 in scallop striated and catch muscle fibres contributed to differences in myosin ATPase activity.

The squid *Doryteuthis pealeii* possesses muscle fibres with unique contractile properties. The squid has eight arms and two tentacles, the latter of which can double their length in only 15–35 ms to capture

prey (Kier, 1982; Van Leeuwen and Kier, 1997). This rapid strike is caused by contraction of transverse muscle in the core of the tentacular stalk. The tentacle transverse muscle is unusual amongst cephalopods, as it includes fibres that are cross-striated with short thick filaments and sarcomeres (~0.8 μm thick filament length), whereas most cephalopod muscle is obliquely striated with long thick filaments [e.g. the thick filaments of arm transverse muscle fibres are ~7.4 μm (Kier and Curtin, 2002)]. Measurements of shortening velocities in transverse muscle tissue preparations revealed that the transverse muscle of the tentacle contracts nearly 10 times faster than the transverse muscle of the arm [15 *vs* 1.5 muscle lengths s^{-1}, respectively (Kier and Curtin, 2002)]. This 10-fold greater contractile speed in the tentacles is expected given the observed differences between the transverse muscle ultrastructure in the arms and tentacles. Because the tentacle fibres contain 10-fold more sarcomeres per unit length than the arm muscle and the shortening velocity is proportional to the number of elements in series (Huxley and Simmons, 1973; Josephson, 1975), the contractile speed of the transverse muscle of the tentacle is predicted to be 10-fold faster than that of the arm, which indeed it is (Kier and Curtin, 2002).

As demonstrated by Barany (Barany, 1967), myosin isoforms and ATPase activity play an important role in regulating shortening velocities in striated muscle fibres, but there has been relatively little research on the biochemistry of myosin isoforms in squid muscles. Kodama and Konno (Kodama and Konno, 1983) isolated myosin from mantle and brachial muscles of the squid *Todarodes pacificus* and found that myosin from both tissues exhibited similar Ca^{2+}, Mg^{2+} and EDTA-ATPase activity. Several groups have measured the myosin ATPase activities of myosin isolated from the mantle musculature, but no comparisons were made with other squid tissues (Konno, 1978; Tsuchiya et al., 1978; Konno et al., 1981). Lehman and Szent-Gyorgyi (Lehman and Szent-Gyorgyi, 1975) measured the ATPase activity of myosin isolated from the ventral pharynx retractor muscle, but again no comparisons were made with other squid musculature. Kier and Schachat (Kier and Schachat, 1992) found that myosin extracted from arm and tentacle muscles was identical based on comparisons of V8 protease and cyanogen bromide digested SDS-PAGE bands and low-percentage polyacrylamide and Neville gels (Neville, 1971). Matulef et al. (Matulef et al., 1998) sequenced the myosin heavy chain transcript from the funnel retractor muscle of *D. pealeii* and found two isoforms that varied in sequence in surface loop 1, suggesting that these myosin isoforms may display different ATPase activities (Matulef et al., 1998). Even if the two isoforms do possess different ATPase activities, Kier and Schachat (Kier and Schachat, 2008), using semi-quantitative RT-PCR, found that the expression of these isoforms was approximately the same in the arm and tentacle muscles, again providing evidence that squid arm and tentacle muscles possess similar myosin heavy chain proteins.

Despite the evidence that the arm and tentacles do not differ with regards to myosin heavy chain composition, a full analysis of the myosin heavy chain sequence from arm and tentacle tranverse muscle fibres has not been completed. In addition to the previously described A and B isoforms that differ in the actin binding site of surface loop 1 (Matulef et al., 1998), other differences in the sequence may exist that could influence myosin ATPase activity. To resolve this issue, we sequenced transcripts for the myosin heavy chain from cross-striated muscle fibres (tentacle) and obliquely striated muscle fibres (arm, mantle, fin and funnel retractor) of the squid *D. pealeii*. We found that the myosin heavy chain was identical in all muscles studied. Both previously identified myosin heavy chain isoforms (A and B) were found in all muscles studied, in addition to a third isoform (referred to as isoform C). These results suggest that the contractile properties of squid muscles are not regulated by myosin heavy chain isoforms (and therefore variable ATPase activities); rather, they are likely regulated by differences in muscle-fibre ultrastructure, in particular the arrangement and dimensions of the myofilaments. This study provides evidence for an alternative method of the regulation of muscle contractile properties.

MATERIALS AND METHODS

Animal preparation

Specimens of *Doryteuthis pealeii* (Lesueur 1821) were collected at night in shallow water with cast nets near the Darling Marine Center, University of Maine (Walpole, ME, USA). They were maintained in a flow-through seawater system and fed daily until used. Animals were anesthetized in cold seawater and then killed. Tissue samples from the tentacular stalk, arm, fin, mantle and funnel retractor musculature were removed, flash frozen in liquid nitrogen and stored at –80°C until use.

RNA extraction and cDNA synthesis

Total RNA was extracted from 30–50 mg of each tissue using the GeneJET RNA Purification Kit (Fermentas, Glen Burnie, MD, USA) and treated with DNase I (Fermentas). First strand cDNA synthesis was performed using 125–250 ng total RNA with the Moloney murine leukemia virus reverse transcriptase (M-MuLV RT) from the Phusion RT-PCR Kit according to the manufacturer's instructions (Finnzymes, Espoo, Finland). An oligo$(dT)_{15}$ primer was used to synthesize cDNA for conventional PCR, an anchor primer (Q_total) was used for 3′ rapid amplification of cDNA ends (RACE) cDNA synthesis and a myosin-specific primer (MHC_RT) was used for 5′ RACE cDNA synthesis (Table 1). The MHC_RT primer was based on the previously published myosin heavy chain sequence from the funnel retractor muscle of *D. pealeii* (Matulef et al., 1998).

Table 1. Primers used for 5′ and 3′ RACE to clone squid (*Doryteuthis pealeii*) myosin heavy chain sequences

Primer name	Primer sequence
Q_total	CCAGTGAGCAGAGTGACGAGGACTCGAGCTCAAGC$(T)_{17}$
Q_outer	CCAGTGAGCAGAGTGACG
Q_inner	GAGGACTCGAGCTCAAGC
MHC_RT	CGGAGCTGATGCAGTACAAG
MHC_5′RACE_1	CTCTTCATCAGTGAAGCCCAAG
MHC_5′RACE_2	TTTTACCTGCTCCAGACTCTCC
MHC_3′RACE_1	GGAACACTACCAGATGGCTG
MHC_3′RACE_2	AGACCCAGAAGAGCATGC

All primers are written 5′ to 3′.

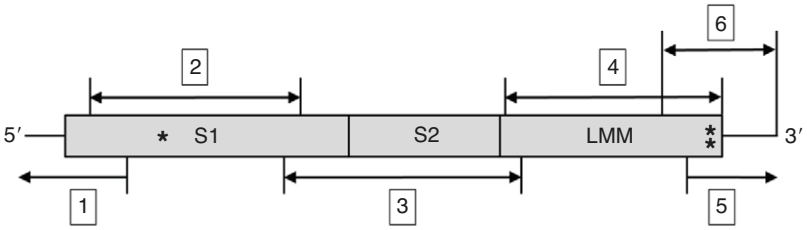

Fig. 1. Schematic of the strategy used to clone and sequence the myosin heavy chain from *Doryteuthis pealeii*. The numbers in boxes indicate different cloning regions that were amplified by PCR or RACE and sequenced. The single asterisk indicates where isoforms A and B differ, and the double asterisk indicates where isoforms A and C differ. S1 and S2, subfragments 1 and 2 of the myosin heavy chain; LMM, light meromyosin.

Cloning of squid myosin heavy chain

Because the myosin heavy chain transcript is large (~6600 bp in length), we used a cloning strategy that divided the myosin heavy chain transcript into six regions (Fig. 1). Conventional PCR was used to clone regions 2, 3 and 4 (internal coding sequences) and region 6, which spanned the internal coding region and the 3′ untranslated region (UTR). 5′ and 3′ RACE methods were used to clone regions 1 and 5. For conventional PCR, 1 µl of oligo(dT)$_{15}$-synthesized cDNA was used as a template with Phusion Hot Start II High-Fidelity DNA Polymerase (Finnzymes). PCR was carried out according to the manufacturer's instructions using primers specific for myosin heavy chain regions 2, 3, 4 and 6 (Table 2) [based on the previously published myosin heavy chain sequence for the funnel retractor muscle of *D. pealeii* (Matulef et al., 1998)]. The cycling conditions were as follows: an initial denaturation of 98°C for 30 s; 40 cycles of 98°C for 10 s, 59–67°C for 10 s (depending on the primer pair), 72°C for 2.5 min; and a final extension of 72°C for 5 min. PCR products were separated on 1% agarose gels, extracted and purified using the GeneJET Gel Extraction Kit (Fermentas). The purified PCR products were mixed in a 3:1 molar ratio with the pJET1.2/blunt vector (Fermentas) and ligated with T4 DNA ligase according to the manufacturer's instructions (CloneJET PCR Cloning Kit, Fermentas). The ligation mixture was transformed into competent JM107 *Escherichia coli* cells prepared using the TransformAid Bacterial Transformation Kit (Fermentas). Colonies that contained the proper insert were selected using colony PCR and grown overnight. Plasmid DNA was isolated using miniprep kits (Fermentas and Qiagen, Valencia, CA, USA).

5′ and 3′ RACE methods based on those described previously (Scotto-Lavino et al., 2006a; Scotto-Lavino et al., 2006b) were used to clone myosin heavy chain regions 1 and 5, respectively (Fig. 1). For 5′ RACE, the MHC_RT-synthesized cDNA was purified using the GeneJET PCR Purification Kit (Fermentas) to remove dNTPs. A synthetic polyA tail was added to the 3′ end of 15 µl of cDNA by adding 0.5 µl terminal deoxynucleotidyl transferase (TdT; Fermentas), 1 µl of 1 mmol l^{-1} dATP, 4 µl of 5× buffer, and heating at 37°C for 15 min, followed by 70°C for 10 min to deactivate the TdT. The first complementary strand to the polyA-tailed cDNA was generated using the anchor primer (Q_total, Table 1) and the Phusion polymerase (Finnzymes) by heating at 98°C for 1 min, 48°C for 2 min and 72°C for 25 min. The first PCR reaction was then carried out using the Phusion polymerase (Finnzymes), the outer anchor primer (Q_outer) and the MHC_5′RACE_1 primer (Table 1), with 30 cycles of 98°C for 10 s, 57°C for 10 s and 72°C for 90 s, and a final extension of 72°C for 10 min. A second, nested PCR was performed to increase specificity of the final PCR product. An aliquot of the first PCR reaction was used as the template with the Phusion polymerase (Finnzymes), the inner anchor primer (Q_inner) and the MHC_5′RACE_2 primer (Table 1). The following cycling conditions were used: 98°C for 1 min; 30 cycles of 98°C for 10 s, 57°C for 10 s and 72°C for 90 s; and a final extension of 72°C for 10 min. The PCR products were then purified, cloned and analyzed as described above.

The 3′ RACE methods used were based on previously published protocols (Scotto-Lavino et al., 2006a). The first complementary strand to the anchor primer (Q_total) synthesized cDNA was generated using the outer anchor primer (Q_outer, Table 1) and the Phusion polymerase (Finnzymes) by heating at 98°C for 1 min, 61°C for 2 min and 72°C for 25 min. The first PCR reaction was then carried out using the Phusion polymerase (Finnzymes), the outer anchor primer (Q_outer) and the MHC_3′RACE_1 primer (Table 1) with 30 cycles of 98°C for 10 s, 61°C for 10 s and 72°C for 90 s, and a final extension of 72°C for 10 min. A second, nested PCR was performed using an aliquot of the first PCR reaction, the Phusion polymerase (Finnzymes), the inner anchor primer (Q_inner) and the MHC_3′RACE_2 primer (Table 1). The following cycling conditions were used: 98°C for 1 min; 30 cycles of 98°C for 10 s, 61°C for 10 s and 72°C for 90 s; and a final extension of 72°C for 10 min. The PCR products were then purified, cloned and analyzed as described above.

Table 2. Primers used for conventional PCR to clone squid myosin heavy chain sequences

Primer name	Primer sequence
MHC_2_F	GGTCAAAGGAATCCACCAAA
MHC_2_R	CGGAGCTGATGCAGTACAAG
MHC_3_F	TCCAGACCATCTCCTCTGTC
MHC_3_R	CCCTGATTCCATCGATGTCT
MHC_4_F	GAACACAATGTCGGTCAACTC
MHC_4_R	TTACATCGGACTGGTACGG
MHC_6_F	AGACCCAGAAGAGCATGC
MHC_6_R	CTTTGAAAAACTAATGAAATCAGAA
MHC_A-B_F	AGCTTGGCTGGAAAGAAAGATAA
MHC_A-B_R	CAGCACCGGCAATTTTACCTT
MHC_IsoformC_F	GAAGAAAGGGCCGATCA
MHC_IsoformC_R	AAGACTTCATTTTCACTGAGTTGA

All primers are written 5′ to 3′.

Detection of myosin heavy chain isoforms

Matulef et al. (Matulef et al., 1998) identified two myosin heavy chain isoforms (A and B) that differ within a 90 bp region. Conventional PCR was used to amplify DNA surrounding this region (total 202 bp for isoform A and 187 bp for isoform B) using primers described previously (Table 2) (Kier and Schachat, 2008). In addition, a third isoform (isoform C) was detected that consisted of 36 additional base pairs at the 3′ end of the coding region as well as a shorter 3′ UTR (see Results). Conventional PCR was used to amplify DNA specific for isoform C (total 141 bp) using primers listed in Table 2. The PCR products were cloned and analyzed as described above.

DNA sequencing and analysis

Plasmid DNA was sequenced using pJET forward and reverse primers (Fermentas) as well as primers specific to internal myosin

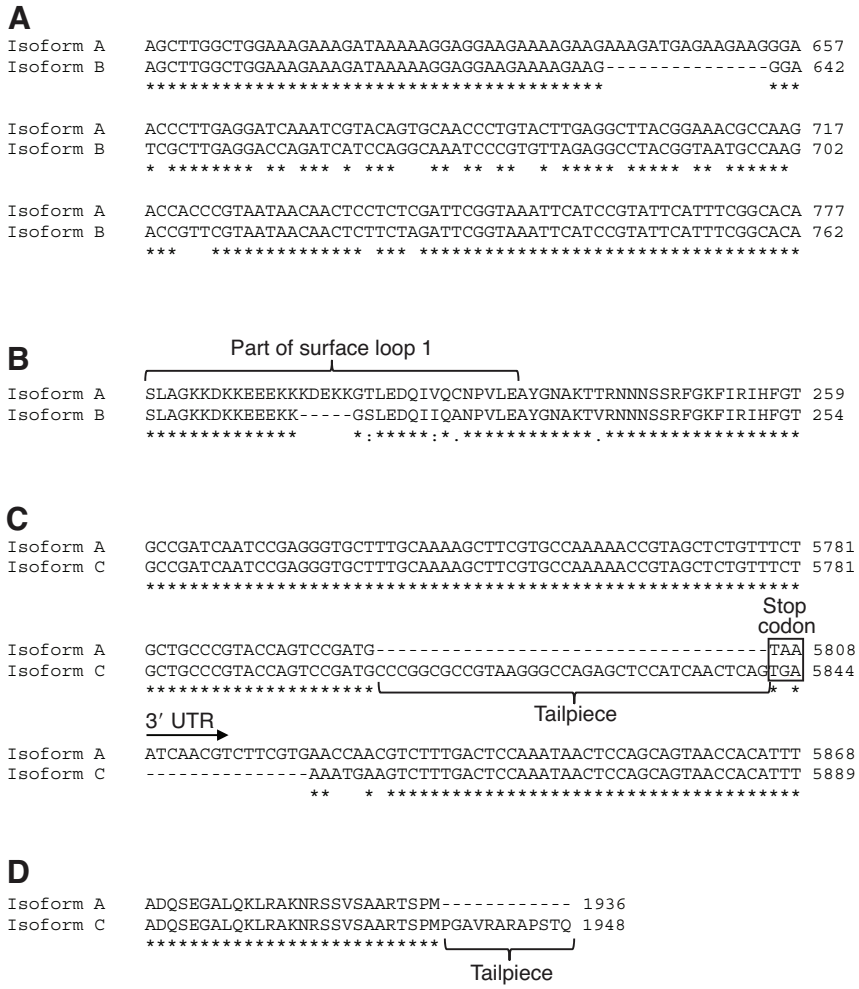

Fig. 2. Comparisons of myosin isoforms A, B and C. Nucleotide (A) and translated amino acid sequences (B) of *Doryteuthis pealeii* myosin heavy chain isoforms A and B. Isoform B contains a 15-nucleotide (five-amino-acid) deletion and four amino acid substitutions compared with isoform A. Nucleotide (C) and translated amino acid sequences (D) of *D. pealeii* myosin heavy chain isoforms A and C. Isoform C contains a 36-nucleotide (12-amino-acid) extension compared with isoform A, as well as differences in the stop codon and 3′ UTR. The remainder of the 3′ UTR is not shown because it is the same as that for isoforms A and C. Asterisks (*) denote identical amino acids (or nucleotides), colons (:) denote well-conserved amino acids, periods (.) denote somewhat conserved amino acids, and a blank space denotes no conservation. Nucleotide numbering starts from the start codon.

heavy chain sequences. DNA sequencing was carried out by Eton Biosciences (Durham, NC, USA). Sequence data were analyzed with Sequence Scanner software (v1.0, Applied Biosystems, Carlsbad, CA, USA) as well as tools on the ExPASy web server (Gasteiger et al., 2003). Sequences were compared with the previously published *D. pealeii* myosin heavy chain sequence (Matulef et al., 1998) and other myosin heavy chain sequences and aligned using the webtools LALIGN (Pearson, 2006) and ClustalW2 (Chenna et al., 2003).

RESULTS

Determination of myosin heavy chain sequence

The transcripts encoding the myosin heavy chain protein from five muscular tissues of the squid *D. pealeii* were sequenced using conventional PCR, 5′ and 3′ RACE, and standard molecular cloning methods. Strikingly, the nucleotide and translated amino acid sequences for the coding region were 100% identical amongst all five myosin heavy chain transcripts studied, suggesting that the squid *D. pealeii* expresses the same myosin heavy chain isoforms in the mantle, arm, tentacle, fin and funnel retractor muscular tissues.

The complete amino acid sequence of the *D. pealeii* myosin heavy chain protein is shown in supplementary material Fig. S1, aligned with closely related cephalopod and molluscan sequences (because the myosin heavy chain sequence was identical for all five tissues studied, only one sequence is shown and labeled *D. pealeii*). The coding region for the *D. pealeii* myosin heavy chain was found to comprise 5808 nucleotides encoding 1936 amino acids and yielding a protein with expected molecular mass of 221,891 Da. The sequence is highly similar to that previously described as myosin heavy chain isoform A for *D. pealeii* funnel retractor muscle by Matulef et al. (Matulef et al., 1998), and thus is referred to as isoform A here. The sequence obtained in the present study was one amino acid longer than that described previously, and also contained 16 amino acid substitutions. Differences in experimental techniques (e.g. the use of the proofreading Phusion polymerase in this study rather than the *Taq* polymerase used previously) and advances in sequencing technologies are likely responsible for these differences.

The 5′ UTR sequences of the myosin heavy chain transcripts were sequenced using 5′ RACE (region 1, Fig. 1). The 5′ UTR sequence was 152–157 nucleotides in length, depending on the tissue-specific transcript (supplementary material Fig. S2A). The differences in the 5′ UTR were found at the most extreme 5′ end of the transcript, likely because of experimental or technical limitations [i.e. premature termination of the reverse transcription reaction (Scotto-Lavino et al., 2006b)]. The 5′ UTRs of the different tissue myosin heavy chain transcripts were identical except for the first six nucleotides, with one exception: one sequencing reaction of the 5′ UTR from the tentacle myosin heavy chain transcript yielded a four-nucleotide insert 26 nucleotides downstream from the 5′ end of the transcript (Tentacle2, supplementary material Fig. S2A). This insert was only found once and may represent a different transcript, but further investigation is needed. The 5′ UTR

has been reported for the Japanese flying squid, *T. pacificus* (Genbank GU338005.1), and has a length of 153 bp. The 5′ UTR for *D. pealeii* was 92.2% identical to that of *T. pacificus*.

The 3′ UTR sequences of the myosin heavy chain transcripts were sequenced using 3′ RACE and conventional PCR (regions 5 and 6, Fig. 1). Initial results with 3′ RACE yielded a 3′ UTR of 408 nucleotides (not counting the polyA tail), which was identical for the myosin heavy chain transcript for all five tissues studied. This 3′ UTR sequence was 273 nucleotides shorter than that reported by Matulef et al. (Matulef et al., 1998) for the *D. pealeii* funnel retractor myosin heavy chain transcript (681 nucleotides, GenBank AF042349). In order to attempt to detect the remainder of the 3′ UTR, a reverse primer specific for the 3′ UTR detected by Matulef et al. (Matulef et al., 1998) was used for conventional PCR along with a forward primer in the coding sequence (Table 2). Using this primer set, a longer 3′ UTR was identified in all five myosin heavy chain transcripts (675 nucleotides in length, not counting the polyA tail, supplementary material Fig. S2B). The 3′ UTRs differed by only two nucleotides between all five myosin heavy chain transcripts studied.

Identification of multiple myosin isoforms

Matulef et al. (Matulef et al., 1998) determined the nucleotide and amino acid sequence (isoform A) of the myosin heavy chain from the funnel retractor muscle of *D. pealeii*, and also found a second isoform (isoform B) that differed in surface loop 1 of the S1 head region of myosin. Isoform B contained a five-amino-acid deletion and seven amino acid substitutions compared with isoform A. In this study, we wished to determine whether isoform B was found in muscular tissues of *D. pealeii* in addition to isoform A. A short region of the myosin heavy chain transcript was cloned *via* PCR, which would yield a 202 bp product for isoform A and a 187 bp product for isoform B [because of the five-amino-acid (15-nucleotide) deletion]. Plasmid DNA from multiple clones for each tissue type (mantle, arm, tentacle, fin and funnel retractor) was prepared and sequenced, and isoform B was found in all five muscle types studied; the sequence was identical for all tissues. Isoform B contained a five-amino-acid deletion as well as four amino acid substitutions compared with isoform A. Sequence comparisons for isoforms A and B identified in this study are shown in Fig. 2A,B (only one sequence is shown for each because the sequences were identical for all five muscle tissues studied). These results do not provide any information regarding the relative abundances of isoforms A and B in the tissues studied. However, they do suggest that all five tissues studied express both myosin isoforms A and B.

While cloning and sequencing the 3′ end of the myosin heavy chain transcript from RNA isolated from arm tissue, we found a third myosin heavy chain isoform (referred to as isoform C) that has not been previously described. Using primers specific for isoform C that flanked the 3′ end of the coding region and the beginning of the 3′ UTR, we amplified a short region of the myosin heavy chain transcript using PCR and analyzed multiple clones from each tissue to determine whether isoform C was expressed in the other four tissues. Myosin isoform C was detected in all five tissues studied, and all had 100% identical sequences. Isoform C contained a 36 bp extension to the 3′ end of the coding region, equivalent to a 12-amino-acid extension, or tailpiece, to the C terminus of the translated myosin heavy chain protein. Isoform C also contained a different stop codon than observed in isoform A, and a unique 7 bp region to begin the 3′ UTR, followed by 100% homology to the 3′ UTR described above (supplementary material Fig. S2B). Together, isoform C is 21 bp longer than isoform A (36 bp longer in the coding region, and 15 bp shorter in the 3′ UTR). The nucleotide and amino acid sequences for isoform C compared with isoform A are shown in Fig. 2C,D, respectively. As with the detection of isoforms A and B described above, we could not determine the relative quantities of these isoforms from these data. Table 3 summarizes the predicted properties of the three squid muscle myosin isoforms. Sequence data for all three isoforms were uploaded to the EMBL Nucleotide Sequence Database under the following accession numbers: isoform A, HE599217; isoform B, HE599218; and isoform C, HE599219.

Table 3. Predicted properties of *Doryteuthis pealeii* muscle myosin heavy chain isoforms

Myosin heavy chain isoform	Length (no. of amino acids)	Molecular mass (Da)	EMBL accession no.
A	1936	221,891	HE599217
B	1931	221,229	HE599218
C	1948	223,084	HE599219

Comparison with other molluscs

The relationship of the *D. pealeii* myosin heavy chain protein sequence to that of other molluscs was examined by aligning the sequences one pair at a time using LALIGN (Pearson, 2006). The *D. pealeii* myosin heavy chain sequence was highly similar to that of other squids, e.g. *Doryteuthis bleekeri* (unknown tissue, 96.9% identity) and *T. pacificus* (unknown tissue, 93.6% identity). It was also highly similar (94.2% identity) to that of the golden cuttlefish, *Sepia esculenta* (unknown tissue), another coleoid cephalopod. No other cephalopod sequences were available in databases, so comparisons were made with a scallop, *Aequipecten irradians* (striated adductor muscle, 73.6% identity). The complete amino acid sequence alignment for these species is shown in supplementary material Fig. S1.

DISCUSSION

The major result from this study is that identical myosin heavy chain isoforms are expressed in five different muscular tissues of the squid *D. pealeii*. This is the first time that myosin sequence information has been obtained for squid muscular tissues other than the funnel retractor muscle, and this is also the first description of a third myosin heavy chain isoform in squid muscular tissues. These results indicate that squid muscle fibres do not express different myosin isoforms with expected different ATPase activities and contractile speeds; rather, squid appear to use ultrastructural differences to achieve a range of shortening velocities (Kier and Curtin, 2002).

Transverse muscle fibres of squid tentacles shorten 10-fold faster than transverse muscle fibres of the arms (Kier and Curtin, 2002). This could be the result of two possible mechanisms. First, tentacle muscle fibres might contain a myosin isoform with a higher ATPase activity than the myosin isoform found in the slower-contracting arm muscles, as described for other muscles by Barany (Barany, 1967). Alternatively, the ultrastructural arrangement of the sarcomeres of the arm and tentacle transverse muscle fibres could differ, with shorter sarcomeres and, consequently, higher shortening velocity in the tentacle fibres. This type of ultrastructural arrangement has been described previously for squid musculature (Kier, 1985; Kier, 1991; Kier and Curtin, 2002). In the present study, our goal was to determine definitively whether the mechanism of multiple myosin isoform expression applies to differences in squid muscle shortening velocities. To do so we determined the primary amino acid sequence of the myosin heavy chain from five different muscles of the squid *D. pealeii* and found that the myosin heavy

chain sequence was 100% identical in all tissues studied (mantle, arm, tentacle, fin and funnel retractor), suggesting that the ATPase activities are identical and thus differences in ultrastructure, not differences in myosin isoforms, are likely responsible for the observed difference in contractile properties. Consistent with our results, Kodama and Konno (Kodama and Konno, 1983) measured ATPase activities from squid mantle and arm musculature and found that they were not significantly different. To our knowledge, this is the first report of an animal that uses differences in the ultrastructural arrangement of contractile proteins and sarcomeres to modulate muscle-fibre contractile properties in the absence of expression of tissue-specific myosin isoforms. This may represent a new mechanism for controlling the contractile properties of muscle.

Other molluscs employ variations in sarcomere length to modulate shortening velocity and force generation, but they also use different myosin isoforms. The adductor muscles of scallops are composed of fast fibres used for swimming and slow catch fibres used for prolonged shell closure with minimal metabolic expenditure. The fast fibres are cross-striated with relatively short sarcomeres, whereas the catch fibres are smooth and contain long myofilaments (Nunzi and Franzini Armstrong, 1981). The striated and catch fibres also shorten at dramatically different speeds; the striated fibres shorten ~20 times faster than the catch fibres (Millman, 1967). In addition to longer myofilaments and slower shortening velocities, catch muscle fibres possess a myosin isoform that has a twofold to fivefold lower ATPase activity than that of a myosin isoform from striated adductor fibres (Ruegg, 1971; Perreault-Micale et al., 1996a; Perreault-Micale et al., 1996b). The difference in ATPase activity between molluscan striated and catch fibres appears to be due to the presence of distinct striated and catch myosin isoforms within these tissues. Perreault-Micale et al. (Perreault-Micale et al., 1996a) found that the amino acid sequence of surface loop 1 of the myosin heads differed between striated and catch adductor fibres of the scallop *Placopecten magellanicus* (Perreault-Micale et al., 1996b). These isoforms are produced *via* alternative splicing of a single gene (Nyitray et al., 1994). Although molluscs such as bivalves vary myofilament and sarcomere lengths, they also express different myosin isoforms with varying myosin ATPase activities, and thus the differences in contractile properties observed are the result of both biochemical and ultrastructural differences.

Other invertebrates also employ both sarcomere length and myosin isoform variation to tune muscle contractile properties. Crustaceans have both long-sarcomere slow and short-sarcomere fast muscles (Fahrenbach, 1967). In addition to differences in sarcomere length, the ATPase activity of myosin from fast crustacean muscles is at least fourfold higher than that of myosin from slow muscles (Lehman and Szent-Gyorgyi, 1975; Mykles, 1985). This difference is likely a result of differences in myosin heavy chain amino acid sequence (Li and Mykles, 1990). Medler and Mykles (Medler and Mykles, 2003) demonstrated using RT-PCR that fast and slow muscles of the American lobster (*Homarus americanus*) express distinct fast and slow myosin heavy chain isoforms (Medler and Mykles, 2003). Distinct myosin heavy chain isoforms with differences in the surface loop 1 sequence have also been described in fast and slow muscles of gammarid amphipods (Whiteley et al., 2010), which display corresponding fast and slow ATPase activities (Ogonowski and Lang, 1979).

Sarcomeric proteins other than myosin may play a role in modulating shortening velocity. Each myosin molecule is composed of two heavy chains, two essential light chains and two regulatory light chains. Both types of light chains are required for physiological shortening velocities in vertebrate skeletal muscles (Lowey et al., 1993). Two isoforms of the essential light chain are found in vertebrate skeletal muscles, and variations in the ratio of these essential light chain isoforms have been shown to affect shortening velocities in skeletal muscle fibres (Sweeney et al., 1988). Variations in troponin isoforms may also influence contractile dynamics. Reconstitution of mouse and rat cardiac fibres with the other species' recombinant troponin complexes resulted in differences in contractile kinetics (Chandra et al., 2007). It is unknown whether differences in light chain and troponin isoforms exist in cephalopod muscle. Kier and Schachat (Kier and Schachat, 1992) showed with SDS-PAGE that the myofibrillar proteins from arm and tentacle musculature were remarkably similar, suggesting that there is little variation in troponin or light chain isoforms between these tissues. It is possible, however, that these techniques might not detect minor sequence variation that could nevertheless contribute to the observed differences in shortening velocities between arm and tentacle transverse musculatures.

External biochemical factors also have been found to affect shortening velocity in vertebrate muscles. Nitric oxide (NO) is known to decrease shortening velocities in both vertebrate skeletal and cardiac muscle fibres (Galler et al., 1997; Perkins et al., 1997). Recently, Evangelista et al. demonstrated that NO acts directly on skeletal and alpha-cardiac myosin heavy chains by S-nitrosylating cysteine residues in the myosin heavy chain, resulting in slowed actin velocity in an *in vitro* motility assay and increased myosin force production in a laser trap (Evangelista et al., 2010). NO has been well studied in cephalopods, most notably the cuttlefish *Sepia officinalis*, and has been linked to variations in manipulative behavior, regulation of blood flow and pressure, statocyst activity and the ink defense response (reviewed in Palumbo, 2005). Despite the evidence of physiological effects of NO in *Sepia* and other cephalopods, no studies have directly examined the effects of NO on muscular contractile properties in these animals. It is thus an interesting possibility that NO may differentially affect the myosin heavy chain of the arm and tentacle musculature of squid, thereby resulting in differences in shortening velocity.

We have identified three myosin isoforms (A, B and C) in squid muscles. Although it is clear that all myosin isoforms are found in all five tissues studied, a limitation of this study is that we have not measured the relative abundance of these isoforms in the different tissues. Kier and Schachat (Kier and Schachat, 2008) demonstrated using semi-quantitative RT-PCR that myosin B comprises less than 5 and 10% of the total myosin heavy chain transcripts in the arm and tentacle musculature of *D. pealeii*, respectively, suggesting that similar quantities of isoform A and B transcripts are found in each of these tissues. An interesting possibility, however, is that myosin isoforms A, B and C may be differentially expressed in the various muscle groups within each tissue type. For instance, the arms and tentacles include, in addition to the transverse muscle fibres, longitudinal, helical and oblique muscle fibres (Kier, 1982). Because of the small size of these appendages and the interdigitation of many of the fibres, it is difficult to dissect pure samples from each muscle mass. We thus used entire cross-sections, which prevented us from assessing potential differences in expression between the muscle groups. Additional studies that measure transcript levels, detect transcript expression patterns, and investigate protein levels and expression patterns using isoform-specific antibodies are warranted to determine the distribution and possible functional importance of myosin isoforms A, B and C in squid muscles.

Myosin heavy chain isoform C contains a 12-amino-acid C-terminal extension, or tailpiece. A myosin heavy chain isoform with a tailpiece has been described previously for scallop catch muscle

myosin (Perreault-Micale et al., 1996b). The myosin heavy chains of *Argopecten* and *Placopecten* sequenced from catch muscle both contained a 10-amino-acid extension at the C terminus. The scallop tailpiece is not homologous in sequence to that of the squid tailpiece described in this study. However, the existence of a myosin heavy chain isoform with an extended C terminus in squid muscles may indicate that this isoform (isoform C) is fibre specific, similar to fibre expression in the scallop. In addition, vertebrate smooth muscle myosin is expressed as two isoforms with different tailpiece lengths (SM1, 43 amino acids; SM2, nine amino acids), which affects thick filament assembly (Rovner et al., 2002). Although the functional implications of the myosin heavy chain isoform C tailpiece in squid musculature are unknown, its existence opens the interesting possibility of additional muscular specialization in cephalopods.

One of the striking differences between the ultrastructure of the squid arm and tentacle transverse muscle fibres is the length of the thick filaments: thick filaments are nearly 10 times longer in the arm than the tentacle fibres (Kier and Curtin, 2002). In addition, the tentacle transverse muscle fibres are cross-striated, whereas the arm transverse muscle fibres are obliquely striated. There are thus significant differences in thick filament organization and arrangement, and it is possible that differences in the rod portion of myosin (where thick filament assembly occurs) are necessary or even dictate these differences, as has been suggested by a previous study (Rovner et al., 2002). In the present study, however, the myosin isoforms found in the arm and tentacle were 100% identical, suggesting that differences in myosin sequence do not contribute to differences in filament assembly. It is possible that other proteins, such as paramyosin, which has been found in different levels in arm and tentacle muscle fibres (Kier and Schachat, 1992), contribute to thick filament length and organization.

Although the present study has revealed that squid muscle fibres likely express identical myosin isoforms, further investigation is needed to determine whether other squids and cephalopods also possess identical myosin isoforms. Of particular interest is the cuttlefish, a closely related coleoid cephalopod that also possesses eight arms and two tentacles. The transverse muscles of the tentacles are cross-striated, whereas the transverse muscle of the arms are obliquely striated, an arrangement that is homologous to that found in squid tentacle and arm muscles (W.M.K., unpublished).

In conclusion, we have determined the sequences of the myosin heavy chain protein for five muscular tissues of the squid *D. pealeii*. All five tissues express 100% identical myosin heavy chain proteins, suggesting that squid, and possibly cephalopods in general, do not tune muscle contractile performance through different myosin isoforms. Rather, these animals likely use changes in muscle ultrastructure to modulate whole-muscle performance.

LIST OF ABBREVIATIONS

LMM	light meromyosin
NO	nitric oxide
RACE	rapid amplification of cDNA ends
S1	myosin subfragment 1
S2	myosin subfragment 2
TdT	terminal deoxynucleotidyl transferase
UTR	untranslated region

ACKNOWLEDGEMENTS

We thank Joe Thompson and the staff of the Darling Marine Center (Walpole, ME, USA) for assistance in animal collection and handling, and Kathleen Smith for the loan of equipment. We also thank Fred Schachat for helpful discussions regarding our results.

FUNDING

This work was supported by the Seeding Postdoctoral Innovators in Research and Education (SPIRE) program funded by the National Institutes of Health, Minority Opportunities in Research division of the National Institute of General Medical Sciences [K12GM000678 to J.F.S.], and the National Science Foundation [IOS-0951067 to W.M.K.]. Deposited in PMC for release after 12 months.

REFERENCES

Barany, M. (1967). ATPase activity of myosin correlated with speed of muscle shortening. *J. Gen. Physiol.* **50**, 197-218.
Bottinelli, R., Schiaffino, S. and Reggiani, C. (1991). Force–velocity relations and myosin heavy-chain isoform compositions of skinned fibres from rat skeletal muscle. *J. Physiol.* **437**, 655-672.
Chandra, M., Tschirgi, M. L., Ford, S. J., Slinker, B. K. and Campbell, K. B. (2007). Interaction between myosin heavy chain and troponin isoforms modulate cardiac myofiber contractile dynamics. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **293**, R1595-R1607.
Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G. and Thompson, J. D. (2003). Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* **31**, 3497-3500.
Evangelista, A. M., Rao, V. S., Filo, A. R., Marozkina, N. V., Doctor, A., Jones, D. R., Gaston, B. and Guilford, W. H. (2010). Direct regulation of striated muscle myosins by nitric oxide and endogenous nitrosothiols. *PLoS ONE* **5**, e11209.
Fahrenbach, W. H. (1967). Fine structure of fast and slow crustacean muscles. *J. Cell Biol.* **35**, 69-79.
Galler, S., Hilber, K. and Gobesberger, A. (1997). Effects of nitric oxide on force-generating proteins of skeletal muscle. *Pflugers Arch.* **434**, 242-245.
Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D. and Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **31**, 3784-3788.
Gauvry, L., Mohan-Ram, V., Ettelaie, C., Ennion, S. and Goldspink, G. (1997). Molecular motors designed for different tasks and to operate at different temperatures. *J. Therm. Biol.* **22**, 367-373.
Huxley, A. F. and Simmons, R. M. (1973). Mechanical transients and origin of muscular force. *Cold Spring Harb. Symp. Quant. Biol.* **37**, 669-680.
Huxley, H. E. (1969). Mechanism of muscular contraction. *Science* **164**, 1356-1366.
Josephson, R. K. (1975). Extensive and intensive factors determining performance of striated muscle. *J. Exp. Zool.* **194**, 135-154.
Kier, W. M. (1982). The functional morphology of the musculature of squid (Loliginidae) arms and tentacles. *J. Morphol.* **172**, 179-192.
Kier, W. M. (1985). The musculature of squid arms and tentacles: ultrastructural evidence for functional differences. *J. Morphol.* **185**, 223-239.
Kier, W. M. (1991). Squid cross-striated muscle: the evolution of a specialized muscle-fiber type. *Bull. Mar. Sci.* **49**, 389-403.
Kier, W. M. and Curtin, N. A. (2002). Fast muscle in squid (*Loligo pealei*): contractile properties of a specialized muscle fibre type. *J. Exp. Biol.* **205**, 1907-1916.
Kier, W. M. and Schachat, F. H. (1992). Biochemical comparison of fast-contracting and slow-contracting squid muscle. *J. Exp. Biol.* **168**, 41-56.
Kier, W. M. and Schachat, F. H. (2008). Muscle specialization in the squid motor system. *J. Exp. Biol.* **211**, 164-169.
Kodama, S. and Konno, K. (1983). Isolation and biochemical properties of myosin from squid brachial muscle. *Bull. Jpn. Soc. Sci. Fish.* **49**, 437-442.
Konno, K. (1978). Two calcium regulation systems in squid (*Ommastrephes sloani pacificus*) muscle preparation of calcium-sensitive myosin and troponin-tropomyosin. *J. Biochem.* **84**, 1431-1440.
Konno, K., Arai, K., Yoshida, M. and Watanabe, S. (1981). Calcium regulation in squid mantle and scallop adductor muscles. *J. Biochem.* **89**, 581-589.
Lehman, W. and Szent-Gyorgyi, A. G. (1975). Regulation of muscular-contraction. Distribution of actin control and myosin control in the animal kingdom. *J. Gen. Physiol.* **66**, 1-30.
Li, Y. L. and Mykles, D. L. (1990). Analysis of myosins from lobster muscles – fast and slow isozymes differ in heavy-chain composition. *J. Exp. Zool.* **255**, 163-170.
Lowey, S., Waller, G. S. and Trybus, K. M. (1993). Skeletal muscle myosin light chains are essential for physiological speeds of shortening. *Nature* **365**, 454-456.
Matulef, K., Sirokman, K., Perreault-Micale, C. L. and Szent-Gyorgyi, A. G. (1998). Amino acid sequence of squid myosin heavy chain. *J. Muscle. Res. Cell Motil.* **19**, 705-712.
Medler, S. and Mykles, D. L. (2003). Analysis of myofibrillar proteins and transcripts in adult skeletal muscles of the American lobster *Homarus americanus*: variable expression of myosins, actin and troponins in fast, slow-twitch and slow-tonic fibres. *J. Exp. Biol.* **206**, 3557-3567.
Millman, B. M. (1967). Mechanism of contraction in molluscan muscle. *Am. Zool.* **7**, 583-591.
Murphy, C. T. and Spudich, J. A. (1998). *Dictyostelium* myosin 25-50K loop substitutions specifically affect ADP release rates. *Biochemistry* **37**, 6738-6744.
Mykles, D. L. (1985). Heterogeneity of myofibrillar proteins in lobster fast and slow muscles: variants of troponin, paramyosin, and myosin light-chains comprise four distinct protein assemblages. *J. Exp. Zool.* **234**, 23-32.
Neville, D. M. (1971). Molecular weight determination of protein-dodecyl sulfate complexes by gel electrophoresis in a discontinuous buffer system. *J. Biol. Chem.* **246**, 6328-6334.
Nunzi, M. G. and Franzini Armstrong, C. (1981). The structure of smooth and striated portions of the adductor muscle of the valves in a scallop. *J. Ultrastruct. Res.* **76**, 134-148.
Nyitray, L., Jancso, A., Ochiai, Y., Graf, L. and Szent-Gyorgyi, A. G. (1994). Scallop striated and smooth muscle myosin heavy-chain isoforms are produced by alternative RNA splicing from a single gene. *Proc. Natl. Acad. Sci. USA* **91**, 12686-12690.

Ogonowski, M. M. and Lang, F. (1979). Histochemical evidence for enzyme differences in crustacean fast and slow muscle. *J. Exp. Zool.* **207**, 143-151.
Palumbo, A. (2005). Nitric oxide in marine invertebrates: a comparative perspective. *Comp. Biochem. Physiol.* **142A**, 241-248.
Pearson, W. R. (2006). Local Protein Alignments (LALIGN). Available at http://fasta.bioch.virginia.edu/fasta_www2/fasta_www.cgi?rm=lalign&pgm=lal.
Perkins, W. J., Han, Y. S. and Sieck, G. C. (1997). Skeletal muscle force and actomyosin ATPase activity reduced by nitric oxide donor. *J. Appl. Physiol.* **83**, 1326-1332.
Perreault-Micale, C. L., Jancso, A. and Szent-Gyorgyi, A. G. (1996a). Essential and regulatory light chains of *Placopecten* striated and catch muscle myosins. *J. Muscle Res. Cell Motil.* **17**, 533-542.
Perreault-Micale, C. L., Kalabokis, V. N., Nyitray, L. and Szent-Gyorgyi, A. G. (1996b). Sequence variations in the surface loop near the nucleotide binding site modulate the ATP turnover rates of molluscan myosins. *J. Muscle Res. Cell Motil.* **17**, 543-553.
Rayment, I., Rypniewski, W. R., Schmidtbase, K., Smith, R., Tomchick, D. R., Benning, M. M., Winkelmann, D. A., Wesenberg, G. and Holden, H. M. (1993). Three-dimensional structure of myosin subfragment-1-a molecular motor. *Science* **261**, 50-58.
Rovner, A. S., Fagnant, P. M., Lowey, S. and Trybus, K. M. (2002). The carboxyl-terminal isoforms of smooth muscle myosin heavy chain determine thick filament assembly properties. *J. Cell Biol.* **156**, 113-123.
Ruegg, J. C. (1971). Smooth muscle tone. *Physiol. Rev.* **51**, 201-248.
Scotto-Lavino, E., Du, G. and Frohman, M. A. (2006a). 3′ end cDNA amplification using classic RACE. *Nat. Protoc.* **1**, 2742-2745.
Scotto-Lavino, E., Du, G. and Frohman, M. A. (2006b). 5′ end cDNA amplification using classic RACE. *Nat. Protoc.* **1**, 2555-2562.
Spudich, J. A. (1994). How molecular motors work. *Nature* **372**, 515-518.
Sweeney, H. L., Kushmerick, M. J., Mabuchi, K., Sreter, F. A. and Gergely, J. (1988). Myosin alkali light chain and heavy chain variations correlate with altered shortening velocity of isolated skeletal muscle fibers. *J. Biol. Chem.* **263**, 9034-9039.
Tsuchiya, T., Kaneko, T. and Matsumoto, J. J. (1978). Calcium sensitivity of mantle muscle of squid. *J. Biochem.* **83**, 1191-1193.
Uyeda, T. Q. P., Ruppel, K. M. and Spudich, J. A. (1994). Enzymatic activities correlate with chimeric substitutions at the actin-binding face of myosin. *Nature* **368**, 567-569.
Van Leeuwen, J. L. and Kier, W. M. (1997). Functional design of tentacles in squid: linking sarcomere ultrastructure to gross morphological dynamics. *Philos. Trans. R. Soc. Lond. B* **352**, 551-571.
Whiteley, N. M., Magnay, J. L., McCleary, S. J., Nia, S. K., El Haj, A. J. and Rock, J. (2010). Characterisation of myosin heavy chain gene variants in the fast and slow muscle fibres of gammarid amphipods. *Comp. Biochem. Physiol.* **157A**, 116-122.

Supplementary Figure 1

```
D.pealeii         MSSYDPSDPDMEFLCLTRQKLMEATSIPFDGKKNCWVPDPDFGFVGAEIQSTKGDEVTVK  60
D.pealeii-M       M-TMDFSDPDMEFLCLTRQKLMEATSIPFDGKKNCWVPDPDFGFVGAEIQSTKGDEVTVK  59
D.bleekeri        M-NMDYSDPDMEFLCLTRQKQLEVTTMPFDGKKNCWVPDPDQGFVGAEIQSTKGDEVTVK  59
T.pacificus       MSSYDPSDPDMEFLCLTRQRQLEITTVPFDGKKNCWVPDPEQGFVSAEIQSTKGDEVTVK  60
S.esculenta       M-NMDYSDPDMEFLCLTRQKQLEITTISFDGKKNCWVPDEKEGFLSAEIQSTKGDEVTVK  59
A.irradians       M-NIDFSDPDFQYLAVDRKKLMKEQTAAFDGKKNCWVPDEKEGFASAEIQSSKGDEITVK  59
                    . * ****:::*.: *:: ::  : .*********** . ** .*****:****:***

D.pealeii         TDKTQETRVVKKDDIGQRNPPKFEMNMDMANLTFLNEASILHNLRSRYESGFIYTYSGLF  120
D.pealeii-M       TDKTQETRVVKKDDIGQRNPPKFEMNMDMANLTFLNEASILHNLRSRYESGFIYTYSGLF  119
D.bleekeri        TDKTQETRVVKKDDIGQKNPPKFEMNMDMANLTFLNEASILHNLRSRYETGFIYTYSGLF  119
T.pacificus       TDKSMEMRTVKKDDVGQMNPPKFEMNMDMANLTFLNEASILHNLRSRYESGFIYTYSGLF  120
S.esculenta       VDKTQEMRTVKKDDVGQMNPPKFEMNMDMANLTFLNEASILHNLRSRYVNGFIYTYSGLF  119
A.irradians       IVADSSTRTVKKDDIQSMNPPKFEKLEDMANMTYLNEASVLYNLRSRYTSGLIYTYSGLF  119
                       . *.*****: . ******   ****:*:*****:*:****** .*:********

D.pealeii         CIAINPYRRLPIYTQGLVDKYRGKRRAEMPPHLFSIADNAYQYMLQDRENQSMLITGESG  180
D.pealeii-M       CIAINPYRRLPIYTQGLVDKYRGKRRAEMPPHLFSIADNAYQYMLQDRENQSMLITGESG  179
D.bleekeri        CVAINPYRRLPIYTQGLVDKYRGKRRAEMPPHLFSVADNAYQYMLQDRENQSMLITGESG  179
T.pacificus       CIAINPYRRLPIYTQGLVDKYRGKRRAEMPPHLFSIADNAYQYMLQDHENQSMLITGESG  180
S.esculenta       CVAINPYRRLPIYTQGLVDKYRGKRRAEMPPHLFSVADNAYQYMLQDRENQSMLITGESG  179
A.irradians       CIAVNPYRRLPIYTDSVIAKYRGKRKTEIPPHLFSVADNAYQNMVTDRENQSCLITGESG  179
                  *:*:**********:.:: ******::*:******:****** *: *:**** *******


D.pealeii         AGKTENTKKVIQYFALVAASLA--GKKDKKEEEKKKDEKKGTLEDQIVQCNPVLEAYGNA  238
D.pealeii-M       AGKTENTKKVIQYFALVAASLA--GKKDKKEEEKKKDEKKGTLEDQIVQCNPVLEAYGNA  237
D.bleekeri        AGKTENTKKVIQYFALVAASLA--GKKEKKEEEKKKDEKKGTLEDQIVQCNPVLEAYGNA  237
T.pacificus       AGKTENTKKVIQYFALVAASLTSGGKDKKKEEEKKKDEKKGTLEDQIVQCNPVLEAYGNA  240
S.esculenta       AGKTENTKKVIQYFALVAASLA--GKDKKKEEEKKKDEKKGTLEDQIVQCNPVLEAYGNA  237
A.irradians       AGKTENTKKVIMYLAKVACAVK-------KKDEEASDKKEGSLEDQIIQANPVLEAYGNA  232
                  *********** *:* **.::        *::*: .*:*:*:*****:*.**********

D.pealeii         KTTRNNNSSRFGKFIRIHFGTQGKIAGADIETYLLEKSRVTYQQSAERNYHIFYQLLSPA  298
D.pealeii-M       ETTRNNNSSRFGKFIRIHFGTQGKIAGADIETYLLEKSRVTYQQSAERNYHIFYQLLSPA  297
D.bleekeri        KTTRNNNSSRFGKFIRIHFGTQGKIAGADIETYLLEKSRVTYQQSAERNYHIFYQLLSPA  297
T.pacificus       KTTRNNNSSRFGKFIRIHFGTQGKIAGADIETYLLEKSRVTYQQSAERNYHIFYQLLSPA  300
S.esculenta       KTTRNNNSSRFGKFIRIHFGTQGKIAGADIETYLLEKSRVTFQQSAERNYHIFYQILSPA  297
A.irradians       KTTRNNNSSRFGKFIRIHFGPTGKIAGADIETYLLEKSRVTYQQSAERNYHIFYQICSNA  292
                  :*******************. ******************:*************: * *

D.pealeii         FPENIEKILAVPDPGLYGFINQGTLTVDGIDDEEEMGLTDTAFDVLGFTDEEKLSMYKCT  358
D.pealeii-M       FPENIEKILAVPDPGLYGFINQGTLTVDGIDDEEEMGLTDTAFDVLGFTDEEKLSMYKCT  357
D.bleekeri        FPDLYDKILAVPDPGLYGFINQGTLTVDGMDDEAEMILTDTAYDVLGFTDDEKLSMYKCT  357
T.pacificus       FPANIEKILAVPDPGLYGFINQGHLSVDGIDDEEEMQLTDTAFDVLGFTDDEKLSMYKCT  360
S.esculenta       FPELIEKILAVPDPGLYGFINQGALTVDGIDDEAEMKLTDTAYDVLGFSDDEKLSMYKCT  357
A.irradians       IPELNDVMLVTPDSGLYSFINQGCLTVDNIDDVEEFKLCDEAFDILGFTKEEKQSMFKCT  352
                  :*   : :*..**.***.***** *:**.:**  *: * * *:*:***:.:** **:***

D.pealeii         GCILHLGEMKWKQR--GEQAEADGTAEAEKVAFLLGVNAGDLLKCLLKPKIKVGTEYVTQ  416
D.pealeii-M       GCILHLGEMKWKQR--GEQAEADGTAEAEKVAFLLGVNAGDLLKCLLKPKIKVGTEYVTQ  415
D.bleekeri        GCILHLGEMKWKQR--GEQAEADGTAEAEKVAFLLGVNAGDLLKCLLKPKIKVGTEYVTQ  415
T.pacificus       GCILHLGEMKWKQR--GEQAEADGTAEAEKVAFLLGVNAGDLLKCLLKPKIKVGTEYVTQ  418
S.esculenta       GCILHLGEMKWKQR--GEQAEADGTAEAEKVAFLLGVNAGDLLKCLLKPKIKVGTEYVTQ  415
A.irradians       ASILHMGEMKFKQRPREEQAESDGTAEAEKVAFLCGINAGDLLKALLKPKVKVGTEMVTK  412
                  ..***:****:***   ****:************ *:*******.*****:***** **:

D.pealeii         GRNKDQVTNSIAALAKSLYDRMFNWLVRRVNQTLDTKAKRQFFIGVLDIAGFEIFDFNSF  476
```

```
D.pealeii-M      GRNKDQVTNSIAALAKSLYDRMFNWLVRRVNQTLDTKAKRQFFIGVLDIAGFEIFDFNSF  475
D.bleekeri       GRNKNQVTNSIAALAKSLYDRMFNWLVKRVNTTLDTKAKRQFFIGVLDIAGFEIFDFNSF  475
T.pacificus      GRNKDQVSNSIAALAKSLYDRMFNWLVKRVNTTLDTKAKRQFFIGVLDIAGFEIFDYNSF  478
S.esculenta      GRNKNQVTNSIAALAKSLYDRMFNWLVKRVNQTLDTKAKRQFFIGVLDIAGFEIFDFNSF  475
A.irradians      GQNMNQVVNSVGALAKSLYDRMFNWLVRRVNKTLDTKAKRNYYIGVLDIAGFEIFDFNSF  472
                 *:* :** **:.***************:*** ********:::*************:***

D.pealeii        EQLCINYTNERLQQFFNHHMFVLEQEEYKKEGIVWEFIDFGLDLQACIELIEKPMGILSI  536
D.pealeii-M      EQLCINYTNERLQQFFNHHMFVLEQEEYKKEGIVWEFIDFGLDLQACIELIEKPMGILSI  535
D.bleekeri       EQLCINYTNERLQQFFNHHMFVLEQEEYKKEGIVWEFIDFGLDLQACIELIEKPMGILSI  535
T.pacificus      EQLCINYTNERLQQFFNHHMFVLEQEEYKKEGIVWEFIDFGLDLQACIELIEKPMGILSI  538
S.esculenta      EQLCINYTNERLQQFFNHHMFVLEQEEYKKEGIVWEFIDFGLDLQACIELIEKPMGILSI  535
A.irradians      EQLCINYTNERLQQFFNHHMFILEQEEYKKEGIAWEFIDFGMDLQMCIDLIEKPMGILSI  532
                 **********************:***********.*******:*** **:***********

D.pealeii        LEEECMFPKASDTSFKNKLYDNHLGKNPMFGKP-KPPKAGCAEAHFCLHHYAGSVSYSIA  595
D.pealeii-M      LEEECMFPKASDTSFKNKLYDNHLGKNPMFGKP-KPPKAGCAEAHFCLHHYAGSVSYSIA  594
D.bleekeri       LEEECMFPKASDTSFKNKLYDNHLGKNPMFGKP-KPPKAGCSEAHFALHHYAGSVSYNIE  594
T.pacificus      LEEECMFPKASDLSFKNKLYDNHLGKNPMFGKP-KPPKAGCIEAHFALHHYAGSVSYNIS  597
S.esculenta      LEEECMFPKASDTSFKNKLYDNHLGKNPMFGKP-KPPKAGCSEAHFALHHYAGSVSYNIE  594
A.irradians      LEEECMFPKADDKSFQDKLYQNHMGKNRMFTKPGKPTRPNQGPAHFELHHYAGNVPYSIT  592
                 **********.* **::***:**:*** ** ** **.:..   *** ******.*.*.*

D.pealeii        GWLDKNKDPINENVVELLQNSKEPIVKMLFTPPEDPNP-GGKKKKGKSAAFQTISSVHKE  654
D.pealeii-M      GWLDKNKDPINENVVELLQNSKEPIVKMLFTPPRILTP-GGKKKKGKSAAFQTISSVHKE  653
D.bleekeri       GWLDKNKDPIDENVVELLQSSKEPIVKMLFTPAEDPTPAGGKKKKGKSAAFQTISSVHKE  654
T.pacificus      SWLEKNKDPINENVVELLQTSKEPIVKMLFTPPEDTSPAGGKKKKGKSAAFQTISSTHKE  657
S.esculenta      GWLDKNKDPIDENVVELLQSSKEPIVKMLFTPAEDPTPAGGKKKKGKSAAFQTISSVHKE  654
A.irradians      GWLEKNKDPINENVVALLGASKEPLVAELFKAPEEPAG-GGKKKKGKSSAFQTISAVHRE  651
                 .**:******:**** **  ****:*  **....     *********:******:.*:*

D.pealeii        SLNKLMKNLYSTHPHFVRCIIPNELKTPGLIDAALVLHQLRCNGVLEGIRICRKGFPNRI  714
D.pealeii-M      SLNKLMKNLYSTHPHFVRCIIPNELKTPGLIDAALVLHQLRCNGVLEGIRICRKGFPNRI  713
D.bleekeri       SLNKLMKNLYSTHPHFVRCIIPNELKTPGLIDAALVLHQLRCNGVLEGIRICRKGFPNRI  714
T.pacificus      SLNKLMKNLYSTHPHFVRCIIPNELKTPGMIDAALVLHQLRCNGVLEGIRICRKGFPNRI  717
S.esculenta      SLNKLMKNLYSTHPHFVRCIIPNELKTPGLIDAALVLHQLRCNGVLEGIRICRKGFPNRI  714
A.irradians      SLNKLMKNLYSTHPHFVRCIIPNELKQPGLVDAELVLHQLQCNGVLEGIRICRKGFPSRL  711
                 ************************** **::** ******:***************.*:

D.pealeii        IYSEFKQRYSILAPNAVPSGFADGKVVTDKVLSALQLDPNEYRLGNTKVFFKAGVLGMLE  774
D.pealeii-M      IYSEFKQRYSILAPNAVPSGFADGKVVTDKVLSALQLDPNEYRLGNTKVFFKAGVLGMLE  773
D.bleekeri       IYSEFKQRYSILAPNAIPGGFADGKVVTDKVLSALQLDPNEYRLGNTKVFFKAGVLGMLE  774
T.pacificus      IYSEFKQRYSILAPNAIPSGFADGKVVTDKVLSALQLDTNEYRLGNTKVFFKAGVLGMLE  777
S.esculenta      IYSEFKQRYSILAPNAIPGGFADGKVVTDKVLSALQLDPNEYRLGNTKVFFKAGVLGMLE  774
A.irradians      IYSEFKQRYSILAPNAIPQGFVDGKTVSEKILAGLQMDPAEYRLGTTKVFFKAGVLGNLE  771
                 ****************:* **.***.*::*:*:.**:*. *****.*********** **

D.pealeii        DMRDERLSKIISMFQAHIRGYLMRKAYKKLQDQRIGLTLIQRNVRKWLVLRNWEWWRLFN  834
D.pealeii-M      DMRDERLSKIISMFQAHIRGYLMRKAYKKLQDQRIGLTLIQRNVRKWLVLRNWEWWRLFN  833
D.bleekeri       DMRDERLSKIISMFQAHIRGYLMRKAYKKLQDQRVGLTLIQRNVRKWLVLRNWEWWRLFN  834
T.pacificus      DMRDERLSKIISMFQAHIRGYLMRKAYKKLQDQRIGLTLIQRNIRKWLMLRNWEWWRLFN  837
S.esculenta      DMRDERLSKIIAMFQAHIRGYLMRKAYKKLQDQRIGLTLIQRNIWKWLVLRNWEWWRLFN  834
A.irradians      EMRDERLSKIISMFQAHIRGYLIRKAYKKLQDQRIGLSVIQRNIRKWLVLRNWQWWKLYS  831
                 :**********:**********:***********:**::****:  ***:****:**:*:.

D.pealeii        KVKPLLNIARQEDENKKAQEEFAKMKEEFARCEQMRKELEEQNTVLMQQKNDLVIAMSSG  894
```

```
D.pealeii-M      KVKPLLLNIARQEDENKKAQEEFAKMKEEFASCEQMRKELEEQNTVLMQQKNDLVIAMSSG  893
D.bleekeri       KVKPLLLNIARQEDESKKIQEEFTKMKEEFSRCEQMRKELEEQNTVLMQQKNDLVIAMSSG  894
T.pacificus      KVKPLLLGMARQEEENKKAAEEFAKMKEEFLRCEQMRKELEEQNTVLMQQKNDLVISMSSS  897
S.esculenta      KVKPLLLGIARQEEENKKAQEEFAKMKEEFTRCEQMRKELEEQNTVLMQQKNDLIISMSSG  894
A.irradians      KVKPLLLSIARQEEEMKEQLKQMDKMKEDLAKTERIKKELEEQNVTLLEQKNDLFLQLQTL  891
                 ******.:****:* *:  ::: ****::   *:::*******..*::*****.: :.:

D.pealeii        EDAIGDAEEKIEQLIKQKSDFETQIKELEDKLMDEEDAATELSAQKKKSDAEIGELKKDV  954
D.pealeii-M      EDAIGDAEEKIEQLIKQKSDFETQIKELEDKLMDEEDAATELSAQKKKSDAEIGELKKDV  953
D.bleekeri       EDAIGDAEEKIEQLIKQKSDFETQIKELEDKLMDEEDAAADMSAQKKKSDAEIGELKKDV  954
T.pacificus      DDAIIESEEKIEGLIKQKSDMETHIKELEEKLMDEEDAATDLAAAKKKSEADIKELKQDV  957
S.esculenta      EDALVDAEEKIEQLIKQKSDFESQLKELEDKLMDEEDAAADLGAQKKKSDAEISELKKDV  954
A.irradians      EDSMGDQEERVEKLIMQKADFESQIKELEERLLDEEDAAADLEGIKKKMEADNANLKKDI  951
                 :*:: : **::* ** **:*:*:::****::*:******::: . *** :*:  :**:*:

D.pealeii        EDLEAGLAKAEQEKTTKDNQIKTLQDEMAQQDEHLSKLNKEKKNLEEVQKKTLEDLQAEE  1014
D.pealeii-M      EDLEAGLAKAEQEKTTKDNQIKTLQDEMAQQDEHLSKLNKEKKNLEEVQKKTLEDLQAEE  1013
D.bleekeri       EDLEAGLAKAEQEKTTKDNQIKTLQDEMAQQDEHLSKLNKEKKNLEEVQKKTLEDLQAEE  1014
T.pacificus      EDLEAGLAKAEQEKTTKDNQIKTLQDEMANQDEALSKVNKEKKALEEVQKKTLEDLQAEE  1017
S.esculenta      EDLEAGLAKAEQEKTTKDNQIKTLQDEMAQQDEHISKLNKEKKNLEEVQKKTLEDLQAEE  1014
A.irradians      GDLENTLQKAEQDKAHKDNQISTLQGEISQQDEHIGKLNKEKKALEEANKKTSDSLQAEE  1011
                  ***  * ****:*: *****.***.*:::*** :.*:***** ***.:*** :.*****

D.pealeii        DKVNHLSKLKTKLEQTLDELEDNLEREKKIRGDVDKAKRKVEQDLKTTQETVEDLERVKR  1074
D.pealeii-M      DKVNHLSKLKTKLEQTLDELEDNLEREKKIPGDVDKAKRKVEQDLKTTQETVEDLERVKR  1073
D.bleekeri       DKVNHLSKLKTKLEQTLDELEDNLEREKKIRGDVEKAKRKVEQDLKTTQETVEDLERIKR  1074
T.pacificus      DKVNHLSKLKTKLEQTLDELEDNLEREKKIRADVDKAKRKVEQDLKTTQETVEDLERVKR  1077
S.esculenta      DKVNHLSKLKTKLEQTLDELEDNLEREKKIRGDVEKAKRKLEQDLKTTQETVEDLERVKR  1074
A.irradians      DKCNHLNKLKAKLEQALDELEDNLEREKKVRGDVEKAKRKVEQDLKSTQENVEDLERVKR  1071
                 ** ***.***:****:**************: .**:*****:*****:***.******:**

D.pealeii        DLEDAGRKKDMEINGLNSKLEDEQNLVAQLQKKIKELQARIEELEEELEAERQARTKVEK  1134
D.pealeii-M      DLEDAGRKKDMEINGLNSKLEDEQNLVAQLQKKIKELQARIEELEEELEAERQARTKVEK  1133
D.bleekeri       DLEDSGRKKDMEINGLNSKLEDEQNLVAQLQKKIKELQARIEELEEELEAERQARTKVEK  1134
T.pacificus      DLEDAGRKKDMEINGLNSKLEDEQNLVAQLQKKIKELQARIEELEEELEAERQARSKVEK  1137
S.esculenta      DLEDAGRKKDMEINGLSSKLEDEQNLVSQLQKKIKELQARIEELEEELEAERQARSKVEK  1134
A.irradians      ELEENVRRKEAEISSLNSKLEDEQNLVSQLQRKIKELQARIEELEEELEAERNARAKVEK  1131
                 :**:  *:*: **..*.**********:***:********************:**:****

D.pealeii        QRTELSRELEELGERLDEAGGATAAQMELNKKREQELLRLRRDLEEATMQHESQIATLRK  1194
D.pealeii-M      QRTELSRELEELGERLDEAGGATAAQMELNKKREQELLRLRRDLEEATMQHESQIATLRK  1193
D.bleekeri       QRTELSRELEELGERLDEAGGATAAQMELNKKREQELLRLRRDLEESNMQHESQIATLRK  1194
T.pacificus      QRTELSREMEELGERLDEAGGATAAQMELNKKREQELLRLRRDLEENTMQNESTIATLRK  1197
S.esculenta      QRTELSRELEELGERLDEAGGATAAQMELNKKREQELLRLRRDLEEGTMQHESQIATLRK  1194
A.irradians      QRAELNRELEELGERLDEAGGATSAQIELNKKREAELLKIRRDLEEASLQHEAQISALRK  1191
                 **:**.**:***************:**:******* ***::****** .:*:*: *::****

D.pealeii        KNQEATNELGDQIDQLQKVKSRLEKEKTQLRAEMDDVQSQVEHAGKNRGCSEKMSKQMEA  1254
D.pealeii-M      KNQEATNELGDQIDQLQKVKSRLEKEKTQLRAEMDDVQSQVEHAGKNRGCSEKMSKQMEA  1253
D.bleekeri       KNQEATNELGDQIDQLQKVKSRLEKEKTQLRAELDDVQSQIEHAGKNRGCSEKMSKQMEA  1254
T.pacificus      KNQEATNELGDQIDQLQKVKSRLEKEKNELKEKAEEAMMELKHREKNMGCSEKMSKQMEA  1257
S.esculenta      KNQEATNELGDQIDQLQKVKSRLEKEKSQLRGELDDLQSQIEHAGKNRGCSEKMSKQMEA  1254
A.irradians      KHQDAANEMADQVDQLQKVKSKLEKDKKDLKREMDDLESQMTHNMKNKGCSEKVMKQFES  1251
                 *:*:*:**:.**:********:***:*.:*: : ::   :: *  ** *****: **:*:

D.pealeii        QLSELNAKIDDQARSVSELTSQKSRLQTEAADLTRQLEEAEHNVGQLTKLKSSLGASLED  1314
```

```
D.pealeii-M      QLSELNAKIDDQARSVSELTSQKSRLQTEAADLTRQLEEAEHNVGQLTKLKSSLGASLED  1313
D.bleekeri       QLSELNAKIDDQARSVSELTSQKSRLQTEAADLTRQLEEAEHNVGQLTKLKSSMGANLED  1314
T.pacificus      QLSELNGKIDEQARSISELGSQKSRLQVEAADLTRQLEEAEHNVGQLTKLKSSMGVNLED  1317
S.esculenta      QLSELNGKIDDQSRSISELSSQKSRLQTEAADLTRQLEEAEHNVGQLSKLKSSLNNNLED  1314
A.irradians      QMSDLNARLEDSQRSINELQSQKSRLQAENSDLTRQLEDAEHRVSVLSKEKSQLSSQLED  1311
                 *:*:**.::::. **:.** *******.* :*******:***.*. *:* **.:. .***

D.pealeii        AKRSLEDEGRLRAKLQAEVRNLNSDIDGIRESLEEEAESKSDLQRALSRANAEVQQWRSK  1374
D.pealeii-M      AKRSLEDEGRLRAKLQAEVRNLNSDIDGIRESLEEEAESKSDLQRALSRANAEVQQWRSK  1373
D.bleekeri       AKRSLEDEGRLRAKLQAEVRNLNSDIDGIRESLEEEAESKSDLQRALSRANAEVQQWRSK  1374
T.pacificus      AKRSLEDEARLRAKLQAEVRNLNSDIDGIRESLEEEQESKTDLQRALSRANAEVQQWRSK  1377
S.esculenta      AKRSLEDEARLRAKLQGEVRNLNADIDGIRETLEEEPESKSDLQRALSRANAEVQQWRSK  1374
A.irradians      ARRSLEEETRARSKLQNEVRNMHADMDAIREQLEEEQESKSDVQRQLSKANNEIQQWRSK  1371
                 *:****:* * *:*** ****:::*:*.*** **** ***:*:** **:** *:******

D.pealeii        FESEGTARADELEDAKRKLQAKLSEAEQTADTLHSKCAGLEKAKSRLQGELEDLAIDVER  1434
D.pealeii-M      FESEGAARADELEDAKRKLQAKLSEAEQTADTLHSKCAGLEKAKSRLQGELEDLAIDVER  1433
D.bleekeri       FESEGAARADELEDAKRKLQSKLGEAEQTVDTLHSKCAGLEKAKSRLQGELEDLAIDVER  1434
T.pacificus      FESEGAARADELYDSKRKLQAELSEAEQTADTLHSKCAALEKAKSRLQGELEDLAIDAER  1437
S.esculenta      FESEGAARADELEDSKRKLQAKLSEAEQTADTLHSKCAALEKAKSRLQGELEDLAIDAER  1434
A.irradians      FESEGANRTEELEDQKRKLLGKLSEAEQTTEAANAKCSALEKAKSRLQQELEDMSIEVDR  1431
                 *****: *::** * **** .:*.*****.:: ::**:.********* ****::*:.:*

D.pealeii        SSAHANNLEKKQRNFDKVVSEWQHKCNDLQAELENAQKEARSYSAELFRVRAQCEEVGDT  1494
D.pealeii-M      SSAHANNLEKKQRNFDKVVSEWQHKCNDLQAELENAQKEARSYSAELFRVRAQCEEVGDT  1493
D.bleekeri       SSAHANNLEKKQRNFDKVVSEWQHKCNDLQAELENAQKEARSYSAELFRVRAQCEEVGDT  1494
T.pacificus      SSAHANNLEKKQRNFDKVVSEWQHKCNDLQAELENAQKEARSYSAELFRVRAQCEEVGDT  1497
S.esculenta      STAHANNLEKKQRNFDKVVSEWQHKCNDLQAELENAQKEARSYSAELFRVRAQCEEVGDT  1494
A.irradians      ANASVNQMEKKQRAFDKTTAEWQAKVNSLQSELENSQKESRGYSAELYRIKASIEEYQDS  1491
                 :.* .*::***** ***..:*** * *.**:****:***:*.*****:*::*. **  *:

D.pealeii        VESLRRENKNLADEIHDLTDQLGEGGRNTHELEKARKRLALEKEELQAALEEAEAALEQE  1554
D.pealeii-M      VASLRRENKNLADEIHDLTDQLGEGGRNTHELEKARKHLALEKEELQAALEEAEGALEQE  1553
D.bleekeri       VESLRRENKNLADEIHDLTDQLGEGGRNTHELEKARKRLALEKEELQAALEEAEAALEQE  1554
T.pacificus      VESLRRENKNLADEIHDLTDQLGEGGRSTHELEKARKRLALEKEELQAALEEAEAALEQE  1557
S.esculenta      VESLRRENKNLADEIHDLTDQLGEGGRSTHELEKNRKRLLLEKEELQAALEEAEAALEQE  1554
A.irradians      IGALRRENKNLADEIHDLTDQLSEGGRSTHELDKARRRLEMEKEELQAALEEAEGALEQE  1551
                 : :*******************.****.****:* *::* :*************.*****

D.pealeii        EAKVMRATLEISQIRQEIDRRLQEKEEEFDNTRRNHQRAIESMQASLEAEAKGKAEALRI  1614
D.pealeii-M      EAKVMRATLEISQIRQEIDRRLQEKEEEFDNTRRNHQRAIESMQASLEAEAKGKAEALRI  1613
D.bleekeri       EAKVMRATLEISQIRQEIDRRLQEKEEEFDNTRRNHQRAIESMQASLEAEAKGKAEALRI  1614
T.pacificus      EAKVMRATLEISQIRQEIDRRLQEKEEEFDNTRRNHQRAIESMQASLEAEAKGKAEALRI  1617
S.esculenta      EAKVMRSTLEISQIRQEIDRRLQEKEEEFDNTRRNHQRAIESMQASLEAEAKGKAEALRI  1614
A.irradians      EAKVMRAQLEIATVRNEIDKRIQEKEEEFDNTRRNHQRALESMQASLEAEAKGKADAMRI  1611
                 ******: ***: :*:***:*:*****************:***************:*:**

D.pealeii        KKKLEGDINELEIALDATNRGKAELEKNVKKYQGQIRELQSQVEEEQRQRDEAKEHYQMA  1674
D.pealeii-M      KKKLEGDINELEIALDATNRGKAELEKNVKKYQGQIRELQSQVEEEQAQRDEAKEHYQMA  1673
D.bleekeri       KKKLEGDINELEIALDATNRGKAELEKNVKKYQGQIRELQSQVEEEQRQRDEAKEHYQMA  1674
T.pacificus      KKKLEGDINELEIALDATNRGKAELEKNVKKYQGQIRELQSAVEEEQRQRDEAKEHFMMA  1677
S.esculenta      KKKLEGDINELEIALDATNRGKAELEKNVKKYQGQIRELQSQVEEEQRQRDEAKEHYQMA  1674
A.irradians      KKKLEQDINELEVALDASNRGKAEMEKTVKRYQQQIREMQTSIEEEQRQRDEARESYNMA  1671
                 ***** ******:****:******:**.**:** ****:*: :**** *****:* : **




D.pealeii        ERRCAAINGELEELRTILEQAERARKAAENELADASDRVNELQAQVSTVGSQKRKLEGDV  1734
```

```
D.pealeii-M      ERRCAAINGELEELRTILEQAERARKAAENELADASDRVNELQAQVSTVGSQKRKLEGDV  1733
D.bleekeri       ERRCAAINGELEELRTILEQAERARKAAENELADASDRVNELQAQISTVGSQKRKLEGDV  1734
T.pacificus      ERRCAAINGELEELRTILEQAERARKAAENELADASDRVNELTAQVSTSNSQKRKLEGDV  1737
S.esculenta      ERRCAAINGELEELRTILEQAERARKAAENELADASDRVNELQAQVSTANSQKRKLERDI  1734
A.irradians      ERRCTLMSGEVEELRAALEQAERARKASDNELADANDRVNELTSQVSSVQGQKRKLEGDI  1731
                 ****: :.**:****: **********::******.****** :*:*:  .****** *:

D.pealeii        TAMQSDLDELNNELKDADERAKHAMADATRLADELRQEQDHGLSVEKMRKSLESQVKELQ  1794
D.pealeii-M      TAMQSDLDELNNELKDADERAKHAMADATRLADELRQEQDHGLSVEKMRKSLESQVKELQ  1793
D.bleekeri       TAMQSDLDELNNELKDADERAKHAMADATRLADELRQEQDHGLSVEKMRKSLESQVKELQ  1794
T.pacificus      TAMQSDLDELNNELKDADDRAKHAMGDATRLADELRQEQDHGLSIEKMRKSLESQVKELQ  1797
S.esculenta      AAMQSDLDELNHELKDADDRSKIALADATRLADELRQEQDHGLSIEKMRKSLESQVKELQ  1794
A.irradians      NAMQTDLDEMHGELKGADERCKKAMADAARLADELRAEQDHSNQVEKVRKNLESQVKEFQ  1791
                  ***:****:: ***.**:*.* *:.**:******* ****. .:**:**.*******:*

D.pealeii        VRLDESEAAALKGGKKMIQKLESRVRELEAELDSEQRRHAETQKSMRKVDRRVKELSFQQ  1854
D.pealeii-M      VRLDESEAAALKGGKKMIQKLESRVRELEAELDSEQRRHAETQKSMRKVDRRVKELSFQQ  1853
D.bleekeri       VRLDESEAAALKGGKKMIQKLESRVRELEAELDSEQRRHAETQKSMRKVDRRVKELSFQQ  1854
T.pacificus      VRLDESEAAALKGGKKMIQKLESRVRELEAELDSEQRRHAETQKSMRKVDRRVKELSFQQ  1857
S.esculenta      VRLDESESAALLGGKKMIQKLESRVRELEAELDSEQRRHAETQKSMRKVDRRLKELSFQQ  1854
A.irradians      IRLDEAEASSLKGGKKMIQKLESRVHELEAELDNEQRRHAETQKNMRKADRRLKELAFQA  1851
                 :****:*:::* *************:*******.**********.***.***:***:**

D.pealeii        EEDRKNYERMQELVDKLQNKIKTYKRQVEEAEEIAAINLAKFRKVQQELEDAEERADQSE  1914
D.pealeii-M      EEDRKNYERMQELVDKLQNKIKTYKRQVEEAEEIAAINLAKFRKVQQELEDAEERADQSE  1913
D.bleekeri       EEDRKNYERMQELVDKLQNKIKTYKRQVEEAEEIAAINLAKFRKVQQELEDAEERADQSE  1914
T.pacificus      EEDRKNYERMQELVDKLQNKIKTYKRQVEEAEEIAAINLAKFRKVQQELEDAEERADQSE  1917
S.esculenta      EEDRKNYERMQELVDKLQNKIKTYKRQVEEAEEIAAINLAKFRKVQQELEDAEERADQSE  1914
A.irradians      DEDRKNQERLQELIDKLNAKIKTFKRQVEEAEEIAAINLAKYRKAQHELEEAEERADTAD  1911
                 :***** **:***:***: ****:******************:**.*:***:******  ::

D.pealeii        GALQKLRAKNRSSVSAARTSPM----- 1936
D.pealeii-M      GALQKLRAKNRSSVSAARTSPM----- 1935
D.bleekeri       GALQKLRAKNRSSVSVARTSPM----- 1936
T.pacificus      GALQKLRAKNRSSVSVARTSPM----- 1939
S.esculenta      GALQKLRAKNRSSVSAARASPM----- 1936
A.irradians      STLQKFRAKSRSSVSVQRSSVSVSASN 1938
                 .:***:***.*****. *:*
```

Supplemental Figure 2A

```
Mantle          --GATGAGTCAGAACAGCTTTTTTGACC----TCTTGTGGCTGTTTCTACACTCGGGTGT 54
Arm             -----GAGTCAGAACAGCTTTTTTGACC----TCTTGTGGCTGTTTCTACACTCGGGTGT 51
Tentacle1       -----GAGTCAGAACAGCTTTTTTGACC----TCTTGTGGCTGTTTCTACACTCGGGTGT 51
Tentacle2       --CATCAGTCAGAACAGCTTTTTTGACCACAGTCTTGTGGCTGTTTCTACACTCGGGTGT 58
Fin             -----GAGTCAGAACAGCTTTTTTGACC----TCTTGTGGCTGTTTCTACACTCGGGTGT 51
F.Retractor     GGGATCAGTCAGAACAGCTTTTTTGACC----TCTTGTGGCTGTTTCTACACTCGGGTGT 56
T.pacificus     -----AAGTCAGAACAGCTTTTTTGACC----TCTTG-GGCTGTTTCTACACTCGGGTGT 50
                      **********************    ***** **********************

Mantle          TGACCAGGTTAAGGGGCTGCTGACCTGAGTCGTCT-AGAGTCTTTGTGGTACTCTCTT-A 112
Arm             TGACCAGGTTAAGGGGCTGCTGACCTGAGTCGTCT-AGAGTCTTTGTGGTACTCTCTT-A 109
Tentacle1       TGACCAGGTTAAGGGGCTGCTGACCTGAGTCGTCT-AGAGTCTTTGTGGTACTCTCTT-A 109
Tentacle2       TGACCAGGTTAAGGGGCTGCTGACCTGAGTCGTCT-AGAGTCTTTGTGGTACTCTCTT-A 116
Fin             TGACCAGGTTAAGGGGCTGCTGACCTGAGTCGTCT-AGAGTCTTTGTGGTACTCTCTT-A 109
F.Retractor     TGACCAGGTTAAGGGGCTGCTGACCTGAGTCGTCT-AGAGTCTTTGTGGTACTCTCTT-A 114
T.pacificus     TGACCAGGTTGAGTGGCTACTGACCTGAGGCGTTTTAGAGTCTTTGTGGCACTCTCTTTG 110
                ********** ** **** ********** *** * ************* ********

Mantle          AACTTTTTTCGGAGATTTCGATCCCAATCGAAAAAGCGTAATC 155
Arm             AACTTTTTTCGGAGATTTCGATCCCAATCGAAAAAGCGTAATC 152
Tentacle1       AACTTTTTTCGGAGATTTCGATCCCAATCGAAAAAGCGTAATC 152
Tentacle2       AACTTTTTTCGGAGATTTCGATCCCAATCGAAAAAGCGTAATC 159
Fin             AACTTTTTTCGGAGATTTCGATCCCAATCGAAAAAGCGTAATC 152
F.Retractor     AACTTTTTTCGGAGATTTCGATCCCAATCGAAAAAGCGTAATC 157
T.pacificus     AACTCTTTTCGGAGATTTCGATCCCAATCGAAAAAGCTTAATC 153
                **** ******************************** *****
```

Supplemental Figure 2B

```
Mantle          ATCAACGTCTTCGTGAACCAACGTCTTTGACTCCAAATAACTCCAGCAGTAACCACATTT 60
Arm             ATCAACGTCTTCGTGAACCAACGTCTTTGACTCCAAATAACTCCAGCAGTAACCACATTT 60
Tentacle        ATCAACGTCTTCGTGAACCAACGTCTTTGACTCCAAATAACTCCAGCAGTAACCACATTT 60
Fin             ATCAACGTCTTCGTGAACCAACGTCTTTGACTCCAAATAACTCCAGCAGTAACCACATTT 60
F.Retractor     ATCAACGTCTTCGTGAACCAACGTCTTTGACTCCAAATAACTCCAGCAGTAACCACATTT 60
                ************************************************************

Mantle          CAACTCTTCGAGGCCGCAGTCCAGGATTTTCCACATCTTACAGATCTACTCATATGTCCA 120
Arm             CAACTCTTCGAGGCCGCAGTCCAGGATTTTCCACATCTTACAGATCTACTCATATGTCCA 120
Tentacle        CAACTCTTCGAGGCCGCAGTCCAGGATTTTCCACATCTTACAGATCTACTCATATGTCCA 120
Fin             CAACTCTTCGAGGCCGCAGTCCAGGATTTTCCACATCTTACAGATCTACTCATATGTCCA 120
F.Retractor     CAACTCTTCGAGGCCGCAGTCCAGGATTTTCCACATCTTACAGATCTACTCATATGTCCA 120
                ************************************************************

Mantle          ATGGAGGAAACGACGATCAATATTTGTAAAATCAGGGATCGATAATCGAACGCATACTTA 180
Arm             ATGGAGGAAACGACGATCAATATTTGTAAAATCAGGGATCGATAATCGAACGCATACTTA 180
Tentacle        ATGGAGGAAACGACGATCAATATTTGTAAAATCAGGGATCGATAATCGAACGCATACTTA 180
Fin             ATGGAGGAAACGACGATCAATATTTGTAAAATCAGGGATCGATAATCGAACGCATACTTA 180
F.Retractor     ATGGAGGAAACGACGATCAATATTTGTAAAATCAGGGATCGATAATCGAACGCATACTTA 180
                ************************************************************

Mantle          AAATCCGTGCCACCAATCCACTTTCTCTGCATTCGCCTTCTTTTCCAATGATTTTTTTT- 239
Arm             AAATCCGTGCCACCAATCCACTTTCTCTGCATTCGCCTTCTTTTCCAATGATTTTTTTTT 240
Tentacle        AAATCCGTGCCACCAATCCACTTTCTCTGCATTCGCCTTCTTTTCCAATGATTTTTTTT- 239
Fin             AAATCCGTGCCACCAATCCACTTTCTCTGCATTCGCCTTCTTTTCCAATGATTTTTTTT- 239
F.Retractor     AAATCCGTGCCACCAATCCACTTTCTCTGCATTCGCCTTCTTTTCCAATGATTTTTTTT- 239
                ************************************************************

Mantle          ACCAATCCTGATAAAAATAGAAATATTATATTATTTAATTTATAACACAATGGGAAATGG 299
Arm             ACCAATCCTGATAAAAATAGAAATATTATATTATTTAATTTATAACACAATGGGAAATGG 300
Tentacle        ACCAATCCTGATAAAAATAGAAATATTATATTATTTAATTTATAACACAATGGGAAATGG 299
Fin             ACCAATCCTGATAAAAATAGAAATATTATATTATTTAATTTATAACACAATGGGAAATGG 299
F.Retractor     ACCAATCCTGATAAAAATAGAAATATTATATTATTTAATTTATAACACAATGGGAAATGG 299
                ************************************************************

Mantle          GACTGACAAAGAAAAAAAATTGAATATCTTTGCAAAAAAGAAGGTAAATGCTGGATATTT 359
Arm             GACTGACAAAGAAAAAAAATTGAATATCTTTGCAAAAAAGAAGGTAAATGCTGGATATTT 360
Tentacle        GACTGACAAAGAAAAAAAATTGAATATCTTTGCAAAAAAGAAGGTAAATGCTGGATATTT 359
Fin             GACTGACAAAGAAAAAAAATTGAATATCTTTGCAAAAAAGAAGGTAAATGCTGGATATTT 359
F.Retractor     GACTGACAAAGAAAAAAAATTGAATATCTTTGCAAAAAAGAAGGTAAATGCTGGATATTT 359
                ************************************************************

Mantle          TTATACGACCAAAAAGGTGTCAAAGACAATTCGATTCGATCAGGAAAGCAAAAAAAAGAA 419
Arm             TTATACGACCAAAAAGGTGTCAAAGACAATTCGATTCGATCAGGAAAGCAAAAAAAAGAA 420
Tentacle        TTATACGACCAAAAAGGTGTCAAAGACAATTCGATTCGATCAGGAAAGCAAAAAAAAGAA 419
Fin             TTATACGACCAAAAAGGTGTCAAAGACAATTCGATTCGATCAGGAAAGCAAAAAAAAGAA 419
F.Retractor     TTATACGACCAAAAAGGTGTCAAAGACAATTCGATTCGATCAGGAAAGCAAAAAAAAGAA 419
                ************************************************************

Mantle          AAATATTTAAAAAGAAAAGCCTTAAAATAACAACAAAACTCAATTTCTTTGTTTTTATTT 479
Arm             AAATATTTAAAAAGAAAAGCCTTAAAATAACAACAAAACTCAATTTCTTTGTTTTTATTT 480
Tentacle        AAATATTTAAAAAGAAAAGCCTTAAAATAACAACAAAACTCAATTTCTTTGTTTTTATTT 479
Fin             AAATATTTAAAAAGAAAAGCCTTAAAATAACAACAAAACTCAATTTCTTTGTTTTTATTT 479
F.Retractor     AAATATTTAAAAAGAAAAGCCTTAAAATAACAACAAAACTCAATTTCTTTGTTTTTATTT 479
                ************************************************************
```

```
Mantle          TCATAGTATAGAGCTTTGTGAAAAGATTTTTTTTAATATCAAACATCATATGGACAACGT 539
Arm             TCATAGTATAGAGCTTTGTGAAAAGATTTTTTTTAATATCAAACATCATATGGACAACGT 540
Tentacle        TCATAGTATAGAGCTTTGTGAAAAGATTTTTTTTAATATCAAACATCATATGGACAACGT 539
Fin             TCATAGTATAGAGCTTTGTGAAAAGATTTTTTTTAATATCAAACATCATATGGACAACGT 539
F.Retractor     TCATAGTATAGAGCTTTGTGAAAAGATTTTTTTTAATATCAAACATCATATGGACAACGT 539
                ************************************************************

Mantle          TCACAAAGTTATCAAATCTTTTTTAAAAATATATATATATTTATACACATGGAATGAATA 599
Arm             TCACAAAGTTATCAAATCTTTTTTAAAAATATATATATATTTATACACATGGAATGAATA 600
Tentacle        TCACAAAGTTATCAAATCTTTTTAAAAAATATATATATATTTATACACATGGAATGAATA 599
Fin             TCACAAAGTTATCAAATCTTTTTAAAAAATATATATATATTTATACACATGGAATGAATA 599
F.Retractor     TCACAAAGTTATCAAATCTTTTTTAAAAATATATATATATTTATACACATGGAATGAATA 599
                ********************** ***********************************

Mantle          TACGCAGATGTGAAAAAAGAGAAAGGATATCTGCAAGATCGTTAAACAAATTTCTGATTT 659
Arm             TACGCAGATGTGAAAAAAGAGAAAGGATATCTGCAAGATCGTTAAACAAATTTCTGATTT 660
Tentacle        TACGCAGATGTGAAAAAAGAGAAAGGATATCTGCAAGATCGTTAAACAAATTTCTGATTT 659
Fin             TACGCAGATGTGAAAAAAGAGAAAGGATATCTGCAAGATCGTTAAACAAATTTCTGATTT 659
F.Retractor     TACGCAGATGTGAAAAAAGAGAAAGGATATCTGCAAGATCGTTAAACAAATTTCTGATTT 659
                ************************************************************

Mantle          CATTAGTTTTTCAAAG 675
Arm             CATTAGTTTTTCAAAG 676
Tentacle        CATTAGTTTTTCAAAG 675
Fin             CATTAGTTTTTCAAAG 675
F.Retractor     CATTAGTTTTTCAAAG 675
                ****************
```