Gene xxx (2009) xxx–xxx

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Linking functional molecular variation with environmental gradients: Myosin gene diversity in a crustacean broadly distributed across variable thermal environments

J. Rock [a,*], J.L. Magnay [b], S. Beech [c], A.J. El Haj [b], G. Goldspink [c], D.H. Lunt [d], N.M. Whiteley [a]

[a] School of Biological Sciences, College of Natural Sciences, Bangor University, Bangor, Gwynedd, LL57 2UW, UK
[b] School of Medicine, Keele University, Stoke on Trent, North Staffordshire, ST4 7QB, UK
[c] Department of Surgery, Royal Free and University College Medical School, University College London, NW3 2PF, UK
[d] Department of Biological Sciences, University of Hull, Hull HU6 7RX, UK

ARTICLE INFO

Article history:
Received 17 September 2008
Received in revised form 9 January 2009
Accepted 7 February 2009
Available online xxxx

Received by J.G. Zhang

Keywords:
Skeletal muscle
Myosin heavy chain
Temperature
Latitude
Adaptation
Amphipoda

ABSTRACT

To investigate the molecular basis of temperature adaptation in natural populations we used the candidate gene approach, targeting the myosin heavy chain (*MyHC*) gene. The functional effects of genetic variation in *MyHC* have been well characterised, and changes in the flexibility of the surface loops 1 and 2, caused by modulations in length, amino acid composition and charge can play an important role in thermal acclimation in fish. However, the extent that *MyHC* diversity is influenced by natural thermal gradients is largely unknown. Sequence variation in *MyHC* cDNA was examined in 7 species of gammarid amphipod with broad latitudinal distributions and differing intertidal thermal habitats in the NE Atlantic and Arctic Oceans. A high degree of diversity was detected in the loop 1 nucleotide sequences, although not all are likely to be functional transcripts, and their deduced amino acid sequences indicated no differences in the length and charge of loop 1 and associated binding kinetics. Four isoforms for loop 2 were detected which differed in sequence length and charge distribution, suggesting functional differences in sliding velocities and ATPase activities. While all species, and indeed most individuals, expressed multiple loop 2 isoforms, analysis of the two species with the greatest number of sequenced clones revealed that *G. duebeni*, a high-shore species with the highest thermal tolerance, expressed a greater diversity of forms than *G. oceanicus*, a low intertidal species more sensitive to temperature change. Latitude further influenced *MyHC* loop 2 diversity in *G. duebeni*, as the number of isoforms increased in the northern populations. Species-specific variations in *MyHC* diversity were observed, irrespective of phylogenetic associations revealed by analysis of the mitochondrial cytochrome oxidase 1 (*CO1*) gene. Overall, it appears that the temporal temperature variations associated with higher intertidal habitat may be a greater selective agent for *MyHC* isoform diversity in gammarid muscles than broad spatial changes with latitude.

1. Introduction

Natural thermal gradients in coastal marine environments occur with change in latitude and with vertical position in the intertidal zone. Species distribution patterns across such gradients are influenced and maintained by the ability of marine ectotherms to exert physiological and biochemical compensation for temperature change (Hochachka and Somero, 2002; Somero, 2002). Underlying these compensatory responses are alterations in the structure and function of proteins determined by changes in genotype (reviewed in Schulte, 2004). Much of what we understand about compensatory responses to temperature change has been examined in the laboratory, but studies that assess the effects of natural temperature variation on protein function are few and far between, and there are relatively few systems in which the link between physiological compensation and a particular gene has been established in species in the wild. From the limited information available, it appears that metabolic enzymes show small changes in sequence leading to changes in kinetic properties among species experiencing differences in average habitat temperature by only 3–8°C (Graves and Somero, 1982). Comparisons between conspecifics have also provided evidence of natural selection acting upon specific gene products in natural populations. From the best characterised system to date, the latitudinal cline in lactate dehydrogenase-B (*LHD-B*) in *Fundulus heteroclitus*, differences in enzyme activity between populations are due to genetic changes that are best explained as adaptation to temperature rather than the result of

Abbreviations: A, adenosine; aa, amino acid(s); ADP, adenosine diphosphate; ATP, adenosine triphosphate; bp, base pair(s); BLAST, Basic Local Alignment Search Tool; cDNA, DNA complementary to RNA; *CO1*, cytochrome *c* oxidase; C-terminal, carboxyl terminal; DNase, deoxyribonuclease; gDNA, genomic DNA; *G.d.*, *Gammarus duebeni*; *G.o.*, *Gammarus oceanicus*; *LHD-B*, lactate dehydrogenase B; MEGA, Molecular Evolutionary Genetics Analysis; mRNA, messenger RNA; *MyHC* or MyHC, myosin heavy chain; N-terminal, amino terminal; Oligo, oligodeoxyribonucleotide; PCR, polymerase chain reaction; RNase, ribonulcease; S1, myosin subfragment-1 domain.

* Corresponding author.
E-mail address: j.rock@bangor.ac.uk (J. Rock).

0378-1119/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2009.02.008

phylogenetic relatedness between groups (Powers et al., 1993; Pierce and Crawford, 1997). Moreover, evolutionary adaptation in *LHD-B* has been shown to be associated with changes in gene expression brought about by sequence differences in the regulatory regions of the gene (Schulte et al., 2000).

The biochemical and physiological properties of skeletal muscle are remarkably plastic, enabling the modulation of muscle performance to a range of cellular demands and endogenous signals (Goldspink, 1998; Johnston and Temple, 2002). In eurythermal temperate fish, such as carp, acclimation temperature determines muscle contractile properties by altering the type of myosin cross-bridge and the molecular motors that produce force (Gerlach et al., 1990; Watabe, 2002; Tao et al., 2004). The functional changes are associated with the selective expression of different fast skeletal myosin isoforms, which vary in several integral loop structures and are coded for by different genes (Hirayama and Watabe, 1997; Hirayama et al., 2000). In general, these loop structures are located on the surface of the globular head, or subfragment 1 (S1) domains, of the myosin heavy chains (*MyHC*), with loop 1 associated with the site for ATP hydrolysis and loop 2 with the actin-binding site. Unlike the majority of the S1 domain, the loop sequences are highly variable among species and fibre types, varying in length, flexibility and overall charge (Goodson et al., 1999). Sequence variations have been shown to play an important role in determining the functional characteristics of myosin, with loop 1 modulating the release of ADP and loop 2 influencing actin-activated myosin ATPase activity (Uyeda et al., 1994; Knetsch et al., 1999; Murphy and Spudich, 1999). Although temperature acclimation in carp muscle is well characterised and clearly associated with modulation in *MyHC* primary structure, these changes have been documented in the laboratory under controlled conditions. Only one study has examined temperature-dependent patterns of MyHC expression in seasonally acclimatised carp subjected to annual changes in temperature, photoperiod and oxygen levels (Tao et al., 2004). In this case, seasonal changes in the expression of MyHC isoforms were observed, but the patterns of expression did not match temperature variation alone, suggesting that the relationship between MyHC expression and thermal gradients in ectotherms living in their natural environment is complex. Our study aimed to address this issue using congenerics and conspecifics of a fauna particularly well-distributed across environmental thermal gradients.

To determine whether there is a correlation between natural thermal gradients and functional sequence variation in the surface loop regions of *MyHC*, we targeted an ecologically diverse group of crustaceans, the gammarid amphipods. Species belonging to this genus are distributed across marine coastal habitats of the North Atlantic and Arctic Ocean, with individual species extending from temperate regions in the British Isles to boreal and polar regimes in northern Norway and Svalbard (Lincoln, 1979; Klekowski and Węsławski, 1991). Species also vary in intertidal distribution; some are low- or subtidal, experiencing relatively stable sea water temperatures, while others extend from intertidal, through brackish, to freshwater habitats where thermal conditions are highly variable over short time scales. Seven species of gammarids were included in the study with overlapping distribution patterns of latitude and intertidal zonation to represent a thermal gradient of at least 14°C in summer microhabitat temperature across the sampling sites. A multi-level approach was used to examine sequence diversity within and among individuals, latitudinal populations and species. Phylogenetic relationships within the gammarid amphipods are poorly resolved, so to control for phylogenetic signal in our assessment of interspecific variation we also examined molecular diversity in the mitochondrial cytochrome oxidase (*CO1*) gene. In this way, we set out to examine the extent to which environmental variation in temperature is associated with sequence variability in a specific locus of known functional significance.

2. Materials and methods

2.1. Sample collection

Seven species of gammarid amphipod were sampled, including *Gammarus duebeni*, *G. oceanicus*, *G. setosus*, *G. finmarchicus*, *G. marinus*, *G. locusta* and *Eulimnogammarus obtusatus*. These species have differing thermal habitats and a variety of thermal tolerances (Table 1). When present, individuals from each species were collected at key sites to include temperate populations in North Wales (Anglesey; 53°N) and Scotland (Isle of Skye; 58°N), boreal populations in northern Norway (Tromsø; 70°N), and Arctic populations in Svalbard (Ny-Ålesund; 79°N). Adult specimens were collected in the summer months between June and August. Species were identified by microscopical examination of their morphology after Lincoln (1979), with confirmation of voucher specimens by M. Lowe at The Natural History Museum, London. After amphipods were decapitated, gutted and the legs were removed, the remaining body, mainly comprised of abdominal muscle, was stored in RNAlater (Ambion Europe Ltd) at −80°C.

2.2. Characterisation of MyHC cDNAs

Total RNA was extracted from individual amphipods using either TRIzol reagent (Sigma-Aldrich) or RNeasy spin columns (Qiagen). RNA was treated with RNase-free DNase (DNase 1, Ambion) to remove any contaminating genomic DNA and quantified with Ribogreen (Molecular Probes) on a multilabel counter (Wallac 1420 VICTOR^2). Poly(A+)RNA, obtained after incubation with Oligo dT (Invitrogen) was reverse-transcribed with Superscript II (Gibco) into cDNA. Amplification by PCR was performed with two sets of degenerate primers. The first set targeted a 210 bp fragment encompassing the loop 1 site (Forward GGTGCTGGTAAAACCGAAAA after Holmes et al. (2002); Reverse ACCGAAGTGGATACGGATGA). The second set of primers produced a 650 bp amplicon encompassing several actin binding regions, including the loop 2 site (Forward CARTTYTTYAAY-CAYCAYATG; Reverse CCYTCIARIACICCRTTRCA). All amplification reactions contained 1–2 µl first-strand cDNA, 1× Taq DNA Polymerase buffer containing 60 nmol $MgSO_4$, 6 nmol dNTP mix, 15 pmol of each forward and reverse primer and 0.6 U Platinum Taq DNA polymerase High Fidelity (Invitrogen), in a total reaction volume of 30 µl. The PCR conditions included 35 cycles of 94°C denaturation for 30 s, annealing for 30 s at 60 and 58°C, for loops 1 and 2, respectively, and 72°C extension for 30 s. Amplifications from aforementioned PCR reactions were electrophoresed in 1.5% agarose, and target amplicons eluted

Table 1
Summary of the latitudinal distribution patterns of seven gammarid species examined in this study, along with a brief description of their position on the shore.

Species	Latitudinal range	Marine habitat	Environmental tolerance
G. oceanicus	N. Atlantic to high Arctic, 55–80°N	Shallow subtidal/ low-intertidal	4
G. locusta	Atlantic, 40–65°N	Shallow subtidal/ low-intertidal	4
G. setosus	Arctic, 70–80°N	Mid-intertidal	3
E. obtusatus	N. Atlantic to Arctic, 48–70°N	Low-intertidal to mid-intertidal	3
G. finmarchicus	N. Atlantic to Arctic, 48–66°N	Mid-intertidal	3
G. marinus	N. E. Atlantic to Arctic, 43–70°N	Mid-intertidal to high-water neaps	2
G. duebeni	N. Atlantic, 48–70°N	Mid-intertidal to high-water neaps, freshwater influence	1

The general ability of each species to tolerate environmental change is given on a scale of 1–4, with the most tolerant ranked as 1, based on intertidal location and tolerances to intertidal stressors such as temperature (Bulnheim, 1979; Fenchel and Kolding, 1979; Kolding, 1981; Kolding and Fenchel, 1981; Kolding, 1985; Gaston and Spicer, 2001).

Please cite this article as: Rock, J., et al., Linking functional molecular variation with environmental gradients: Myosin gene diversity in a crustacean broadly distributed across variable thermal environments, Gene (2009), doi:10.1016/j.gene.2009.02.008

Table 2
Nucleotide sequence diversity for loop 1 as represented by the number of forms (A–C) and their variants (A_{1-5}, C_{1-2}) found in each species.

	Form A							Form B	Form C		
	A_1	A_2	A_3	$A_{4.0}$	$A_{4.1}$	$A_{4.2}$	A_5		C_1	C C_2	*n*
G. oceanicus	5	2	7	9	2	12	2	3		4	46
G. duebeni	3	3	4		3		4	1	14		32
G. setosus			2				1		1		4
G. finmarchicus	1		1						1		3
G. marinus			1		3	1					5
E. obtusatus	1				1		1		1		4
G. locusta	2										2
Percent of total clones	11	5	16	10	9	13	9	4	9	4	

Total number of sequenced clones for each species is given as n (see Results for the number of individual amphipods *n* is comprised of).

from the gel (Millipore spin column kit) and cloned directly into the pCR2.1 vector by a T-A cloning method (Invitrogen Corporation, USA). Inserts were harvested from plasmids using either Wizard® SV 96 (Promega) or QIAprep miniprep kits (Qiagen), and sequenced bi-directionally using M13 universal primers (MWG Biotech). Chromatograms were checked and sequences aligned using CodonCode Aligner v. 1.3.4 (CodonCode, Dedham, Massachusetts, USA). Representatives of all sequence types were deposited in GenBank (Accession nos. EU877449–EU877459). The amino acid sequence was deduced with CodonCode Aligner, and charge differentials quantified for each loop site.

Two methods were used to examine intra-individual variation in MyHC transcripts. In the first, multiple clones were sequenced from randomly picked individuals for each species, with additional sequencing of clones for the two species with the broadest latitudinal distribution: *G. oceanicus* and *G. duebeni* (up to 12 clones per amphipod for 11 and 21 individuals, respectively). In the second, identification of *MyHC* amplicons of differing size was achieved using high performance liquid chromatography using a WAVE™ Nucleic Acid Fragment Analysis System under non-denaturing conditions (Transgenomic Limited, Glasgow). This enabled the screening of PCR amplifications to identify possible *MyHC* variants expressed within individual amphipods. In each case, WAVE analysis was carried out on two separate PCR amplifications, representing two individuals of each of the seven species, at each of their latitudinal populations. To maintain appropriate size-sensitivity levels for this analysis, new primers were designed from consensus sequence for the gammarid loop 2 region to generate a smaller amplicon of 179 bp (Forward GGATGGCTCGAGAAGAACAA; Reverse TTCTCCACTGTGTCGTCTGG) using the same PCR conditions as above. Cloned *MyHC* cDNAs of known sequence were used as standard size markers, and peak retention times compared between standards and samples with WAVEMAKER™ software.

In order to identify associations between the loop 1 and 2 sequences generated above, primers were designed from gammarid sequence alignments to span the ~1.5 kb sequence between loops 1 and 2 (Forward: AAAGACTGAGAACACCAAGAAG; Reverse: GGCACAATGCAACGGATGAAG). Amplification conditions were as above but with 55°C annealing and a 90 s extension. Amplicons from *G. duebeni* from different latitudes were cloned and sequenced as described above.

2.3. Resolution of phylogenetic relationships between species

Total genomic DNA was extracted from a minimum of five individuals for each species and a fragment of *CO1* gene amplified as described in Rock et al. (2007). Phylogenetic tree reconstructions were carried out in MEGA4 (Tamura et al., 2007) using neighbor-joining methods with bootstrap tests performed to 1000 replications.

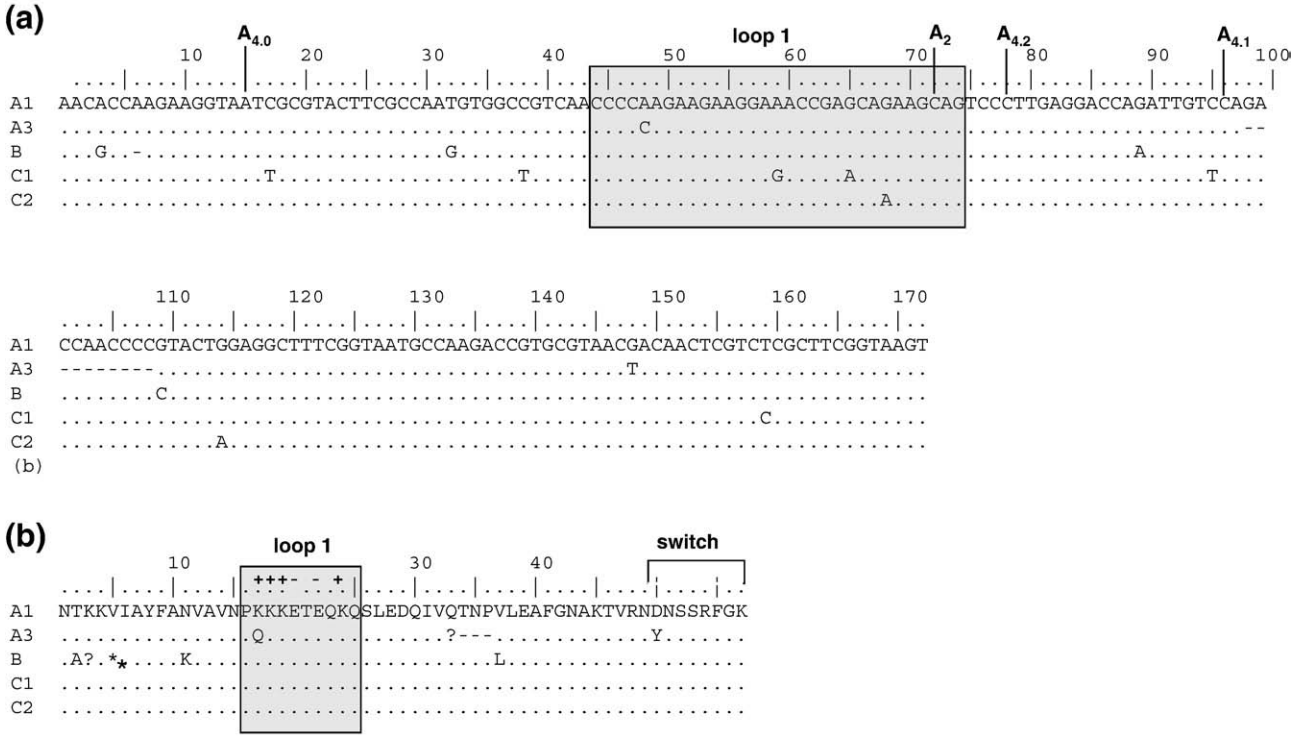

Fig. 1. Comparison of the nucleotide (a) and deduced amino acid (aa) (b) sequences of cDNA clones encoding the common MyHC forms and associated variants for the loop 1 region in the abdominal muscles of gammarid amphipods. Nucleotides and aa residues identical to the conserved form (Variant A_1) are shown by dots, with gaps represented by dashed lines. The loop sequence is boxed. The aa sequence for A_3 and B does not include the frame shifts in order to illustrate the conservation of downstream sequence. For Form B, * is used to indicate placement of the stop codon induced if the frame shift (noted with "?") was included. A_2 presents the site of point mutation eliciting a stop codon in Variant A_2; $A_{4.0}$, $A_{4.1}$ and $A_{4.2}$ indicate the single bp deletion sites characterising these variants. Residue numbers are given in (b) for cross-reference with the description in the text.

3. Results

3.1. Characterisation of MyHC loop 1 cDNAs

3.1.1. Nucleotide variation in loop 1

A total of 96 *MyHC* loop 1 cDNAs were analysed representing different individuals ($n = 38$), species, and latitudinal populations. Sequence diversity for the loop 1 region was high (Table 2). Nucleotide sequence alignment of clones revealed a set of diagnostic base identities and other defining characteristics that identified 10 distinct sequence types. These were grouped into three common forms (A, B, C) (Fig. 1a) consisting of seven variants of Form A (A_1, A_2, A_3, $A_{4.0-4.2}$, A_5), two variants of Form C (C_{1-2}), and a single sequence type for Form B (Table 2). Variant A_1 was considered the most conserved sequence, based on similarities revealed by mixed-taxa alignment with GenBank BLAST (Basic Local Alignment Search Tool; Altschul et al., 1997; highest identity score of 81% with *Anopheles gambiae*, GenBank accession no. BX060640). Variant A_1 was also the most frequently detected, along with A_3 and $A_{4.2}$. Only Variants A_1 and C_{1-2} were considered likely to be functional transcripts because the remainder (A_2–A_4 and B) were characterised by mutations either producing a frame-shift or a truncated amino acid sequence. These putatively non-functional variants were found in most gammarid species, and were considered to be authentic because of the stringent methodology used to synthesise and amplify the *MyHC* cDNAs, including use of Oligo dT during cDNA synthesis to target messenger RNA, the removal of contaminating genomic DNA (gDNA) before amplification, the use of a high fidelity Taq for PCR, and bi-directional sequencing.

Variant A_3 was the most divergent of the loop 1 sequences, being characterised by multiple non-synonymous changes and a 10 bp deletion without a subsequent stop codon in downstream sequence that was highly conserved (Fig. 1a). Variant A_2 was characterised by a point mutation that induced a stop codon at the end of loop 1, and A_4 involved a single bp deletion with no subsequent stop codon in the amplified downstream sequence. Variants classed as A_5 were singletons observed only once and characterised by 1–2 point mutations generally located outside the loop site and resulting in synonymous substitutions and/or the introduction of a stop codon (data not included on Fig. 1a). Form B diverged from Form A_1 with one synonymous and three nonsynonymous substitutions, as well as a single bp deletion upstream of the loop 1 site followed by a stop codon as shown in Fig. 1a. Due to its extremely high prevalence and characteristic frame-shift deletion, Variant A_3 was examined for evidence of alternative splicing, as observed in other invertebrate myosins (George et al., 1989; Nyitray et al., 1994). To this end, loop 1 was amplified from gDNA in *G. duebeni*. The resulting 15 sequenced clones (from four individuals at 53°N) were found to share 100% identity with cDNA Forms A and C, with no apparent coding signals for

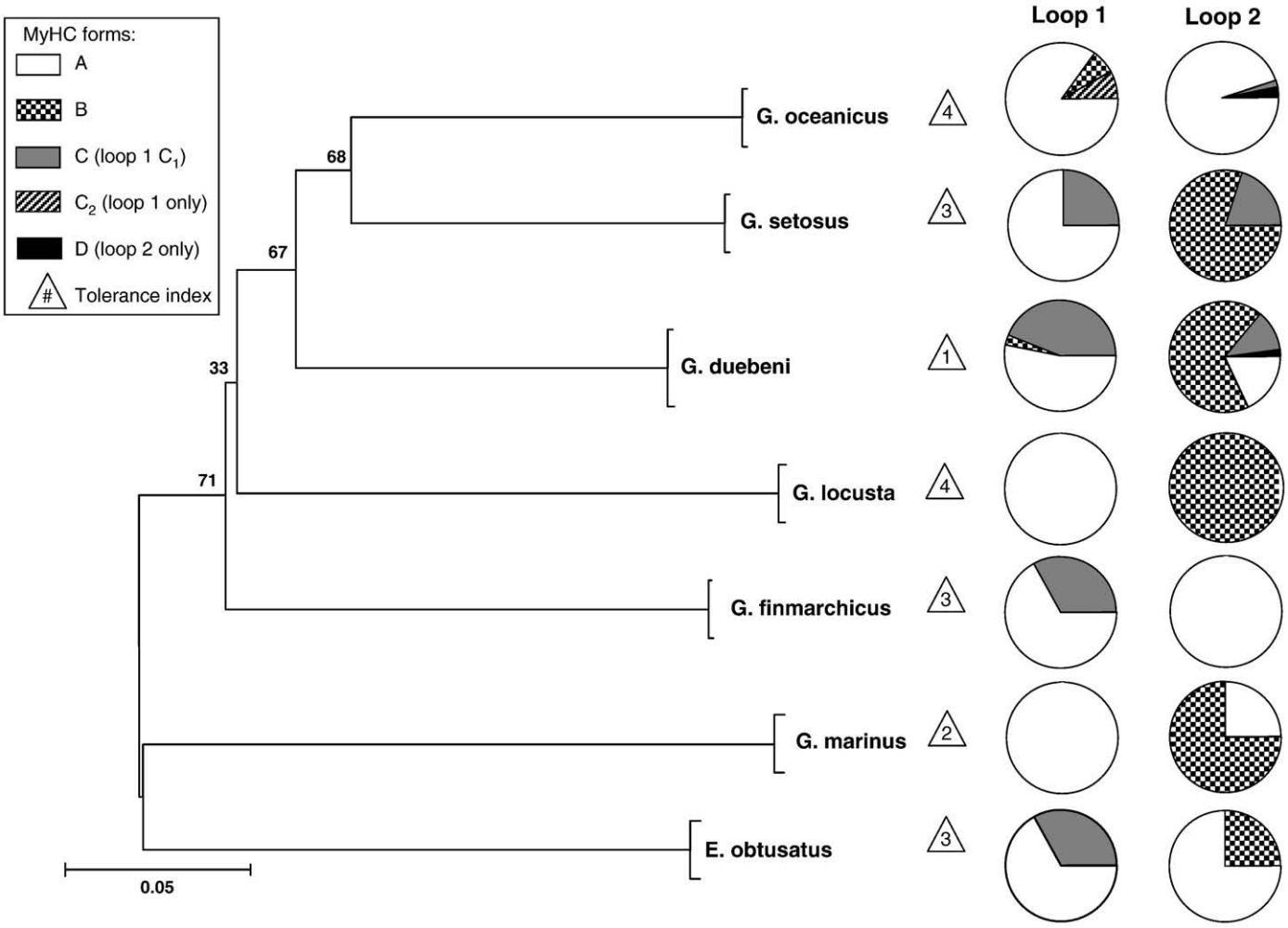

Fig. 2. Phylogenetic relationships and MyHC loop 1 and 2 molecular variation. Tree constructed by neighbor-joining analysis for seven species, with bootstrap analysis consisting of 1000 pseudo-replicates. Pie charts represent percent total composition for each MyHC form for each species (see Table 2 for sample sizes). A measure of environmental tolerance (see Table 1) is also given to allow comparison between thermal ecotypes.

Please cite this article as: Rock, J., et al., Linking functional molecular variation with environmental gradients: Myosin gene diversity in a crustacean broadly distributed across variable thermal environments, Gene (2009), doi:10.1016/j.gene.2009.02.008

splice site junctions. A similar result was obtained when 10 additional gDNA amplicons were electrophoresed on either 3.5% Nuseive or 8% polyacrylamide gels. All gDNA products resolved to the same molecular weights as those expected for the cDNA amplicons. When the deduced amino acid (aa) sequence incorporating the frame shift mutation (YWRLSVMPRPCVTTTRLASVSSSVSTS) was subjected to a BLAST search, similarity was indicated with several hypothetical or unconventional proteins, and residues 15–19 shared 73% identity with a putative unconventional *Dictyostelium MyHC* (GenBank accession no. AF090533).

3.1.2. Amino acid variation in loop 1

Synonymous changes within surface loop 1 for all major forms and variants (A_1, A_3, B, C_1 and C_2) meant that the deduced aa sequence composition, length, net charge and charge distribution were identical, apart from Variant A_3 (Fig. 1b). In this variant, the mutation from lysine to glutamine inside loop 1 resulted in a reduction in net charge from +2 in the conserved sequence to +1 in A_3. In addition, a negatively charged aspartine residue was replaced by tyrosine inside the switch site of A_3, downstream from the loop.

3.1.3. Intra-individual variation in loop 1

Examination of multiple clones from individual amphipods revealed that 79% (15/19 individuals) possessed more than one loop 1 form and/or variant: all individuals had between 1 and 3 variants of Form A, whereas 20% of the individuals co-expressed Form B and 47% co-expressed Form C (one individual possessed all three forms). Screening of PCR reactions by WAVE analysis confirmed the routine co-expression of multiple forms/variants in individual amphipods. This method was optimised to discriminate between the 210 bp amplicons associated with the majority of loop 1 forms and the 200 bp amplicon representing A_3. Screening 29 separate amplicons revealed that individuals generally co-expressed both size classes, with the 210 bp amplicon being the most dominant peak (data not shown).

3.1.4. Inter-specific variation in loop 1

Although greatest in *G. oceanicus* and *G. duebeni*, the overall diversity of forms expressed was notable for most species (Table 2; Fig. 2). Both *G. oceanicus* (46 clones from 11 individuals) and *G. duebeni* (32 clones from 20 individuals) had multiple Form A variants as well as Form B, and either C_1 or C_2 (Table 2). Variants A_1 and A_3 were detected in nearly every species, while Form B only occurred in species where multiple clones were sequenced, indicating that it was relatively rare. Interestingly, C_2 was exclusively found in *G. oceanicus*, whereas C_1 was expressed in *G. duebeni* and most other species, including *G. setosus*, which is phylogenetically closest to *G. oceanicus* (Fig. 2). The proportion of different loop 1 forms/variants also differed between species. For example, the most likely functional variants, A_1 and C_1/C_2, made up 11 and 9% of the total number of clones in *G. oceanicus*, versus 9 and 44% of the clones in *G. duebeni*, respectively (Table 2). Overall patterns of species-specific form composition for loop 1, however, did not appear to be influenced either by phylogenetic associations, or by differences in environmental tolerances (Fig. 2). This remained true for analyses restricted to only putatively functional Variants A_1 and C_1/C_2 (see Table 2 and Fig. 2).

Table 3
Nucleotide sequence diversity for loop 2 as represented by the number of forms (A–D) found in each species.

	Form A	Form B	Form C	Form D	n
G. oceanicus	59		1	2	62
G. duebeni	6	23	4	1	34
G. setosus		4	1		5
G. finmarchicus	4				4
G. marinus	1	3			4
E. obtusatus	3	1			4
G. locusta		2			2
Percent of total clones	64	28	5	3	

See Table 2 for description of *n*.

3.1.5. Latitudinal variation in loop 1

The effect of latitude on loop 1 diversity was analysed for the two species in which data was collected at three different latitudes: *G. oceanicus* and *G. duebeni* (Fig. 3). Form A was predominant at all latitudes for both species. In *G. duebeni*, Forms A and C_1 were detected at all latitudes, but diversity of form type was greatest at mid-latitudes (58°N) due to the appearance of Form B. In contrast, for *G. oceanicus* greater diversity of forms was found at higher latitudes (although this result was biased by the concomitant increase in clone numbers). When analyses are restricted to putatively functional Variant A_1 and variants of Form C, the latter generally dominates across all latitudes for *G. duebeni*, but is equally represented at higher latitudes for *G. oceanicus*.

3.2. Characterisation of MyHC loop 2 cDNAs

3.2.1. Nucleotide variation in loop 2

In total, 115 loop 2 cDNAs were analysed representing different individuals ($n = 35$), species, and latitudinal populations

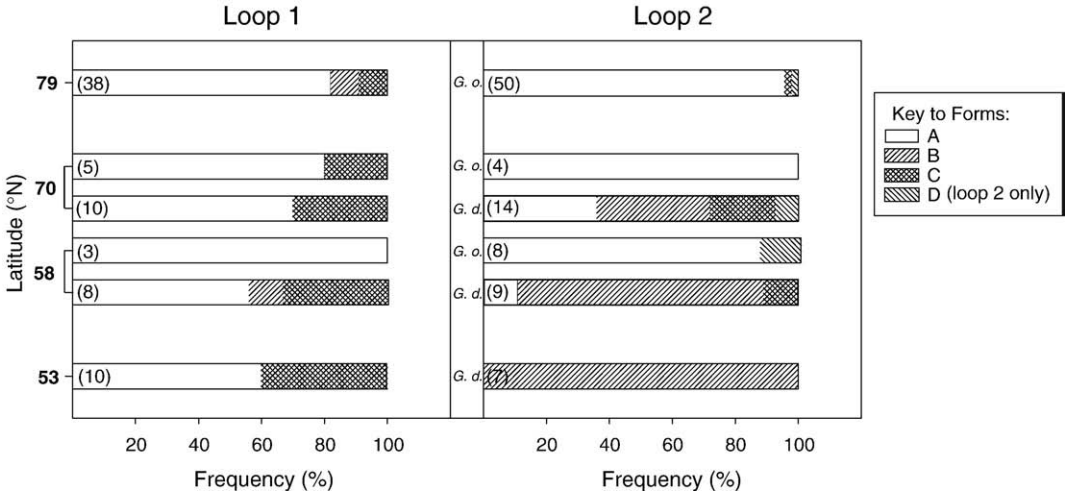

Fig. 3. Variation in frequency of *MyHC* forms for loops 1 and 2 across latitude for two species of gammarid, *Gammarus oceanicus* (*G. o.*) and *G. duebeni* (*G. d.*). Samples sizes are given at the base of each column.

Please cite this article as: Rock, J., et al., Linking functional molecular variation with environmental gradients: Myosin gene diversity in a crustacean broadly distributed across variable thermal environments, Gene (2009), doi:10.1016/j.gene.2009.02.008

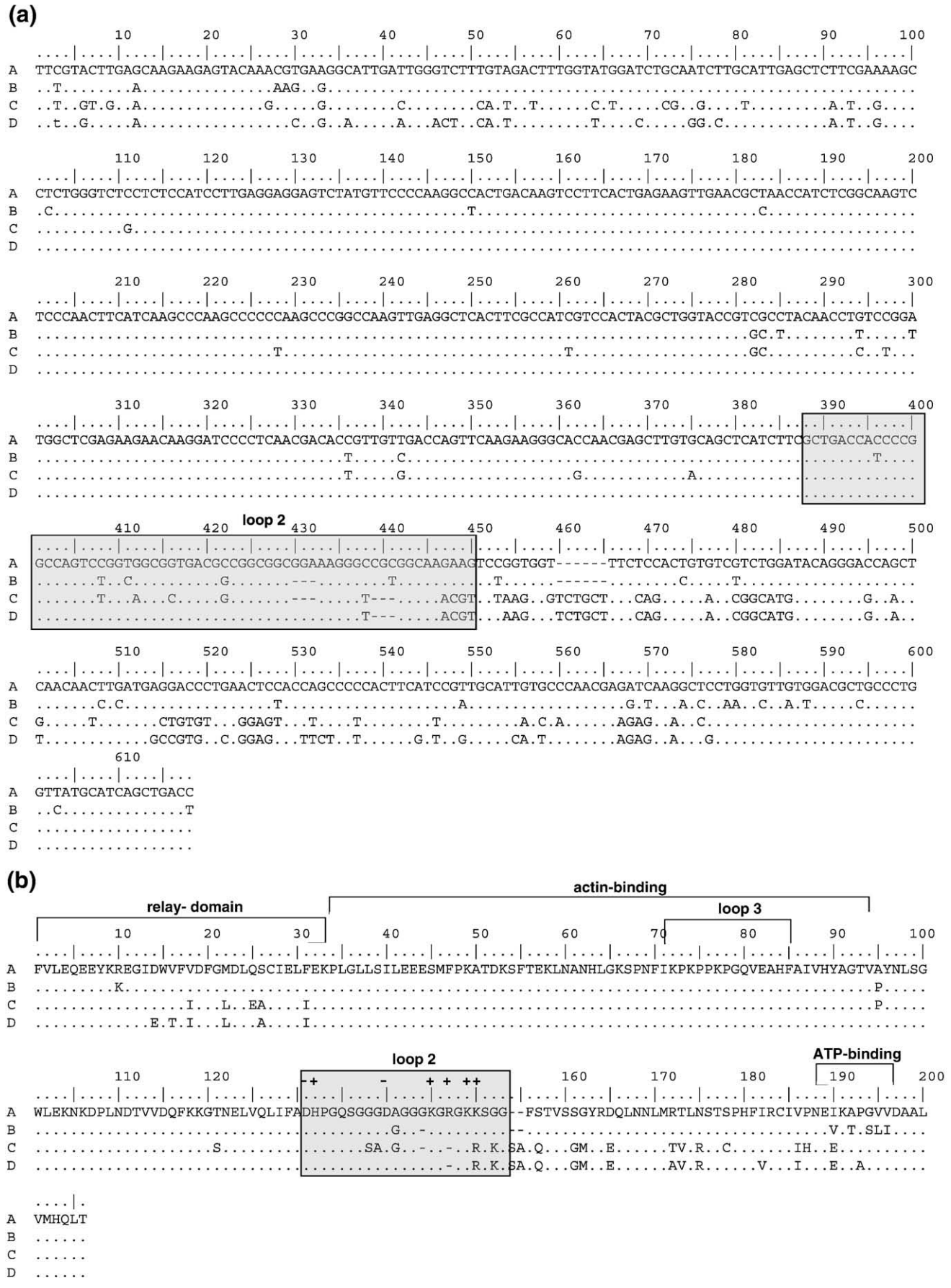

Fig. 4. Comparison of the nucleotide (a) and deduced amino acid (b) sequences of cDNA clones encoding the common MyHC forms for the loop 2 region in the abdominal muscles of gammarid amphipods.

(Table 3). Four common forms were detected (A–D; Fig. 4a). Form A was detected the most frequently (64% of all clones, Table 3), and was considered to represent the most conserved sequence based on BLAST multi-taxa alignments (highest identity 95% with *Eulimnogammarus verrucosus*, GenBank accession no. AF474964). Although nucleotide changes were numerous, most were synonymous as expected under normal purifying selection. We outline the most notable non-synonymous changes below in Section 3.2.2.

3.2.2. Amino Acid variation in loop 2 sequences

The deduced aa sequence was divergent for all four loop 2 forms, however, the net charge for the loop site was conserved at +3 (Fig. 4b). Of the eight aa substitutions separating Form B from Form A, the most significant was a mutation from alanine to proline at residue 95, upstream from the surface loop, and the deletion of glycine at residue 144 within the loop. The latter reduced the length of the loop itself but preserved the charge distribution. Form C diverged from Form A at 27 residues, including those described for Form B. Of note was a second deletion at residue 147 causing a further shortening of the loop and loss of a positively charged arginine (Fig. 4b). Although the net charge of the loop in Form C was conserved by a subsequent mutation from glycine to a positively charged lysine at residue 152, the charge distribution pattern shifted, resulting in a more positively charged 3′ end of the loop. Five additional mutations in Form C (residues 165, 172,175,187, 190) altered the charge distribution of the region immediately downstream of loop 2, some notably within an ATP-binding site (Fig. 4b). Other changes of potential structural significance included an insertion of serine and alanine immediately after the loop site (the previously described multi-taxa sequence alignment indicates that this insertion does not join the loop), and a mutation from glycine to methionine at residue 162. Form D differed from Form A at 22 residues and although it shared many of the same mutations and charge distribution as Form C, it had a longer loop length and several structural mutations including the loss of a proline to an alanine at residue 193 (Fig. 4b).

3.2.3. Intra-individual variation in loop 2

Sequencing of multiple clones from 16 individual amphipods across a range of species demonstrated that while Form A was predominant, 56% of the individuals expressed more than one loop 2 isoform. Screening by WAVE analysis against standard clones allowed discrimination of the 3–6 bp difference between loop 2 forms and confirmed that most individuals expressed two, and sometimes three size-classes.

3.2.4. Analysis of linkage between loop 1 and loop 2 forms

To identify associations between the molecular forms/variants for both loop regions, fifteen cDNA and two gDNA cloned sequences encompassing both surface loop regions (337 aa residues long) were analysed. Two forms were identified in both cDNA and gDNA: one which always linked loop 1 Form A_1 with loop 2 Form A, and another linking loop 1 Form C_1 with loop 2 Form C. In both cases, the nucleotide sequence spanning the loop regions contained no introns or deletion sites. Comparison of the corresponding deduced aa sequences between the two forms revealed 25 aa differences (Fig. 5), the majority occurring in the less conservative MyHC regions identified by multi-taxa BLAST sequence alignment. No variation was observed in the actin-binding site at residues 151–166, and even though changes occurred in the ATP-binding site (residues 71–87), they were functionally conservative. At residue 63, tyrosine was substituted for the positively charged histidine in Form A, thereby reducing the positive charge of the region in Form C which, due to its proximity to the nucleotide binding pocket, may lessen nucleotide binding efficiency (Rayment et al., 1993; Bernstein and Milligan, 1997). A shift from a non-polar to a polar residue at position 73 within the ATP-binding site in Form C may have similar functional significance (Fig. 5).

3.2.5. Inter-specific variation in loop 2

Most species expressed multiple forms of loop 2 (Table 3, Fig. 2). Diversity was highest in *G. duebeni*, where all four forms were detected, followed by three forms in *G. oceanicus*. Form A occurred in both species but was dominant in *G. oceanicus* where it accounted for 95% of the clones, while in *G. duebeni*, Form B was dominant and accounted for 68% of the clones. In both species, Forms C and D were relatively rare, and virtually absent in the remaining gammarid species (Table 3, Fig. 2). A strong species-specific difference in expression patterns was observed as Form B was not detected in *G. oceanicus*, though it was present in the closely related species, *G. setosus*. Evidence that further species-specific patterns were not directly linked to phylogeny was demonstrated by strikingly different expression patterns in *E. obtusatus* and *G. marinus*, two species which comprise a clade highly divergent from other gammarids (Fig. 2).

3.2.6. Latitudinal variation in loop 2

Loop 2 *MyHC* isoform diversity in the eurythermal species, *G. duebeni*, was influenced by latitude, because there was a trend towards greater diversity of loop 2 forms with increasing latitude (Fig. 3). Only Form B was detected in individuals collected at the southern latitude of 53°N, whereas 3 forms were detected at 58°N, although Form B was the most abundant comprising 56% of the sequenced

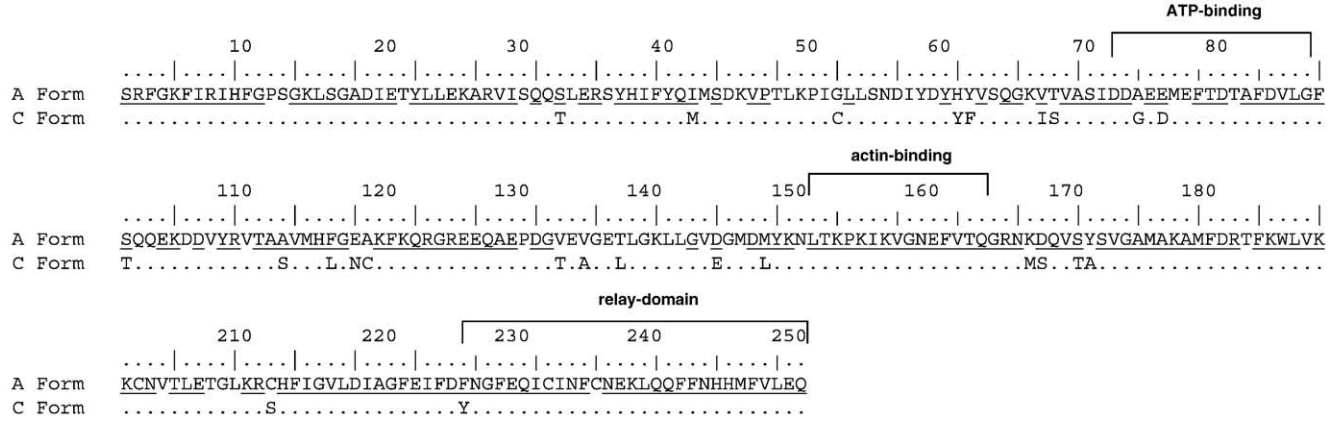

Fig. 5. Deduced amino acid sequence for the region between loops 1 and 2. Sequences were amplified from two *G. duebeni* individuals from 70°N. A five aa overlap with sequence given for loop 1 (Fig. 1) and loop 2 (Fig. 4) is preserved on either end of these sequences to clarify orientation. Areas of high (>80%) sequence conservation, ascertained from BLAST multi-taxa alignment of scallop, *Drosophila* and *C. elegans* (GenBank Accession no. X55714, XP002014739 and U59295, respectively) are underlined.

Please cite this article as: Rock, J., et al., Linking functional molecular variation with environmental gradients: Myosin gene diversity in a crustacean broadly distributed across variable thermal environments, Gene (2009), doi:10.1016/j.gene.2009.02.008

clones. At 70°N, all four loop 2 forms were present. In contrast, latitude had less effect on the diversity of loop 2 forms in *G. oceanicus*, with Form A predominating at all latitudes, and being most dominant at higher latitudes due to the decrease in frequency of Form D (Fig. 3).

3.3. Molecular variability of gammarid MyHC loop cDNAs

The analysis of sequence variation in the mitochondrial *CO1* gene among the 7 gammarid species revealed high levels of genetic divergence (generally greater than 20%) between species. Further analyses of CO1 sequence diversity appear elsewhere (Rock et al. 2007; submitted ms).

4. Discussion

4.1. Characterisation of MyHC loop 1 cDNAs

Out of the 10 distinct nucleotide sequence types detected for MyHC loop 1 in gammarid muscle, only 3 of the variants were likely to be translated into functional MyHCs (A_1, C_1 and C_2). Sequence conservation between the 3 variants, however, suggests similar kinetic roles (Goodson et al., 1999), although functional differences could be modulated by changes in associated downstream sequences. Divergent catalytic activities for rat and human MyHC β/slow isoforms, for instance, cannot be explained by loop 1 primary sequences, despite differences in ATPase activities and *in vitro* motilities (Canepari et al., 2000). Similarly in fish, MyHC actin motility is not regulated by loop 1 sequence diversity (Hirayama et al., 2000; Watabe, 2002). Collectively, these studies implicate sequence variability in other regions of the MyHC subunit. Possibilities include loop 3, the secondary actin-binding loop, and sequence variation in the convertor and relay domains which affect the overall kinetic cycle (Swank et al., 2002; Kronert et al., 2008). In addition, mutations in light meromyosin have been shown to affect the speed of shortening (Watabe et al., 2008), and variation in the net charge of loop 2 can influence the affinity of the myosin head for actin and affect sliding velocities (Furch et al., 1998; Murphy and Spudich, 2000) (see Section 3.2).

The remainder of the loop 1 nucleotide sequences was either not translated, or translated into unconventional and/or non-functional myosin proteins due to their truncated sequence or frame-shifts. As they were amplified from mRNA transcripts, they could represent processed pseudogenes, the biological function of which is still debated but may include co-option in the regulation of mRNA stability as antisense/microRNAs (Zheng and Gerstein, 2007; Hirotsune et al., 2003). The prevalence of transcripts of unknown function and their potential to encode short peptides is only now being fully recognised, even though their enormous genomic contribution is estimated to out-number that of protein coding genes (Gingeras, 2007). The role and prevalence of *MyHC* pseudogenes in gammarid muscles is far from clear, but their identification in medaka fast muscle by Liang et al. (2007) suggests that their occurrence is a more general phenomenon.

The loop 1 Variant A_3 deserves a special mention because this is the first time that a truncated MyHC loop 1 sequence has been detected in crustacean muscle. Despite the 10 bp deletion, it is possible that processing of A_3 might proceed by translational frameshifting, or ribosomal slippage resulting in a −1 or +1 frameshift functional transcript (Shigemoto et al., 2001; Meng et al., 2004). If this is the case, then the deduced aa sequence for the loop 1 region in A_3 could signify a functional change due to the reduction in loop net charge. Such a reduction matches that observed between fast and slow *MyHC* loop 1 sequences in the lobster, *Homarus gammarus* (Holmes et al., 2002), which is known to lower rates of ADP release and reduce shortening velocities in *Dictylostelium* chimeras (Sweeney et al., 1998). These observations support the involvement of loop 1 in determining the well-characterised biochemical and contractile differences between fast and slow muscle fibres in crustaceans (e.g. Ogonowski and Lang, 1979; Holmes et al., 1999). Parallel work on the biochemical and molecular characterisation of *MyHC* loop 1 transcripts in the deep flexor muscles (fast) and the superficial extensor muscles (slow) of gammarids provides further evidence that A_3 is associated with slow fibre types, while A_1 is a fast fibre isoform (N.M. Whiteley and J. Rock, unpubl. obs.).

4.2. Characterisation of MyHC loop 2 cDNAs

Unlike the MyHc loop 1 sequences, there were 4 distinct loop 2 sequences in gammarids, representing 4 different isoforms due to non-synonymous changes in the loop region. Although it is beyond the scope of this paper to pinpoint the functional significance of every individual aa substitution or deletion between the loop 2 forms, there is enough background information on the structure/function relationships in MyHCs to make some general predictions on loop flexibility and catalytic efficiency between the isoforms. For example, the net positive charge from lysine and arginine residues is responsible for influencing the interaction between the myosin cross-bridges and the negatively charged residues on the actin molecule N-terminus (Yamamoto, 1989). The lysine residues are of particular importance as they regulate association of myosin with actin (Uyeda et al., 1994) and their distribution within the loop region has been found to vary between crustacean species (Holmes et al., 2002). Even minor changes in charge separation can translate into significant biochemical differences, particularly if they occur at the C-terminal end where positive lysine residues interact with negatively charged N-terminal residues of actin (Uyeda et al., 1994; Furch et al., 1998; Murphy and Spudich, 1999; Hirayama et al., 2000). Consequently, the shift in positive charge distribution observed towards the C-terminal of the loop 2 site in Forms C and D, is expected to augment actin activated-ATPase activity, perhaps compensating for the shorter loop size in these Forms. Loop length decreased from 23 aa in Form A to 22 aa in Forms B and D, and then to 21 aa in Form C. The greater loop length in Form A is expected to increase both loop flexibility and rate of ADP release (Furch et al., 1998).

The functional repercussions of the amino acid changes in the regions flanking the loop 2 site are less well understood. The mutation at residue 95, which introduces a proline residue in Forms B and C, comes at the end of a highly conserved sequence that forms one component of the actin binding surface including another surface loop, loop 3 (Rayment et al., 1993); such a mutation is expected to affect the stability and flexibility of the actin-binding surface. Mutations observed downstream from loop 2 occur in exon 13 which is alternatively spliced in muscle sequences from the scallop to create myosins with different functions (Nyitray et al., 1994; Perreault-Micale et al., 1996). This region is a long α-helix (S^{650} to R^{665} in Chicken S1; Rayment et al., 1993) which leads to another large surface loop, capping one end of the nucleotide binding pocket. Therefore, any aa differences in exon 13 occur in the region that connects part of the actin binding site to part of the ATP binding site in loop 1. In gammarid MyHCs, significant sequence diversity occurred in this region near, or in the ATP binding site, which is likely to contribute to biochemical differences, as predicted in the scallop, where aa differences in exon 13 may affect both ATPase activity and phosphate release (Perreault-Micale et al., 1996; Murphy and Spudich, 2000). In gammarid Form D, for instance, the loss of a proline residue at residue 193 is expected to alter stability of the region, and the introduction of a negatively charged glutamic acid at residue 190 in Forms C and D imposes a change in charge. An additional site of importance is the relay domain upstream of the loop, where several significant mutations occurred among Forms A–D at residues 10–31. Distinguishing features of this region include mutations at residues 9 and 10, and 14 and 16, which contribute to the functional variation observed among *Drosophila* MyHC isoforms (Kronert et al., 1991, 2008).

Please cite this article as: Rock, J., et al., Linking functional molecular variation with environmental gradients: Myosin gene diversity in a crustacean broadly distributed across variable thermal environments, Gene (2009), doi:10.1016/j.gene.2009.02.008

4.3. Linkage between loop 1 and loop 2 forms

Analysis of the sequences encompassing both loop regions in gammarid MyHC revealed 2 forms (A and C). By association, the pairing of loop 1 and 2 forms suggests that the A and C loop 2 sequences were fast muscle isoforms. The conservation in loop net charge between the various forms suggests that all loop 2 isoforms were fast, as slow fibres have been shown to be associated with lowered loop charges in other crustaceans (Holmes et al. 2002).

4.4. Diversity of MyHC loop 1 and loop 2 cDNAs

The presence of multiple *MyHC* transcripts for both loops 1 and 2 demonstrates the importance of examining multiple clones from each individual. As relatively large numbers of clones were sequenced per individual in the current study, it was possible to assess *MyHC* diversity among individuals, as well as populations and species. Such an approach showed that many *MyHC* loop 1 transcripts were present in gammarid muscle, and that there was a diversity of transcripts in all species studied. Any functional implications of loop 1 diversity, however, cannot be explained by loop 1 sequences alone and downstream sequences are required for a fuller analysis. Diversity of MyHC loop 2 sequences are more significant due to the presence of 4 distinct isoforms, and observed to vary between species and with latitude, based on the sequence data from the two species with the highest number of clones: *G. oceanicus* and *G. duebeni*.

Interspecific sequence similarity was generally high in both loop regions, which demonstrates significant conservation across evolutionary time in *Gammarus* (see Hou et al., 2007). However, there appear to be some inter-specific differences in patterns of loop 2 *MyHC* form diversity between *G. oceanicus* and *G. duebeni*. These may be linked to differences in intertidal habitat and thermal experiences: *G. oceanicus* is a shallow subtidal species living in relatively stable thermal conditions, whereas *G. duebeni* inhabits the high intertidal and experiences a greater range and temporal variability of temperature. Therefore, at a given latitude, *G. duebeni* is more eurythermal than *G. oceanicus*, and this is reflected in their differences in environmental tolerances (Table 1). The predominance of only one *MyHC* loop 2 form in the muscles of *G. oceanicus* and the presence of 4 forms in *G. duebeni*, suggests that gammarid *MyHC* forms are temperature-sensitive. The maintenance of muscle performance in *G. duebeni*, despite rapid changes in temperature, may be possible by switching from one temperature *MyHC* form to another. In the more stenothermal species, Form A may function more generally by maintaining muscle performance over a narrower range of temperatures and during more progressive thermal changes. Forms B to D could be more specialised, maintaining muscle function at specific temperatures as found in eurythermal fish in the laboratory (Imai et al., 1997; Watabe, 2002). For instance, in the carp, *Cyprinus carpio*, temperature-sensitive fast skeletal MyHC isoforms are characterised by aa differences in the loop 2 regions resulting in higher sliding velocities and actin-activated ATPase activities in 10°C versus 30°C fish (Chaen et al., 1996; Hirayama and Watabe, 1997; Hirayama et al., 2000). An increase in muscle ATPase activities in the cold indicates compensation to improve swimming performance in cyprinids at low temperatures (Johnston and Temple, 2002); abdominal muscles in *G. duebeni* may respond in a similar fashion as swimming velocities were shown to be remarkably uncoupled with acclimation temperature (J. Rock, unpubl. data).

In addition, the change in isoform diversity with latitude could be a reflection of the changing thermal regimes experienced by populations of *G. duebeni*. Northern, boreal populations of *G. duebeni* experience different thermal ranges to their temperate counterparts, especially over an annual cycle. At 70°N, *G. duebeni* continue to show effective escape responses even when living in pools of water covered in ice during the prolonged winters, but can experience habitat temperatures upwards of 16°C for short periods at low tide on warm days during the summer. More temperate populations experience a seasonal thermal range of 7 to 20°C (J. Rock and S. Rastrick, pers. obs.). Thermal ranges increase with latitude for the intertidal species, supporting the need for various temperature-sensitive MyHC isoforms. Initial studies also provide biochemical evidence for a latitudinal shift in MyHC isoforms in the abdominal muscles of *G. duebeni* as myofibrillar ATPase activities in animals from 70 to 53°N differ in their sensitivities to temperature (L.S. Faulkner and N.M. Whiteley, unpubl. obs.). In sharp contrast, only one loop 2 form, Form A, dominated in *G. oceanicus* muscles at all latitudes. Again, muscle function as determined by loop 2 Form A appears to be possible over the relatively narrow range of water temperatures experienced by *G. oceanicus* at 70°N (i.e. 0–5°C).

These data suggest that isoform diversity in *G. duebeni* and *G. oceanicus* is indicative of adaptive variation associated with environmental temperature. The dominance of loop 2 Form A, and the lack of loop 2 Form B in *G. oceanicus*, the more cold-adapted species with a high latitudinal distribution, suggests that Form A is a general cold-temperature isoform. The frequency of occurrence of Form A also increases with latitude in *G. duebeni*, and it is the dominant form in the boreal populations of both *G. finmarchicus* and *E. obtusatus*. Form A has the longest loop length suggesting higher flexibility, lower thermostability and reduced activation energy that would favour increased sliding velocities at low temperatures (Chaen et al., 1996; Watabe, 2002). Furthermore, Form A is similar to fast loop 2 sequences from other cold-adapted gammaridean amphipods including *Eulimnogammarus verrucosus* from Lake Baikal (Holmes et al., 2002), and *Tryphosella sp.* from the Antarctic (J. Rock, unpubl. data). Likewise, the data suggest that Form B might be a warm-temperature isoform, as it was detected most frequently at lower latitudes and in the high-shore species with the broadest southerly distribution, *G. duebeni*. Form B was also dominant in the mid-high intertidal species *G. marinus*, and in *G. locusta*, the species with lowest latitudinal distribution. The prevalence of Form B in the Arctic population of *G. setosus* is less easy to explain, although it may be indicative of a broader latitudinal range of this species (versus its Arctic congeneric *G. oceanicus*).

4.5. Molecular variability of gammarid MyHC in phylogenetic and ecological context.

The high levels of mitochondrial genetic divergence between the 7 gammarid species are consistent with other phylogenetic work on north Atlantic gammarids (Hou et al. 2007; Rock et al., 2007). Even though there was no overarching phylogenetic basis to the interspecific patterns of MyHC form diversity, *G. oceanicus* displayed unique patterns of *MyHC* expression for both loop 1 and loop 2. The maintenance of these patterns across latitudes implicates a taxonomic signal in gammarid *MyHC* sequences. The pattern is, however, not shared by the most closely related species, *G. setosus*, which forms a distinct phylogenetic clade with *G. oceanicus*, and also shares a high latitudinal distribution. Links between physiological adaptation, phylogeny, and molecular variation within the gammarids have been previously investigated with isozyme analysis, and have indicated some differentiation between *G. oceanicus*, *G. duebeni*, and *G. locusta* (Siegismund et al., 1985; Skadsheim and Siegismund 1986). These studies suggested that the level of genetic differentiation was correlated with habitat differentiation with respect to salinity, as species that tended to have more isolated populations, and thus lower migration rates (e.g. *G. duebeni* because of its tolerance to low salinity), were more differentiated than less tolerant species (e.g. *G. oceanicus*). Our work certainly confirms such patterns of differentiation in MyHC isoforms between these two species, although our phylogenetic analyses suggest that *G. locusta* has the longest history of

Please cite this article as: Rock, J., et al., Linking functional molecular variation with environmental gradients: Myosin gene diversity in a crustacean broadly distributed across variable thermal environments, Gene (2009), doi:10.1016/j.gene.2009.02.008

reproductive isolation of the aforementioned species (Rock et al. ms submitted elsewhere).

Thus, despite deep phylogenetic divergence between species, and significant variation in intertidal habitat, latitudinal distribution patterns and predicted levels of temperature tolerance, overall myosin diversity among most gammarids appears to consist of a largely conserved gene family. The patterns of *MyHC* form expression described here appear to be stable at least across short term thermal shifts. Laboratory acclimated individuals did not demonstrate any variation in form expression from wild-caught animals for temperate *G. duebeni* acclimated for three weeks to temperatures differing by +5 or −13°C from their habitat temperature, or for Arctic *G. setosus* acclimated for a week to temperature 3°C warmer than their habitat temperature (J. Rock and N.M. Whiteley, unpubl. data). Thus, there appear to be two strategies adopted by gammarids depending on differences in thermal experiences over spatial and temporal scales. Due to the wide range of temperatures experienced in the intertidal habitat plus the relatively rapid rate of change, it seems likely that evolution has favoured greater MyHC isoform diversity with a number of isoforms enabling performance at differing thermal ranges. In the low intertidal species, evolution has favoured the expression of a general cold-water MyHC isoform capable of functioning over a relatively wide temperature range, but possibly sub-optimally. It seems that the temporal temperature variations associated with higher intertidal habitat may be a greater selective agent for MyHC isoform diversity than broad spatial changes with latitude. Ultimately, more information about other regions of the *MyHC* gene is needed to assess the contribution of this gene to adaptive divergence in gammarids. However, as our study demonstrates, characterising the adaptive response of wild organisms is complex and using the candidate gene approach is not always straightforward. Indeed, it is now becoming clear that a system-wide transcriptional response is involved in temperature adaptation in ectotherms, controlling a complex adaptive phenotype (Gracey et al., 2004; Cui et al., 2007) with the majority of adaptations attributable to changes in regulation rather than in gene sequences (Hoekstra and Coyne, 2007; Marden 2008). Regardless of the methodology employed, our data clearly demonstrate the necessity of a multi-level approach for assessing molecular variability, including examination of diversity at the intra- and inter-individual, population, and interspecific levels to identify gene–environment interactions.

Acknowledgments

We thank, Nick Cox, The Station Manager at NERC's Arctic Research Station in Ny-Ålesund, Svalbard and NERC for supporting the collection of Arctic gammarids. We also thank Dr Jørgen Christiansen at the Norwegian College of Fishery Science for the logistical support during the sampling of gammarids in Tromsø, and to Miranda Lowe at the Natural History Museum, London, for the taxonomic advice. Our research was supported by the Natural Environmental Research Council (NERC), UK (grant no. 01319).

References

Altschul, S.F., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Bernstein, S.I., Milligan, R.A., 1997. Fine tuning a molecular motor: the location of alternative domains in the *Drosophila* myosin head. J. Mol. Biol. 271, 1–6.

Bulnheim, H.P., 1979. Comparative studies on the physiological ecology of five euryhaline *Gammarus* species. Oecologia 44, 80–86.

Canepari, M., Rossi, R., Pellegrino, M.A., Bottinelli, R., Schiaffino, S., Reggiani, C., 2000. Functional diversity between orthologous myosins with minimal sequence diversity. J. Muscle Res. Cell Motil. 21, 375–382.

Chaen, S., Nakaya, M., Guo, X.F., Watabe, S., 1996. Lower activation energy for sliding velocity of F-actin on a less thermostable isoform of carp myosin. J. Biochem. 120, 788–791.

Cui, Q., Yu, Z., Purisima, E.O., Wang, E., 2007. MicroRNA regulation and interspecific variation of gene expression. Trends Evol. Ecol. 23, 372–375.

Fenchel, T.M., Kolding, S., 1979. *Habitat selection* and distribution. patterns of five species of the amphipod genus *Gammarus*. Oikos 33, 316–322.

Furch, M., Geeves, M.A., Manstein, D.J., 1998. Modulation of actin affinity and actomyosin adenosine triphosphate by charge changes in the motor domain. Biochemistry 37, 6317–6326.

Gaston, K.J., Spicer, J.I., 2001. The relationship between range size and niche breadth: a test using five species of *Gammarus* (Amphipoda). Global Ecol. Biogeog. 10, 179–188.

George, E.L., Ober, M.B., Emerson Jr, C.P., 1989. Functional domains of the *Drosophila melanogaster* muscle myosin heavy chain gene are encoded by alternatively spliced exons. Mol. Cell. Biol. 9, 2957–2974.

Gerlach, G.F., Turray, L., Malik, T., Lida, J., Scutt, A., Goldspink, G., 1990. Mechanisms of temperature-acclimation in the carp — a molecular biology approach. Am. J. Physiol. 259, 237–244.

Gingeras, T.R., 2007. Origin of phenotypes: genes and transcripts. Genome Res. 17, 682–690.

Goldspink, G., 1998. Selective gene expression during adaption of muscle in response to different physiological demands. Comp. Biochem. Physiol. 120, 5–15.

Goodson, H.V., Warrickm, H.M., Spudich, J.A., 1999. Specialised conservation of surface loops of myosin: evidence that loops are involved in determining functional characteristics. J. Mol. Biol. 287, 173–185.

Gracey, A., 2004. Coping with cold: an integrative, multitissue analysis of the transcriptome of a poikilothermic vertebrate. Proc. Nat. Acad. Sci. 30, 6971–6975.

Graves, J.E., Somero, G.N., 1982. Electrophoretic and functional enzymic evolution in four species of eastern Pacific barracudas from different thermal environments. Evolution 36, 97–106.

Hirayama, Y., Sutoh, K., Watabe, S., 2000. Structure–function relationships of the two surface loops of myosin heavy chain isoforms from thermally acclimated carp. Biochem. Biophys. Res. Coms. 269, 237–241.

Hirayama, Y., Watabe, S., 1997. Structural differences in the crossbridge head of temperature-associated myosin subfragment-1 isoforms from carp fast skeletal muscle. Eur. J. Biochem. 246, 380–387.

Hirotsune, S., 2003. An expressed pseudogene regulates the messenger-RNA stability of its homologous coding gene. Nature 423, 91–96.

Hochachka, P.W., Somero, G.N., 2002. Biochemical adaption. Mechanisms and process. In: Hochachka, P.W., Somero, G.N. (Eds.), Physiological Evolution. Oxford University Press, Oxford.

Hoekstra, H.E., Coyne, J.A., 2007. The locus of evolution: evo devo and the genetics of adaptation. Evolution 61, 995–1016.

Holmes, J.M., Hilber, K., Galler, S., Neil, D.M., 1999. Shortening velocities of two biochemically defined fibre types of the Norway lobster, *Nephrops norwegicus*. J. Muscle Res. Cell Motil. 20, 265–278.

Holmes, J.M., Whiteley, N.M., Magnay, J.L., El Haj, A.J., 2002. Comparison of the variable loop regions of myosin heavy chain genes from Antarctic and temperate isopods. Comp. Biochem. Physiol. B 131, 349–359.

Hou, Z., Fu, J., Li, S., 2007. A molecular phylogeny of the genus *Gammarus* (Crustacea: Amphipoda) based on mitochondrial and nuclear gene sequences. Mol. Phylogen. Evol. 45, 596–611.

Imai, J.I., Hirayama, Y., Kikuchi, K., Kainuma, M., Watabe, S., 1997. cDNA cloning of myosin heavy chain isoforms from carp fast skeletal muscle and their gene expression associated with temperature acclimation. J. Exp. Biol. 200, 27–34.

Johnston, I.A., Temple, G.K., 2002. Thermal plasticity of skeletal muscle phenotype in ectothermic vertebrates and its significance for locomotory behaviour. J. Exp. Biol. 205, 2305–2322.

Klekowski, R.Z., Węsławski, J.M., 1991. *Atlas* of the *marine fauna* of Southern Spitsbergen. Polish Academy of Science Press, Gdansk.

Knetsch, M.L.W., Uyeda, T.Q.P., Manstein, D.J., 1999. Disturbed communication between actin- and nucleotide-binding sites in a myosin II with truncated 50/20-kDa junction. J. Biol. Chem. 274, 20133–20138.

Kolding, S., 1981. Habitat selection and life cycle characteristics of five species of the amphipod genus *Gammarus* in the Baltic. Oikos 37, 173–178.

Kolding, S., 1985. Genetic adaptation to local habitats and speciation processes within the genus *Gammarus* (Amphipoda: Crustacea). Mar. Biol. 89, 249–255.

Kolding, S., Fenchel, T.M., 1981. Patterns of reproduction in different populations of five species of the amphipod genus *Gammarus*. Oikos 37, 167–172.

Kronert, W.A., Edwards, K.A., Roche, E.S., Wells, L., Bernstein, S.I., 1991. Muscle-specific accumulation of *Drosophila* myosin heavy chains: a splicing mutation in an alternative exon results in an isoform substitution. The EMBO J. 10, 2479–2488.

Kronert, W.A., Dambacher, C.A., Knowles, A.F., Swank, D.M., Bernstein, S.I., 2008. Alternative relay domains of *Drosophila melanogaster* myosin differentially affect ATPase activity, *in vitro* motility, myofibril structure and muscle function. J. Mol. Biol. 379, 443–456.

Liang, C.-S., 2007. Fast skeletal muscle myosin heavy chain gene cluster of medaka *Oryzias latipes* enrolled in temperature adaptation. Physiol. Genomics 29, 201–214.

Lincoln, R.J., 1979. British Marine Amphipoda: Gammaridea. British Museum Natural History, London.

Marden, J.H., 2008. Quantitative and evolutionary biology of alternative splicing: how changing the mix of alternative transcripts affects phenotypic plasticity and reaction norms. Heredity 100, 113–119.

Meng, S.-W., Zhang, Z., Li, J., 2004. Twelve C2H2 zinc-finger genes on human chromosome 19 can be each translated into the same type of protein after frameshifts. Bioinformatics 20, 1–4.

Murphy, C.T., Spudich, J.A., 1999. The sequence of the myosin 50–20 loop effects myosins affinity for actin throughout the actin-myosin ATPase cycle and its maximum ATPase activity. Biochemistry 38, 3785–3792.

Murphy, C.T., Spudich, J.A., 2000. Variable surface loops and myosin activity: accessories to a motor. J. Muscle Res. Cell Motil. 21, 139–151.

Please cite this article as: Rock, J., et al., Linking functional molecular variation with environmental gradients: Myosin gene diversity in a crustacean broadly distributed across variable thermal environments, Gene (2009), doi:10.1016/j.gene.2009.02.008

Nyitray, L., Jancso, A., Ochiai, Y., Graf, L., Szent-Györgyi, 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. U. S. A. 91, 12686–12690.
Ogonowski, M.M., Lang, F., 1979. Histochemical evidence for enzyme differences in crustacean fast and slow muscle. J. Exp. Zool. 207, 143–151.
Perreault-Micale, C.L., Kalabokis, V.N., Nyitray, L., Szent-Györgyi, A.G., 1996. Sequence variations in the surface loop near the nucleotide binding site modulate the ATP turnover rtes of molluscan myosins. J. Muscle Res. Cell Motil. 17, 543–553.
Pierce, V.A., Crawford, D.L., 1997. Phylogenetic analysis of thermal acclimation of the glycolytic enzymes in the genus *Fundulus*. Physiol. Zool. 70, 597–609.
Powers, D.A., Smith, M., Gonzalez-Villasenor, I., Dimichele, L., Crawford, D., 1993. A multidisciplinary approach to the selectionist/neutralist controversy using the model teleost *Fundulus heteroclitus*. In: Futuyuma, D., Antonovics, J. (Eds.), Oxford Survey of Evolutionary Biology, 9. Oxford University Press, Oxford, pp. 43–107.
Rayment, I., 1993. Three-dimensional structure of myosin subfragment-1: a molecular motor. Science 261, 50–58.
Rock, J., Ironside, J., Potter, T., Whiteley, N.M., Lunt, D.H., 2007. Phylogeography and environmental diversification of a highly adaptable marine amphipod, *Gammarus duebeni*. Heredity 99, 102–111.
Schulte, P.M., 2004. Changes in gene expression as biochemical adaptations to environmental change: tribute to Peter Hochachka. Comp. Biochem. Physiol. B. 139, 519–529.
Schulte, P.M., Glemet, H.C., Fiebig, A.A., Powers, D.A., 2000. Adaptive variation in lactate dehydrogenase-B gene expression: role of a stress-responsive regulatory element. Proc. Natl. Acad. Sci. U. S. A. 97, 6597–6602.
Shigemoto, K., 2001. Identification and characterisation of a developmentally regulated mammalian gene that utilises −1 programmed ribosomal frameshifting. Nucleic Acids Res. 29, 4079–4088.
Siegismund, H.R., Simonsen, V., Kolding, S., 1985. Genetic studies of *Gammarus*. 1. Genetic differentiation of local populations. Hereditas 102, 1–13.
Skadsheim, A., Siegismund, H.R., 1986. Genetic relationships among north-western European Gammaridea (Amphipoda). Crustaceana 51, 163–175.
Somero, G.N., 2002. Thermal physiology and vertical zonation of intertidal animals: optima, limits and costs of living. Integ. Comp. Biol. 42, 780–789.
Swank, D.M., 2002. The myosin converter domain modulates muscle performance. Nature Cell Biol. 4, 312–316.
Sweeney, H.L., 1998. Kinetic tuning of myosin via a flexible loop adjacent to the nucleotide binding pocket. J. Biol. Chem. 273, 6262–6270.
Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596–1599.
Tao, Y., Kobayashi, M., Liang, C.-S., Okamoto, T., Watabe, S., 2004. Temperature-dependent expression patterns of grass carp fast skeletal myosin heavy chain genes. Comp. Biochem. Phys. B 139, 649–656.
Uyeda, T.Q.P., Ruppel, K.M., Spudich, J.A., 1994. Enzymic activities correlate with chimeric substitutions at the actin-binding face of myosin. Nature 368, 567–569.
Watabe, S., 2002. Temperature plasticity of contractile proteins in fish muscle. J. Exp. Biol. 205, 2231–2236.
Watabe, S., 2008. Temperature acclimation induces light meromyosin isoforms with different primary structures in carp fast skeletal muscle. Biochem. Biophys. Res. Comm. 208, 118–125.
Yamamoto, K., 1989. Binding manner of actin to the lysine-rich sequence of myosin subfragment-1 in the presence and absence of ATP. Biochemistry 28, 5573–5577.
Zheng, D., Gerstein, M.B., 2007. The ambiguous boundary between genes and pseudogenes: the dead rise up, or do they? Trends Gen. 23, 219–224.

Please cite this article as: Rock, J., et al., Linking functional molecular variation with environmental gradients: Myosin gene diversity in a crustacean broadly distributed across variable thermal environments, Gene (2009), doi:10.1016/j.gene.2009.02.008