© 2017. Published by The Company of Biologists Ltd | Journal of Experimental Biology (2017) 220, 4109-4118 doi:10.1242/jeb.164988

RESEARCH ARTICLE

# Rapid expansion of pigmentation genes in penaeid shrimp with absolute preservation of function

**Alyssa M. Budd[1,*], Tracey M. Hinton[2], Mark Tonks[3], Sue Cheers[1] and Nicholas M. Wade[1,4,‡]**

## ABSTRACT

Crustaceans form their distinct patterns and colours through the interaction of the carotenoid astaxanthin with a protein called crustacyanin (CRCN). Presently, the expression of just two *CRCN* genes is thought to provide the protein subunits that combine to form the crustacyanin complex and associated carotenoid colour change from red to blue. This study aimed to explore the genetic complexity underlying the production of pigmentation and camouflage in penaeid shrimp. We isolated 35 new *CRCN* genes from 12 species, and their sequence analysis indicated that this gene family has undergone significant expansion and diversification in this lineage. Despite this duplication and sequence divergence, the structure of the CRCN proteins and their functional role in shrimp colour production has been strictly conserved. Using *CRCN* isoforms from *Penaeus monodon* as an example, we showed that isoforms were differentially expressed, and that subtle phenotypes were produced by the specific downregulation of individual isoforms. These findings demonstrate that our knowledge of the molecular basis of pigmentation in shrimp was overly simplistic, and suggests that multiple copies of the *CRCN* genes within species may be advantageous for colour production. This result is of interest for the origin and evolution of pigmentation in crustaceans, and the mechanisms by which gene function is maintained, diversified or sub-functionalized.

**KEY WORDS: Colouration, Molecular evolution, Crustacyanin, Crustacean**

## INTRODUCTION

The carotenoprotein complex called crustacyanin (CRCN) is an evolutionary novelty among the animals in class Crustacea as the mechanism for producing shell colour and patterns (Wade et al., 2009). The carotenoid astaxanthin forms the central chromophore of crustacyanin (Britton and Goodwin, 1982), where it is non-covalently bound but twisted by the protein interaction to produce the diverse array of crustacean colours (Chayen et al., 2003; Helliwell, 2010). Five distinct carotenoid-binding proteins (CBPs) were originally identified from the European lobster (*Homarus gammarus*) and these proteins were grouped into the broad classes CRCN-A and CRCN-C based upon amino acid composition, electrophoretic mobility and peptide mapping (Cheesman et al., 1966; Quarmby et al., 1977). The protein forms a dimer composed of two 20 kDa CRCN subunits and two astaxanthin (Axn) molecules, known as β-crustacyanin, which in turn form a large multimeric complex, called α-crustacyanin, composed of eight dimeric β-crustacyanin subunits (Chayen et al., 2003; Cheesman et al., 1966; Habash et al., 2004; Helliwell, 2010; Zagalsky and Cheesman, 1963).

Within the shell and hypodermal tissue of crustaceans, two CRCN proteins and two Axn molecules dimerise to form β-crustacyanin, and these dimers form the building blocks of the much larger multimeric α-crustacyanin (Chayen et al., 2003; Zagalsky, 2003). The 3D crystal structures of both α-crustacyanin and β-crustacyanin have been resolved on several independent occasions, as summarised in Table 1. Current research suggests that the colour shift from red to blue when astaxanthin is bound to CRCN is induced by acid–base change in carotenoid conformation, associated with the relocation of a proton on the end ring of Axn (Begum et al., 2015). This theory is distinct from the physical twisting of Axn or electron coupling within the protein-bound form investigated in other studies (Christensson et al., 2013; Strambi and Durbeej, 2009).

The full protein sequence was identified by protein purification and direct sequencing methods for the predominant CRCN A and C subunits (Keen et al., 1991a,b). It is thought that post-translational modifications, for example amidation or glycosylation, are responsible for the differences in electrophoretic mobility of subunits (Habash et al., 2004). In the American lobster (*Homarus americanus*) only two protein isoforms H1 and H2 have been identified (Zagalsky and Tidmarsh, 1985). However, there are only two genes reported to encode the different CRCN protein subunits (Chayen et al., 2003; Wade et al., 2009). A single *CRCN-A* and *CRCN-C* gene has been identified from a number of different crustaceans, including spiny lobsters (Wade et al., 2009), clawed lobsters (Ferrari et al., 2012; Wade et al., 2009), hermit crabs (Wade et al., 2009) and penaeid shrimp (Ertl et al., 2013; Wade et al., 2009). Only one gene (either A or C) has been identified in some species (Wade et al., 2009; Yang et al., 2011), although the inability to identify genes does not definitively indicate the absence of that gene in that species. Meanwhile, a duplication of the *CRCN-A* gene was identified from the spiny lobster *Panulirus cygnus*, with no corresponding *CRCN-C* gene identified (Wade et al., 2009). As such, the number of genes encoding functional CRCN proteins in crustaceans remains unclear.

In the present study, 35 new genes that encode CRCN proteins were identified across 12 species of penaeid shrimp, using a combination of degenerate PCR and bioinformatics data mining. These sequences were used to establish potential lineage-specific gene duplications, and predict three-dimensional (3D) structures of the encoded proteins using known crystal structures. The giant tiger shrimp (*Penaeus monodon* Fabricius 1798) was used as a model species to understand the functional regulation of three of the putative *CRCN* isoforms (*PmonCRCN-A*, *PmonCRCN-C1* and

[1]CSIRO Agriculture and Food, Queensland Biosciences Precinct, St Lucia, Queensland 4067, Australia. [2]CSIRO Health and Biosecurity, Australian Animal Health Laboratories, Geelong, Victoria 3220, Australia. [3]CSIRO Oceans and Atmosphere, Ecosciences Precinct, Dutton Park, Queensland 4102, Australia. [4]ARC Research Hub for Advanced Prawn Breeding, James Cook University, Townsville, Queensland 4811, Australia.
*Present address: Centre for Sustainable Tropical Fisheries and Aquaculture, College of Science and Engineering, James Cook University, Townsville, Queensland 4811, Australia.

‡Author for correspondence (nick.wade@csiro.au)

N.M.W., 0000-0001-7235-6274

**Received 20 June 2017; Accepted 23 August 2017**

Journal of Experimental Biology (2017) 220, 4109-4118 doi:10.1242/jeb.164988

**Table 1. Summary of the crustacyanin protein entries in the Protein Data Bank (PDB), and their literature references**

| Protein name | PDB code | Organism | Resolution (Å) | Residue count | Macromolecule | Citation |
|---|---|---|---|---|---|---|
| Alpha-crustacyanin C1 | 1I4U | *Homarus gammarus* | 1.15 | 362 | Crustacyanin | Gordon et al., 2001 |
| Alpha-crustacyanin A1 | 1H91 | *Homarus gammarus* | 1.4 | 360 | Crustacyanin A1 subunit | Cianci et al., 2001 |
| Beta-crustacyanin | 1GKA | *Homarus gammarus* | 3.23 | 354 | Crustacyanin A1 subunit<br>Crustacyanin A2 subunit | Cianci et al., 2002 |
| Alpha-crustacyanin C1 | 10BQ | *Homarus gammarus* | 1.85 | 362 | Crustacyanin C1 subunit | Habash et al., 2003 |
| Alpha-crustacyanin C1 | 10BU | *Homarus gammarus* | 2.0 | 362 | Crustacyanin C1 subunit | Habash et al., 2003 |
| Alpha-crustacyanin C2 | 1S44 | *Homarus gammarus* | 1.6 | 360 | Crustacyanin A1 subunit | Habash et al., 2004 |
| Alpha-crustacyanin C2 | 1S2P | *Homarus gammarus* | 1.3 | 362 | Crustacyanin C2 subunit | Habash et al., 2004 |
| Alpha-crustacyanin H1 | 4ALO | *Homarus americanus* | 2.37 | 362 | Crustacyanin H1 | Ferrari et al., 2012 |

*PmonCRCN-C2*) using reverse transcription PCR (RT-PCR) in pigmented hypodermal tissue and across the moult cycle. The systemic injection of double-stranded RNA (dsRNA) against each specific *PmonCRCN* isoform was used to trigger the RNA interference (RNAi) pathway, and the isoform-specific pigment phenotype was tracked in live animals using colour quantification from digital images. The functional integration of each of the *PmonCRCN* isoforms within CRCN complexes was confirmed using microscopy of pigmentory chromatophores.

## MATERIALS AND METHODS

### Animal collection and RNA extraction

Animal samples were opportunistically collected as part of CSIRO trawl surveys in the Gulf of Carpentaria in 2013 and 2014 (Kenyon et al., 2015). Animals were euthanized by immersion in an ice–seawater slurry for several minutes. Species were identified by experienced researchers equipped with species-specific reference material (Grey et al., 1983). Once collected and identified, small pieces of hypodermal tissue were removed and stored in RNA*later*™ (Invitrogen, Carlsbad, CA, USA). Samples were preserved and extracted from three to four individuals, with the exception of the red spot king prawn (*Penaeus longistylus* Kubo 1943), where only one animal was captured.

Total RNA was extracted using Trizol reagent (Invitrogen) according to the manufacturer's instructions, and precipitated by adding 0.5 volumes of isopropyl alcohol and 0.5 volumes of RNA precipitation solution for purity improvement (Green et al., 2012). Total RNA was DNase digested with the Turbo DNA-free kit (Applied Biosystems, Foster City, CA, USA) and quality and quantity were assessed by gel electrophoresis and on a NanoDrop spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). All RNA samples were diluted to 200 ng $\mu l^{-1}$ using an epMotion 5070 (Eppendorf, Hamburg, Germany). Reverse transcription was performed on 1 µg total RNA using Superscript III (Invitrogen) with 25 µmol $l^{-1}$ oligo $(dT)_{20}$ and 25 µmol $l^{-1}$ random hexamers.

### Sequence isolation and analysis

Partial fragments of *CRCN* gene sequences were isolated using several penaeid degenerate *CRCN-A* or *CRCN-C* primer pairs (Table 2) designed on a region of sequence conserved across the Pacific white shrimp [*Litopenaeus vannamei* (Boone 1931): *LvanCRCN-A*, CV468194; *LvanCRCN-C*, CV468290] and the banana prawn [*Fenneropenaeus merguiensis* (de Man 1888): *FmerCRCN-A*, HM370278; *FmerCRCN-C*, HM370279]. Hypodermal cDNA from each species was prepared as described below, and a PCR amplification product of approximately 560 and 460 bp was obtained for *CRCN-A* and *CRCN-C*, respectively (data not shown). PCR fragments were cloned into a pGEM-T Easy (Promega, Madison, WI, USA) and sequenced in both directions using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). Several other *CRCN* gene sequences were reported in the NCBI database, from two previous studies in shrimp unrelated to pigmentation (Leu

**Table 2. Primers used for identification of crustacyanin (*CRCN*) sequences, and analysis in *P. monodon***

| Gene | Primer name | Purpose | Amplicon size (bp) | Sequence |
|---|---|---|---|---|
| *PenaeidCRCN-A* | *Penaeid CRCN-A F2* | Degenerate primer | 563 | 5′-TGTTGAAGGCACTCSTARCTG-3′ |
| | *Penaeid CRCN-A R2* | Degenerate primer | | 5′-GTAARMACAKTCGGATGWRTGAGG-3′ |
| *PenaeidCRCN-C* | *Penaeid CRCN-C F1* | Degenerate primer | 460 | 5′-CGTNGTNCCNGGAARRTGYC-3′ |
| | *Penaeid CRCN-C R2* | Degenerate primer | | 5′-GACGCCGATAYTCWTGAAGG-3′ |
| *PmonCRCN-A* | *PmCRCN-A Tq F1* | Taqman primer | 99 | 5′-AAGTCCGAGTTCGGCTTCGT-3′ |
| | *PmCRCN-A Tq F1* | Taqman primer | | 5′-AAAAGTCGACGCCGTTCCT-3′ |
| | *PmCRCN-A Tq FAM* | Taqman probe | | 5′-FAM-TACTCCACAGAATGCC-MGB-3′ |
| *PmonCRCN-C1* | *PmCRCN-C1 Tq F1* | Taqman primer | 62 | 5′-AAACTTGCAGACCAGTACCTGAGA-3′ |
| | *PmCRCN-C1 Tq R1* | Taqman primer | | 5′-GACGTCGACGCCGATATTCT-3′ |
| | *PmCRCN-C1 Tq FAM* | Taqman probe | | 5′-FAM-CTGCGAGGCCGCCT-MGB-3′ |
| *PmonCRCN-C2* | *PmCRCN-C2 Tq F1* | Taqman primer | 59 | 5′-TCGGGATACAACTTCGGCTATT-3′ |
| | *PmCRCN-C2 Tq R1* | Taqman primer | | 5′-ACTTGGTGAGCGGGAGAAGA-3′ |
| | *PmCRCN-C2 Tq FAM* | Taqman probe | | 5′-FAM-TTCCGACTTTGCCTTC-MGB-3′ |
| *PmonEF1a* | *PmEF1a Tq F1* | Taqman primer | 57 | 5′-TCGCTTCCGACTCGAAGAA-3′ |
| | *PmEF1a Tq R1* | Taqman primer | | 5′-ACCTGGGCGGTGAAGTCA-3′ |
| | *PmEF1a Tq VIC* | Taqman probe | | 5′-VIC-CCCAGCCAAGGAA-MGB-3′ |
| *PmonCRCN-A* | *PmCRCN-A dsRNA F1* | dsRNA | 369 | 5′-ATGCTGGTCGCTGGTATCAGG-3′ |
| | *PmCRCN-A dsRNA R1* | dsRNA | | 5′-AGGCACACCTGTCAATCGCTG-3′ |
| *PmonCRCN-C1* | *PmCRCN-C1 dsRNA F1* | dsRNA | 343 | 5′-CGTGGTGCCGGGAAGGTGTC-3′ |
| | *PmCRCN-C1 dsRNA R1* | dsRNA | | 5′-GGAGTAGATGCAGGAGAAGTTC-3′ |
| *PmonCRCN-C2* | *PmCRCN-C2 dsRNA F1* | dsRNA | 343 | 5′-CGTGGTGCCTGGAAGATGTC-3′ |
| | *PmCRCN-C2 dsRNA R1* | dsRNA | | 5′-GCTATAGATGCACGAGAAGTTC-3′ |



Journal of Experimental Biology

Journal of Experimental Biology (2017) 220, 4109-4118 doi:10.1242/jeb.164988

et al., 2007; Robinson et al., 2014). Sequence editing was performed in CLC Main Workbench (CLC Bio, Aarhus, Denmark). New sequences identified in this study have been submitted to GenBank using accession numbers KP790005–KP790007 and MF627611–MF627642, as shown in Table 3.

### *CRCN* alignments and sequence analysis

Protein and nucleotide alignments were performed using CLC Main Workbench (CLC Bio). Nucleotide and protein sequence comparisons across all species were performed on the minimum shared sequence across both *A* and *C* genes (Fig. S1). Only short transcripts could be amplified for several *CRCN-C* isoforms; therefore, the *Melicertus latisulcatus MlatCRCN-C1* and *MlatCRCN-C2*, the *Penaeus esculentus PescCRCN-C1* and *PescCRCN-C2*, and the *PmonCRCN-C4* sequences were compared only across the shorter sequence identified. Sequences were trimmed to equal length for phylogenetic tree construction, and separate alignments were performed for the CRCN-A (Fig. S2) and CRCN-C

**Table 3. Resolved *CRCN* sequences in penaeid shrimp**

| Gene abbreviation | Size | Coding sequence (aa) | Source | Accession number | Additional NCBI sequences |
|---|---|---|---|---|---|
| *Fenneropenaeus merguiensis* (de Man 1888) | | | | | |
| *FmerCRCN-A1* | Complete, 573 bp | 190 | *in silico* | HM370278 | JN683654 |
| *FmerCRCN-C1* | Partial, 420 bp | 139 | PCR | MF627629 | HM370279 |
| *FmerCRCN-C2* | Partial, 420 bp | 139 | PCR | MF627631 | HM370279 |
| *FmerCRCN-C3* | Partial, 420 bp | 139 | PCR | MF627630 | HM370279 |
| *Litopenaeus vannamei* (Boone 1931) | | | | | |
| *LvanCRCN-A1* | Complete, 634 bp | 190 | *in silico* | CV468194 | DQ858916 |
| *LvanCRCN-C1* | Complete, 694 bp | 201 | *in silico* | CV468290 | |
| *LvanCRCN-C2* | Complete, 755 bp | 197 | *in silico* | JR494407 | JR494425 |
| *LvanCRCN-C3* | Complete, 669 bp | 198 | *in silico* | FE049586 | |
| *Melicertus latisulcatus* (Kishinouye 1896) | | | | | |
| *MlatCRCN-A1* | Partial, 518 bp | 172 | PCR | MF627634 | |
| *MlatCRCN-C1* | Partial, 301 bp | 99 | PCR | MF627632 | |
| *MlatCRCN-C2* | Partial, 301 bp | 99 | PCR | MF627633 | |
| *Metapenaeus bennettae* (Racek and Dall 1965) | | | | | |
| *MbenCRCN-C1* | Partial, 420 bp | 139 | PCR | MF627635 | |
| *Metapenaeus endeavouri* (Schmidt 1926) | | | | | |
| *MendCRCN-A1* | Partial, 518 bp | 172 | PCR | MF627637 | |
| *MendCRCN-A2* | Partial, 518 bp | 172 | PCR | MF627636 | |
| *Metapenaeus ensis* (de Haan 1850) | | | | | |
| *MensCRCN-A1* | Partial, 509 bp | 169 | PCR | MF627625 | |
| *MensCRCN-A2* | Partial, 518 bp | 172 | PCR | MF627626 | |
| *MensCRCN-A3* | Partial, 518 bp | 172 | PCR | MF627627 | |
| *MensCRCN-A4* | Partial, 515 bp | 171 | PCR | MF627628 | |
| *MensCRCN-C1* | Partial, 460 bp | 153 | PCR | MF627621 | |
| *MensCRCN-C2* | Partial, 462 bp | 154 | PCR | MF627620 | |
| *Parapenaeopsis sculptilis* (Heller 1862) | | | | | |
| *PscuCRCN-C1* | Partial, 420 bp | 139 | PCR | MF627641 | |
| *PscuCRCN-C2* | Partial, 420 bp | 139 | PCR | MF627642 | |
| *Penaeus esculentus* (Haswell 1879) | | | | | |
| *PescCRCN-A1* | Partial, 558 bp | 185 | PCR | MF627614 | |
| *PescCRCN-C1* | Partial, 343 bp | 114 | PCR | MF627623 | |
| *PescCRCN-C2* | Partial, 343 bp | 114 | PCR | MF627622 | |
| *Penaeus longistylus* (Kubo 1943) | | | | | |
| *PlonCRCN-A1* | Partial, 546 bp | 182 | PCR | MF627615 | |
| *PlonCRCN-C1* | Partial, 462 bp | 154 | PCR | MF627624 | |
| *Penaeus monodon* (Fabricius 1798) | | | | | |
| *PmonCRCN-A1* | Complete, 629 bp | 190 | PCR and *in silico* | KP790005 | GO068457; GO069205; GO071154; GO077931; GO080976; JR225989; FJ498898 |
| *PmonCRCN-A2* | Partial, 622 bp | 185 | *in silico* | MF627617 | GO077347; GO078698; GO079441; GO080966; JR220102 |
| *PmonCRCN-A3* | Partial, 622 bp | 189 | *in silico* | MF627616 | DT624270; GO076474 |
| *PmonCRCN-C1* | Partial, 730 bp | 198 | PCR and *in silico* | KP790006 | JR227352; GO070224, |
| *PmonCRCN-C2* | Partial, 591 bp | 169 | PCR and *in silico* | KP790007 | GO080326 |
| *PmonCRCN-C3* | Partial, 673 bp | 195 | *in silico* | MF627612 | GO070135; GO075578; GO075579; GO076593; GO077222 |
| *PmonCRCN-C4* | Partial, 328 bp | 109 | *in silico* | MF627611 | JR203011; FJ498904 |
| *Penaeus semisulcatus* (de Haan 1850) | | | | | |
| *PsemCRCN-A1* | Partial, 558 bp | 185 | PCR | MF627613 | |
| *PsemCRCN-C1* | Partial, 460 bp | 153 | PCR | MF627618 | |
| *PsemCRCN-C2* | Partial, 462 bp | 153 | PCR | MF627619 | |
| *PsemCRCN-C3* | Partial, 420 bp | 139 | PCR | MF627638 | |
| *Trachypenaeus anchoralis* (Bate 1888) | | | | | |
| *TancCRCN-C1* | Partial, 417 bp | 138 | PCR | MF627640 | |
| *TancCRCN-C2* | Partial, 420 bp | 139 | PCR | MF627639 | |

Journal of Experimental Biology (2017) 220, 4109-4118 doi:10.1242/jeb.164988

(Fig. S3) protein sequences. Maximum likelihood trees from protein alignments were created using the neighbor joining model and WAG protein substitution models over 100 bootstrap replicates, and the *H. americanus* CRCN-A (DV774018) or CRCN-C (CN951851) protein translation as outgroups. The tertiary structures of the identified putative carotenoid-binding proteins were generated using the Modeller software from HHPRED alignments on HHpred servers (Eswar et al., 2007; Soding et al., 2005). Verify3D was used to check the quality of the models generated (Eisenberg et al., 1997), where compatibility of a 3D model was analyzed against the corresponding 1D amino acid sequence. The 3D structures were visualized and superimposed using the Matchmaker application of the UCSF Chimera software (Pettersen et al., 2004).

### Quantitative RT-PCR

Three *CRCN* isoforms identified by degenerate PCR from *P. monodon* (*PmonCRCN-A*, *PmonCRCN-C1* and *PmonCRCN-C2*) were targeted for quantitative RT-PCR analysis. Expression of each *CRCN* gene across the moult cycle was assessed in hypodermal tissue only from three *P. monodon* at each of the moult stages A, B, C, $D_0$, $D_1$, $D_2$ and $D_{3/4}$ [determined according to setal staging and epidermal withdrawal in uropods (Promwikorn et al., 2004)], as used previously (Wade et al., 2012). Expression of *PmonCRCN* genes was analyzed by quantitative RT-PCR using specific Taqman® (Applied Biosystems) primers and probes as shown in Table 2 using primers located outside the region used to create dsRNA constructs. The relative expression of each gene was normalized across samples using *PmonEF1a*, then to the average of a pool of all samples (tissue expression and dsRNA) or to the average expression at moult stage A (moult stage samples) and $\log_2$ transformed.

### Downregulation of *CRCN* isoforms using dsRNA

Partial sequences specific to the *PmonCRCN-A*, *PmonCRCN-C1* or *PmonCRCN-C2* genes were amplified using primers shown in Table 2, cloned into pGEM-T easy (Promega) then sub-cloned into the dual T7 dsRNA expression vector L4440 (Addgene, Cambridge, MA, USA) according to the manufacturer's instructions. To synthesize dsRNA, vector primers were designed to amplify the target genes contained within each plasmid, then the specific PCR products were converted to dsRNA using the MEGAscript T7 kit (Applied Biosystems) according to the manufacturer's instructions. Integrity and size of newly synthesized dsRNAs were quantified by 1.5% agarose/EtBr gel electrophoresis (data not shown). *Luciferase* (*Luc*) dsRNA was used as an exogenous downregulation control, as used previously (De Santis et al., 2011; Sellars et al., 2011).

Juvenile shrimp, *P. monodon*, were obtained from commercial farms and maintained at CSIRO Agriculture and Food laboratories at Bribie Island Research Centre. For all trials, filtered seawater was heated then pumped through the tanks at 0.8 l $min^{-1}$, maintaining water temperatures at 28°C and salinity at 35 g $l^{-1}$. Ten animals were held in each of six red polyethylene tanks that held 80 litres seawater in each. Red tanks were used as they had previously been observed to produce an intermediate coloured shrimp. The experiment was conducted indoors under low artificial light conditions and a 12 h:12 h light:dark photoperiod. Initial mass of the animals across all treatments was 6.13±0.63 g and shrimp in each tank were fed once per day on a commercial diet.

Shrimp (15 per treatment) were tail-muscle injected with the four experimental treatments: *Luc*-dsRNA (5 µg), *PmonCRCN-A*-dsRNA (5 µg), *PmonCRCN-C1*-dsRNA (5 µg) and *PmonCRCN-C2*-dsRNA (5 µg) in a total of 50 µl of shrimp saline solution (10 mmol $l^{-1}$ HEPES, 450 mmol $l^{-1}$ NaCl, 10 mmol $l^{-1}$ KCl, 10 mmol $l^{-1}$ EDTA pH 7.2–7.5). Shrimp colour change over time was quantified on days 0, 2, 4 and 7 using digital images as outlined previously (Wade et al., 2015, 2014). Two shrimp were randomly sampled from each treatment on day 4 and dissected for RNA extraction and quantitative RT-PCR analysis as outlined above. The remaining animals were photographed for colour quantification on day 7 and retained frozen. For microscopy analysis, the shell was removed from abdominal segments and chromatophores were photographed immediately using a Leica M165C stereomicroscope fitted with a Canon EOS 5D digital camera.

### Statistical analysis

Where comparison between individual measurements was required, statistical significance was assessed by single-factor ANOVA, followed by Fisher's test allowing 5% error. Statistical significance of the type of dsRNA injection, time after injection or an interaction of these two variables was assessed by two-way ANOVA, followed by Fisher's test allowing 5% error. All statistical analyses were performed using StatPlus:mac 2009 (AnalystSoft Inc., Walnut, CA, USA).

## RESULTS

### Identification of *CRCN* sequences in *P. monodon*

A 518 bp fragment was isolated by degenerate PCR from *P. monodon* and sequenced. Seven published sequences (GO068457, GO069205, GO071154, GO077931, GO080976, JR225989 and FJ498898) aligned with *CRCN-A1* identified in this study, and were combined with the identified sequence to produce a 607 bp fragment, which included a 555 bp region that encoded a 185 amino acid (aa) PmCRCN-A1 protein. Evidence of another *CRCN-A* isoform was also identified *in silico*, with a group of five sequences (GO077347, GO078698, GO079441, GO080966 and JR220102) that aligned to produce a 622 bp *CRCN-A2* sequence, which included a 558 bp region that encoded a 186 aa protein. The CRCN A1 and A2 proteins shared 87% identity, including a small 2 aa deletion after residue D71. A further four sequences (DT624270, GO076474, JR226983 and JR215171) were identified that were strongly homologous with *CRCN-A*. Together they aligned to form a 622 bp fragment that encoded a 189 aa protein that was highly similar to *PmonCRCN-A2* except it did not contain the deletion noted above. However, this putative *PmonCRCN-A3* isoform contained a large amount of sequence variation and resulted in five unknown amino acids within the coding sequence, which made alignment with each other and other isoforms difficult. Despite the presence *in silico*, no evidence of the expression of the potential *PmonCRCN-A2* or *PmonCRCN-A3* isoforms was detected by degenerate PCR of hypodermal tissue in this study.

Two *PmonCRCN-C* sequences were identified by degenerate PCR, and combined with other publicly available sequences to resolve four putative *CRCN-C* isoforms, named *PmonCRCN-C1*, *PmonCRCN-C2*, *PmonCRCN-C3* and *PmonCRCN-C4*. Four published sequences (JR203011, JR227352, GO070224 and FJ498904) matched with *PmonCRCN-C1*, while one (GO080326) matched with *PmonCRCN-C2*. Five EST sequences (GO070135, GO075578, GO075579, GO076593 and GO077222) formed a distinct potential *PmonCRCN-C3* isoform, and two remaining sequences (JR203011 and FJ498904) formed a potential *PmonCRCN-C4* isoform (Table 3). For the regions of overlap, *PmonCRCN-C1* shared 75%, 74% and 78% protein identity with the *PmonCRCN-C2*, *PmonCRCN-C3* and *PmonCRCN-C4* isoforms, respectively (Fig. S1). Meanwhile, *PmonCRCN-C2* shared 90% protein identity with the *PmonCRCN-C3* isoform and 72% identity with the *PmonCRCN-C4* isoform. The *PmonCRCN-C3* isoform was



Journal of Experimental Biology

Journal of Experimental Biology (2017) 220, 4109-4118 doi:10.1242/jeb.164988

68% similar to the *PmonCRCN-C4* isoform at the protein level. Similar to the *PmonCRCN-A2* isoform, the presence *in silico* of the *PmonCRCN-C3* or *PmonCRCN-C4* isoforms could not be confirmed by RT-PCR.

### Identification of CRCN sequences in penaeid shrimp

Degenerate PCR products were amplified from hypodermal tissue from 12 species of penaeid shrimp. These products were sequenced, and along with *in silico* analysis of transcriptomic data produced a total of 35 unique CRCN sequences (Table 3). A CRCN-A and CRCN-C subunit was identified in eight of those species, whereas only CRCN-C was identified in *Metapenaeus bennettae* Racek and Dall 1965, *Parapenaeopsis sculptilis* (Heller 1862) and *Trachypenaeus anchoralis* (Bate 1888), and only CRCN-A was identified in *Metapenaeus endeavouri* (Schmidt 1926). However, the absence of detection of either isoform does not indicate that the gene does not exist in that species, just that it was not found in this study.

In 10 of the 12 species analyzed, more than one copy of the *CRCN-A* or *CRCN-C* gene was identified by degenerate PCR. This included up to a total of six isoforms in *Metapenaeus ensis* (de Haan 1850) and *P. monodon*, and four isoforms in three other species. Only three isoforms were identified by degenerate PCR in *P. monodon*, while others were identified by resolution of high-throughput sequence data as outlined above. Regardless of species, the *CRCN-A* and *CRCN-C* genes shared approximately 50% nucleotide similarity and 30% amino acid similarity, in line with levels observed in clawed lobsters (Fig. S1). Sequences duplicated within the *CRCN-A* or *CRCN-C* isoforms displayed between 72.1 and 99.8% nucleotide and 68.2 and 99.3% protein similarity with other sequences from the same species. In several examples, sequences were more similar between species than within species. Indeed, no nucleotide or amino acid variance was detected between the sequences determined for *PlonCRCN-C1* and *MensCRCN-C1*.

The aligned protein sequences were trimmed and used to create two independent maximum likelihood trees for the *CRCN-A* and *CRCN-C* genes using the corresponding clawed lobster *H. americanus* (Hamer) sequence as a root node (Fig. 1). There was clear separation between intraspecific isoforms, including those with high sequence homology such as *MensCRCN-A2* and *A3*, *PscuCRCN-C1* and *C2*, and *PmonCRCN-C2* and *C3*. There was evidence of both lineage-specific and cross-lineage duplication of both the *CRCN-A* and *CRCN-C* genes. There did not appear to be a clear duplication event that reflected a recent common ancestor, and the present nomenclature did not attempt to reflect that evolution. One prevailing observation was the broad separation of the *CRCN-C* genes into two clades (Fig. 1B). This may suggest that the majority of penaeid shrimp would be expected to have one copy of the *CRCN-A* gene and at least two copies of the *CRCN-C* gene, with clear evidence that more are present in some species.

### Protein alignments and prediction of CRCN 3D structure

Protein sequences from the identified *CRCN-A* and *CRCN-C* isoforms were aligned with their respective sequences using the clawed lobster *HamerCRCN-A* and *HamerCRCN-C* genes as a

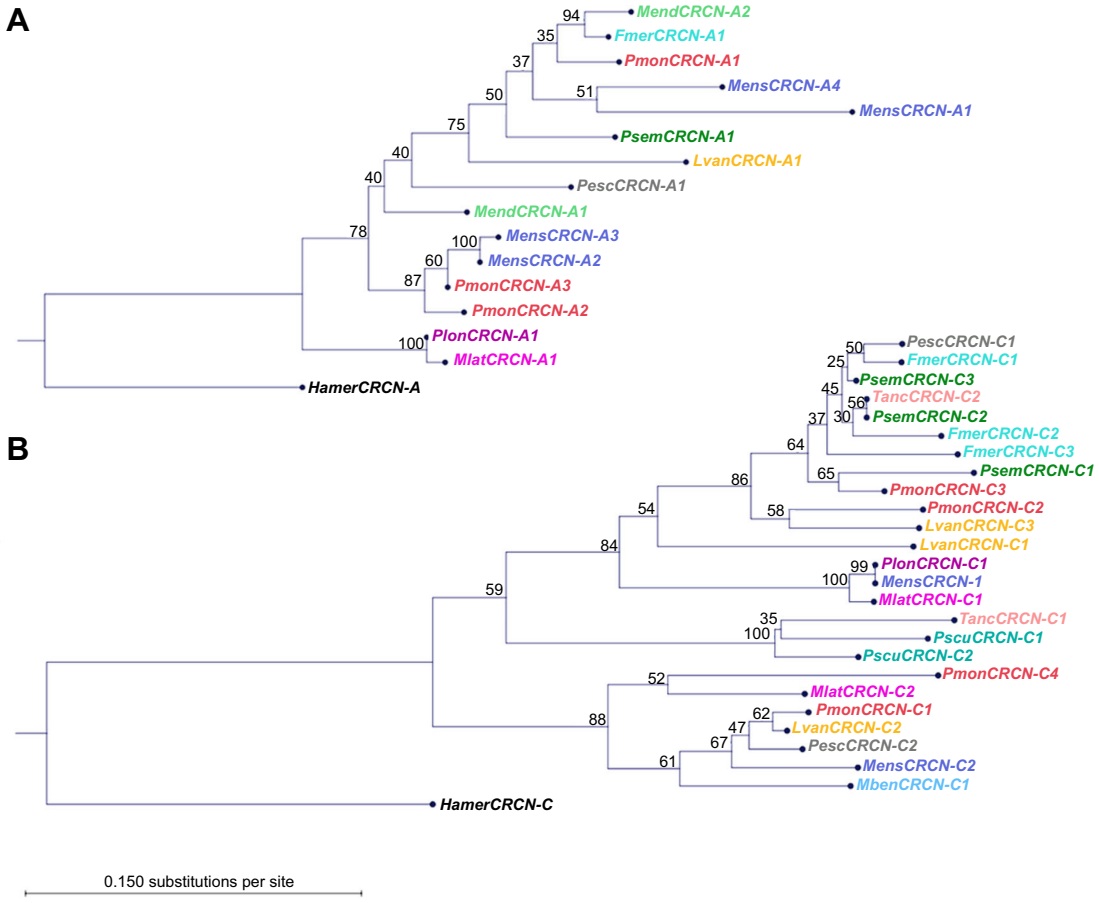



**Fig. 1. Maximum likelihood trees of amino acid alignments of identified *Penaeus monodon CRCN* isoforms identified in this study.** Trees were created independently for CRCN-A (A) or CRCN-C (B), with the root node set to the corresponding gene from *Homarus americanus* (*HamerCRCN*) and bootstrap replicate values at each node junction. Each species is displayed as a different colour, with gene abbreviations shown in Table 3.



Journal of Experimental Biology

Journal of Experimental Biology (2017) 220, 4109-4118 doi:10.1242/jeb.164988

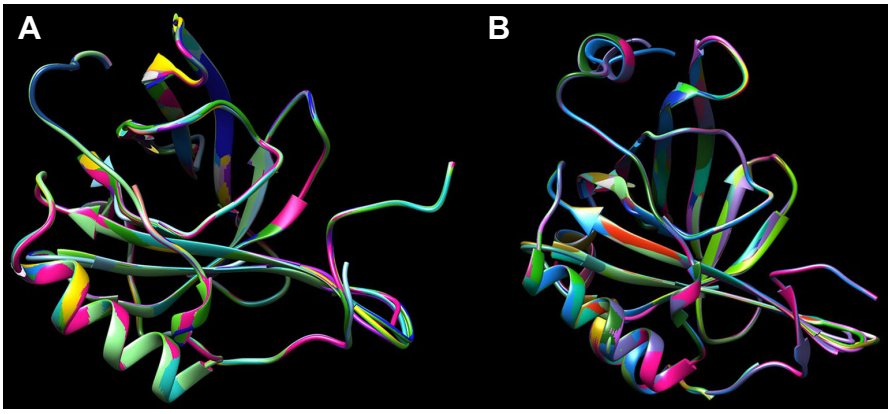



**Fig. 2. Predicted 3D reconstructions of each of two CRCN isoforms identified in this study.** (A) CRCN-A and (B) CRCN-C isoforms. Using the resolved tertiary structure for *Homarus gammarus* as a template (CRCN-A PDB ID: 1gka; CRCN-C PDB ID: 1i4u), each predicted CRCN structure was given a unique colour and superimposed on top of one another using Chimera software (Pettersen et al., 2004).

reference (Figs S2 and S3, respectively). The translated CRCN sequences are known to contain a short signal peptide and these were removed from the alignments. Protein sequence similarity with other penaeid shrimp was typically high, between 80 and 90% for CRCN-A and 70 and 75% for CRCN-C. This included absolute conservation of critical cysteine (*HamerCRCN-A*: C12-C120, C47-C171; *HamerCRCN-C*: C12-C121, C51-C173) and tryptophan (*HamerCRCN-A*: Y52) and histidine (*HamerCRCN-C*: H92) residues required for protein folding and interaction with the carotenoid chromophore. Broad regions of absolute sequence conservation were identified within penaeid shrimp, such as FAAPYEVIETDYDSYSCVYSC for CRCN-A and APYVILDTD-YENFSCIYSC for CRCN-C.

Given that the 3D structure has been resolved for clawed lobsters (PDB ID: 1i4u; Gordon et al., 2001), the 3D structure of the newly resolved sequences was predicted using Modeller. Structures for each gene were overlayed on each other using Chimera software and a different colour for each gene. Despite significant sequence changes, structural predictions revealed a complete preservation of tertiary structure for both the CRCN-A (Fig. 2A) and CRCN-C proteins (Fig. 2B). Any regions that contained variability or small deletions had minimal impact on tertiary structure, apart from slight shifts in loop structures on the exterior of the structure. The hydrophobic pocket that forms to enclose the carotenoid was completely preserved, as was the number and location of internally facing hydrophobic residues (Fig. 2C).

### Expression of PmonCRCN isoforms in shrimp hypodermal tissue

Three *CRCN* isoforms detected by degenerate PCR from *P. monodon* (*A*, *C1* and *C2*) were used to determine any potential isoform-specific expression patterns, and specific quantitative PCR primers were designed for each isoform. Across the moult cycle, the expression of the *PmonCRCN-A* and *PmonCRCN-C2* was stable (Fig. 3A), and was not significantly different. However, the expression of the *PmonCRCN-C1* isoform was significantly less expressed during the $D_0$ and $D_1$ intermoult period (Fig. 3A). Relative gene expression within the same tissue showed that *PmonCRCN-A* was approximately twofold more highly expressed than *PmonCRCN-C2* in shrimp hypodermal tissue, with twofold less expression of *PmonCRCN-C1* relative to *PmonCRCN-C2* (Fig. 3B).

### Functional downregulation of PmonCRCN isoforms

The RGB colour space is an additive colour space based on varying levels of the pigments red (R), green (G) and blue (B). The RGB values can be combined to produce a single colour that is representative of the average colour of a particular treatment that includes all the biological replicates (each box shown in Fig. 4). The

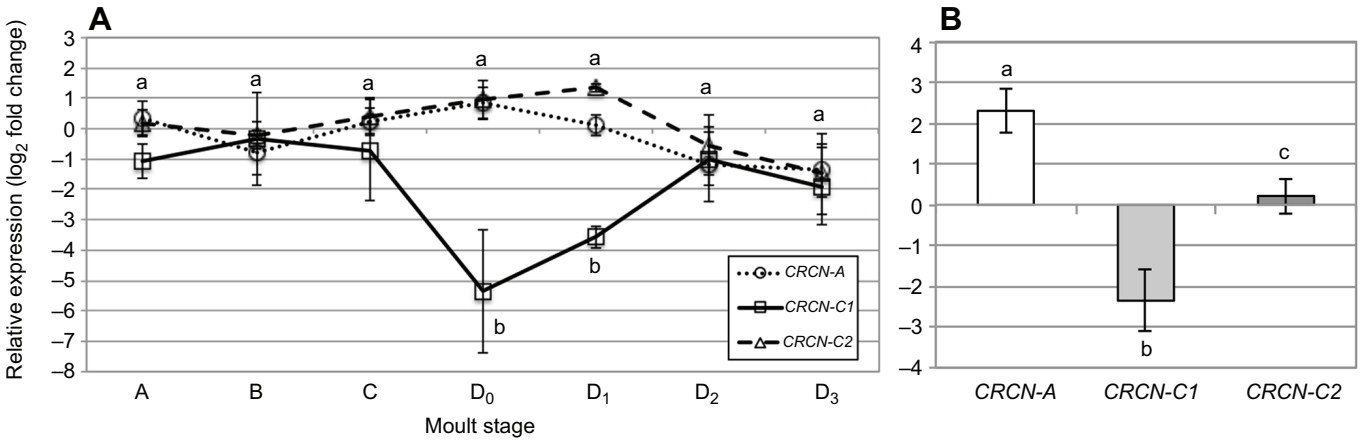



**Fig. 3. Relative expression of *CRCN* isoforms across the moult cycle and in hypodermal tissues of *P. monodon*.** (A) Relative expression of *CRCN* isoforms across the moult cycle of *P. monodon*. Expression values (means±s.e.m.) for each gene at each moult stage ($n$=3) were normalized relative to the expression of that gene across all moult stages. (B) Relative expression of *CRCN* isoforms in hypodermal tissues of *P. monodon*. Expression values (means±s.e.m.) for each gene ($n$=6) were normalized relative to the expression across all samples. Different lowercase letters denote significant ($P$<0.05) differences between the expression of different *CRCN* isoforms at the same moult stage, as well as the expression of the *PmonCRCN-C1* gene at different moult stages.



Journal of Experimental Biology

Journal of Experimental Biology (2017) 220, 4109-4118 doi:10.1242/jeb.164988

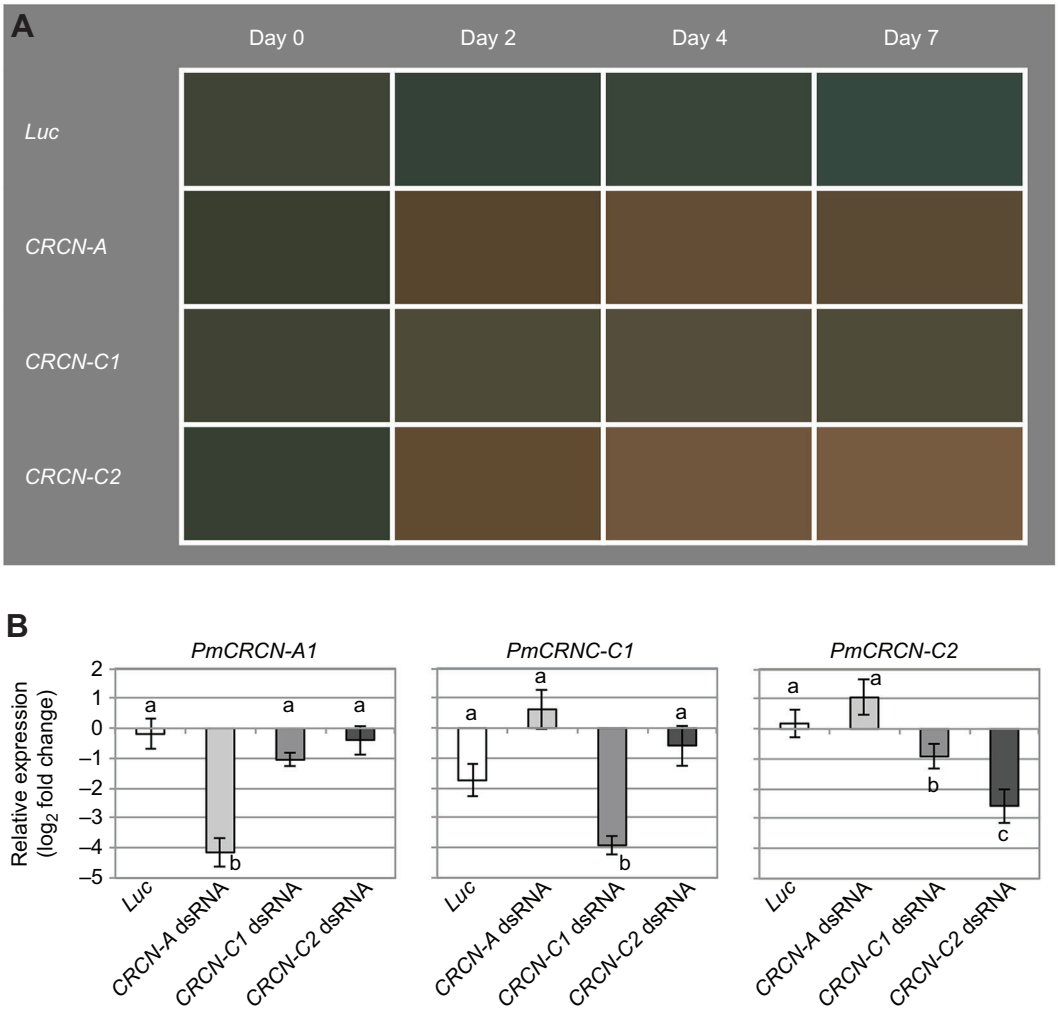



**Fig. 4. Colour change induced by specific downregulation of endogenous *CRCN* isoforms in *P. monodon* using RNA interference.** (A) The average uncooked abdominal colour quantified from all animals (three replicate tanks of five animals for each treatment, $n$=15 total) that were injected with dsRNA constructs targeting *LucLuc*, *PmonCRCN-A*, *PmonCRCN-C1* or *PmonCRCN-C2* to specifically downregulate expression of each isoform. Colour was tracked on the same live animals prior to injection (day 0) and after injection at days 2, 4 and 7 using digital images and the RGB colour space. (B) Expression of each *PmonCRCN* isoform within the abdominal hypodermis 4 days after injection of one of the four dsRNA products. Expression level ($log_2$ fold change, mean±s.e.m.) of each gene within each treatment ($n$=6) is shown relative to all samples across all treatments. Different lowercase letters denote significant ($P$<0.05) differences between the expression levels of the different genes.

average RGB values quantified for each group of replicates demonstrated that injection of dsRNA specific for each of the *PmonCRCN* isoforms produced a strong phenotypic colour change (Fig. 4A). Shrimp colour visibly changed, both relative to the type of dsRNA injected and over time. Analysis of the RGB values by two-way ANOVA showed that the type of dsRNA injected produced a significant change in R ($F$=42.6, $P$<0.001), G ($F$=8.1, $P$<0.001) and B values ($F$=9.1, $P$<0.001; Table S1). Time after injection produced a significant change in R ($F$=29.7, $P$<0.001), G ($F$=8.1, $P$<0.001) and B values ($F$=9.1, $P$<0.001), while a significant interaction effect was found between R ($F$=8.6, $P$<0.001) and G values ($F$=3.0, $P$=0.01). Injection with *PmonCRCN-A* dsRNA caused a significant increase in R and B values compared with values prior to injection (Table S1). Injection with *PmonCRCN-C1* dsRNA only caused a significant increase in R values, while injection with *PmonCRCN-C2* dsRNA caused a significant increase in R and G values, relative to values prior to injection (Table S1). Injection of animals with *Luc* dsRNA had no significant effect on RGB values (Table S1).

The specific downregulation of each isoform by their respective dsRNA target was confirmed by quantitative RT-PCR (Fig. 4B). Injection of dsRNA against each gene caused a significant three- to fourfold reduction in expression of the target gene relative to the *Luc* dsRNA-injected control or the other *PmonCRCN* genes (Fig. 4B). Injection of *PmonCRCN-C2* dsRNA also caused a significant reduction in expression of the *PmonCRCN-C1* gene, although expression of the *PmonCRCN-C2* gene was further reduced. As a result of dsRNA injection, hypodermal chromatophores reflected the hypodermal pigment changes (Fig. 5). Blue pigments were depleted in animals injected with *PmonCRCN-A* (Fig. 5D–F), *PmonCRCN-C1* (Fig. 5G–I) or *PmonCRCN-C2* dsRNA (Fig. 5J–L) compared with *Luc* dsRNA-injected controls (Fig. 5A–C). Injection of *PmonCRCN-C1* dsRNA removed blue pigment from some chromatophores (Fig. 5G–I, white arrowheads), but most were unaffected. Injection of *PmonCRCN-C2* dsRNA removed blue pigment from most of the chromatophores, with only a small number maintaining their distinct blue pigment (Fig. 5J–L, white arrowheads). A similar effect on pigmentation was recorded in the freshwater shrimp *Macrobrachium*

Journal of Experimental Biology (2017) 220, 4109-4118 doi:10.1242/jeb.164988

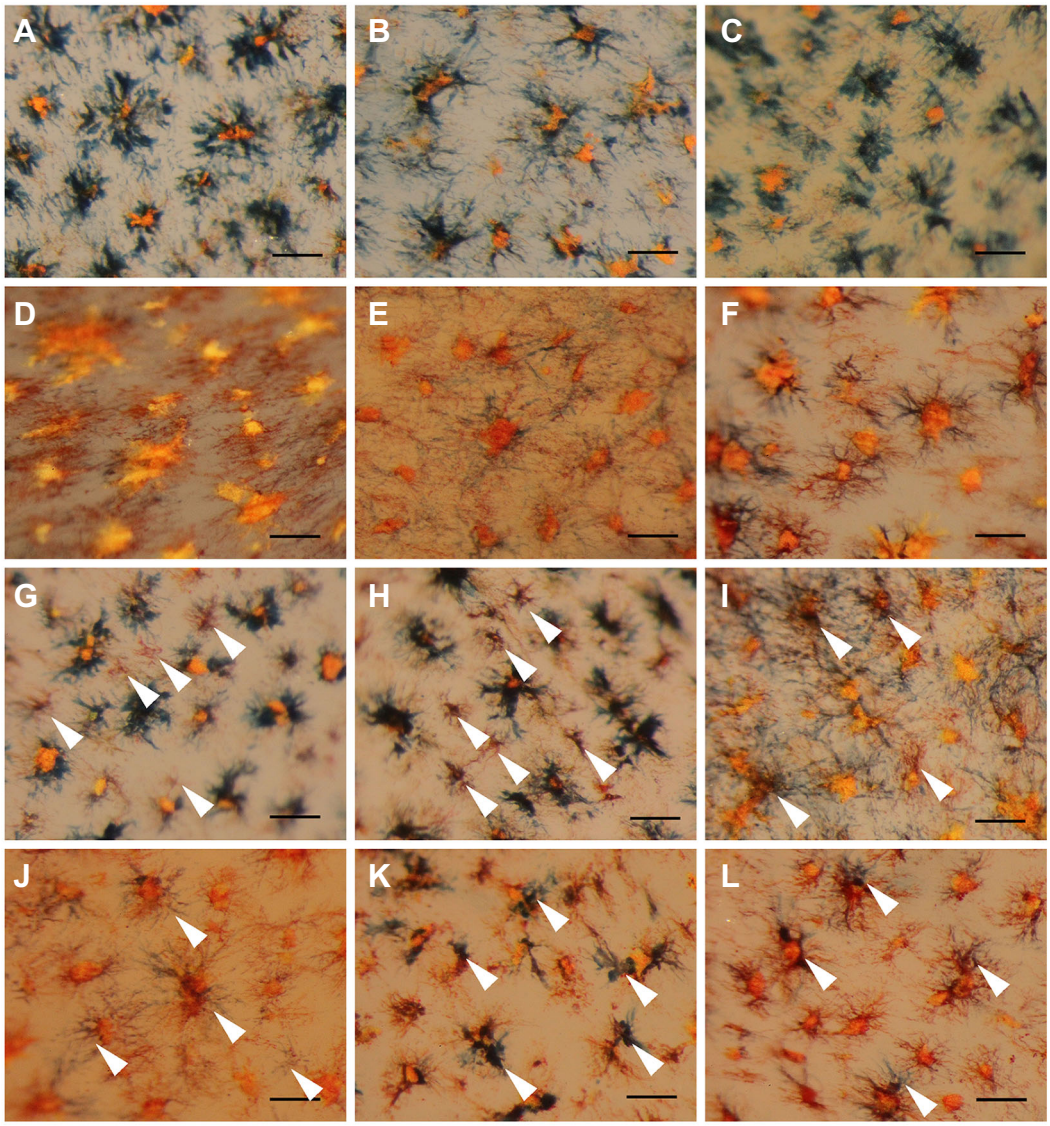



**Fig. 5. Hypodermal chromatophores in *P. monodon* after specific downregulation of endogenous *CRCN* isoforms using RNA interference.** Three replicates were randomly selected from photographs of the animals injected with dsRNA targets *Luciferase* (A–C), *PmonCRCN-A* (D–F), *PmonCRCN-C1* (G–I) or *PmonCRCN-C2* (J–L). Arrowheads denote chromatophores affected by *PmonCRCN-C1* injection (G–I) or not affected by *PmonCRCN-C2* injection (J–L).

*rosenbergii*, where injection of animals with the *MrosCRCN-C* isoform completely removed blue pigmentation from hypodermal chromatophores (Yang et al., 2011), but this pigmentation effect was not analyzed in response to *CRCN-A* dsRNA injection, or using specific *CRCN* isoforms.

## DISCUSSION

This study demonstrated that the *CRCN* genes that control colour production have undergone significant expansion and diversification in penaeid shrimp. However, despite this duplication and sequence divergence, their structure and functional role in shrimp colour production has been strictly conserved. Using *P. monodon* as an example, we show that multiple *CRCN* isoforms play a role in forming colours in the pigmented hypodermis of shrimp, that there is differential expression of these isoforms, and that subtle differential phenotypes are produced by the specific downregulation of each isoform. These duplications are unlikely to be maintained in a population unless they differ in some aspect of their efficacy, mutation rate or function (Hughes, 1994; Nowak et al., 1997). In the case of *CRCN*, the sequence of both *CRCN-A* and *CRCN-C* genes have been strictly conserved, suggesting that both subunits are required to maintain overall function. Multiple copies of the *CRCN* genes within species, particularly evident for *CRCN-C*, may allow increased expression of *CRCN* and potentially be advantageous for colour production. The ability of crustaceans to produce cryptic colours is critical for adaptive camouflage (Wade et al., 2009), although may not be as critical for penaeid shrimp that are often buried in muddy or sandy substrates.

To date, only one *CRCN-A* and one *CRCN-C* gene was thought to encode the various CRCN protein isoforms that have been identified in a range of crustaceans (Chayen et al., 2003; Wade et al., 2009). In a comprehensive transcriptomic assessment of pigmentation genes in the banana shrimp (*F. merguiensis*), only single copies of the *FmerCRCN-A* and *FmerCRCN-C* genes were identified (Ertl et al., 2013). The present study has identified several duplications of the *FmerCRCN-C* gene that were expressed within hypodermal tissue. Under the current *de novo* assemblies of transcriptome data, this type of isoform-specific information is compressed and lost, to the

detriment of our understanding of gene function. This study also putatively identified two further *PmonCRCN-A* (*A2* and *A3*) and two *PmonCRCN-C* (*C3* and *C4*) isoforms from *in silico* data. The presence of these other isoforms and any potential function in shrimp pigmentation awaits confirmation. The sequence difference between the *MensCRCN-A2* and *MensCRCN-A3* isoforms was eight nucleotides (99.8%) and four amino acids (99.3%). Such small sequence variation between isoforms or *in silico* assemblies may not have arisen from gene duplications, but may represent sequencing errors, normal polymorphic variation or the development of specific-sequence variations within individuals from disparate populations. A complete species genome may resolve some of these questions about isoforms, something that may be possible in the near future for some species. In any case, there is ample evidence from the number of sequences identified, and from the functional downregulation of the different *P. monodon* isoforms, to demonstrate that significant lineage-specific duplications have occurred.

The present study suggests that, at least in *P. monodon*, the β-crustacyanin is likely a heterodimer requiring the presence of both *CRCN-A* and *CRCN-C* isoforms. The downregulation of *PmCRNC-C1* expression removed the blue pigment from a small number of hypodermal chromatophores (Fig. 5G–I, arrowheads). Meanwhile, the blue pigment was retained within a complementary number of chromatophores after downregulation of the *PmCRCN-C2* isoform (Fig. 5J–L, arrowheads). This result was consistent with the lower level of endogenous expression of the *PmonCRCN-C1* isoform, and the more subtle colour phenotype of the animals injected with *PmonCRCN-C1* dsRNA. This result demonstrates that two different blue dimeric β-crustacyanin subunits are forming between PmonCRCN-A/C1 and PmonCRCN-A/C2, but essentially these CRCN-C subunits are performing the same function. This concept is supported by a past study with recombinant CRCN protein where both subunits combined to produce the strongest blue colour when reconstituted with astaxanthin (Ferrari et al., 2012), although β-crustacyanin homodimers of either CRCN-A or CRCN-C were also possible *in vitro*.

A small amount of blue colour also remained in the *PmonCRCN-A* dsRNA-injected animals, potentially because of incomplete downregulation of this gene in this tissue, as the expression of this gene was two- to fourfold higher than the other *CRCN* genes. Alternatively, one of the other *in silico PmonCRCN-A* isoforms may be present within these tissues that was not identified by degenerate PCR. The inability to detect other isoforms by PCR in this study does not indicate that they do not exist, simply that their expression is extremely low or that they have different spatial and/or temporal expression patterns. Irrespective of *in silico* analysis, this study demonstrates that there has been significant functional duplication of *CRCN* in penaeid shrimp. This study expands our understanding of the origin and evolution of pigmentation in crustaceans, and the mechanisms by which gene function is maintained, diversified or subfunctionalized.

#### Acknowledgements
The authors thank Richard Thaggard, Kiname Salee, Nick Polymeris and Megan Dearnley for trial maintenance, injections and sampling. The authors also thank Carmel McDougall and Melony Sellars for considered review of this manuscript.

#### Competing interests
The authors declare no competing or financial interests.

#### Author contributions
Conceptualization: A.M.B., T.M.H., N.W.; Methodology: A.M.B., T.M.H.; Validation: A.M.B.; Formal analysis: A.M.B.; Investigation: A.M.B., T.M.H., M.T., S.C.; Resources: M.T., S.C.; Data curation: N.W.; Writing - original draft: N.W.; Writing - review & editing: A.M.B., T.M.H., M.T., S.C.; Visualization: N.W.; Supervision: N.W.; Project administration: N.W.; Funding acquisition: M.T., N.W.

#### Funding
The authors acknowledge funding from DSM Nutritional Products Pty Ltd and the Commonwealth Scientific and Industrial Research Organisation (CSIRO), Agriculture and Food.

#### Data availability
New sequences identified in this study have been submitted to GenBank using accession numbers KP790005–KP790007 and MF627611–MF627642 (https://www.ncbi.nlm.nih.gov/genbank/).

#### Supplementary information
Supplementary information available online at
http://jeb.biologists.org/lookup/doi/10.1242/jeb.164988.supplemental

#### References
**Begum, S., Cianci, M., Durbeej, B., Falklöf, O., Hädener, A., Helliwell, J. R., Helliwell, M., Regan, A. C. and Watt, I. F.** (2015). On the origin and variation of colors in lobster carapace. *Phys. Chem. Chem. Phys.* **17**, 16723-16732.

**Britton, G. and Goodwin, T. W.** (1982). *Carotenoid Chemistry and Biochemistry*. Oxford: Pergamon Press.

**Chayen, N. E., Cianci, M., Grossmann, J. G., Habash, J., Helliwell, J. R., Nneji, G. A., Raftery, J., Rizkallah, P. J. and Zagalsky, P. F.** (2003). Unravelling the structural chemistry of the colouration mechanism in lobster shell. *Acta Crystallogr. Sec. D Biol. Crystallogr.* **59**, 2072-2082.

**Cheesman, D. F., Zagalsky, P. F. and Ceccaldi, H. J.** (1966). Purification and properties of crustacyanin. *Proc. R. Soc. B Biol. Sci.* **164**, 130-151.

**Christensson, N., Žídek, K., Magdaong, N. C. M., LaFountain, A. M., Frank, H. A. and Zigmantas, D.** (2013). Origin of the bathochromic shift of astaxanthin in lobster protein: 2D electronic spectroscopy investigation of beta-crustacyanin. *J. Phys. Chem. B* **117**, 11209-11219.

**Cianci, M., Rizkallah, P. J., Olczak, A., Raftery, J., Chayen, N. E., Zagalsky, P. F. and Helliwell, J. R.** (2001). Structure of lobster apocrustacyanin A1 using softer X-rays. *Acta Crystallogr. Sec. D Biol. Crystallogr.* **57**, 1219-1229.

**Cianci, M., Rizkallah, P. J., Olczak, A., Raftery, J., Chayen, N. E., Zagalsky, P. F. and Helliwell, J. R.** (2002). The molecular basis of the coloration mechanism in lobster shell: beta-crustacyanin at 3.2-A resolution. *Proc. Natl. Acad. Sci. USA* **99**, 9795-9800.

**De Santis, C., Wade, N. M., Jerry, D. R., Preston, N. P., Glencross, B. D. and Sellars, M. J.** (2011). Growing backwards: an inverted role for the shrimp ortholog of vertebrate myostatin and GDF11. *J. Exp. Biol.* **214**, 2671-2677.

**Eisenberg, D., Lüthy, R. and Bowie, J. U.** (1997). VERIFY3D: assessment of protein models with three-dimensional profiles. *Macromol. Crystallogr. Pt B* **277**, 396-404.

**Ertl, N. G., Elizur, A., Brooks, P., Kuballa, A. V., Anderson, T. A. and Knibb, W. R.** (2013). Molecular characterisation of colour formation in the prawn *Fenneropenaeus merguiensis*. *PLoS ONE* **8**, e56920.

**Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M.-Y., Pieper, U. and Sali, A.** (2007). Comparative protein structure modeling using MODELLER. *Curr. Protoc. Protein Sci.* **50**, 2.9.1-2.9.31.

**Ferrari, M., Folli, C., Pincolini, E., McClintock, T. S., Rössle, M., Berni, R. and Cianci, M.** (2012). Structural characterization of recombinant crustacyanin subunits from the lobster *Homarus americanus*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **68**, 846-853.

**Gordon, E. J., Leonard, G. A., McSweeney, S. and Zagalsky, P. F.** (2001). The C1 subunit of alpha-crustacyanin: the de novo phasing of the crystal structure of a 40 kDa homodimeric protein using the anomalous scattering from S atoms combined with direct methods. *Acta Crystallogr. Sec. D Biol. Crystallogr.* **57**, 1230-1237.

**Green, M. R., Sambrook, J. and Sambrook, J.** (2012). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

**Grey, D. L., Dall, W. and BAker, A.** (1983). *A guide to the Australian Penaeid Prawns*. Darwin: The Department of Primary Production of the Northern Territory.

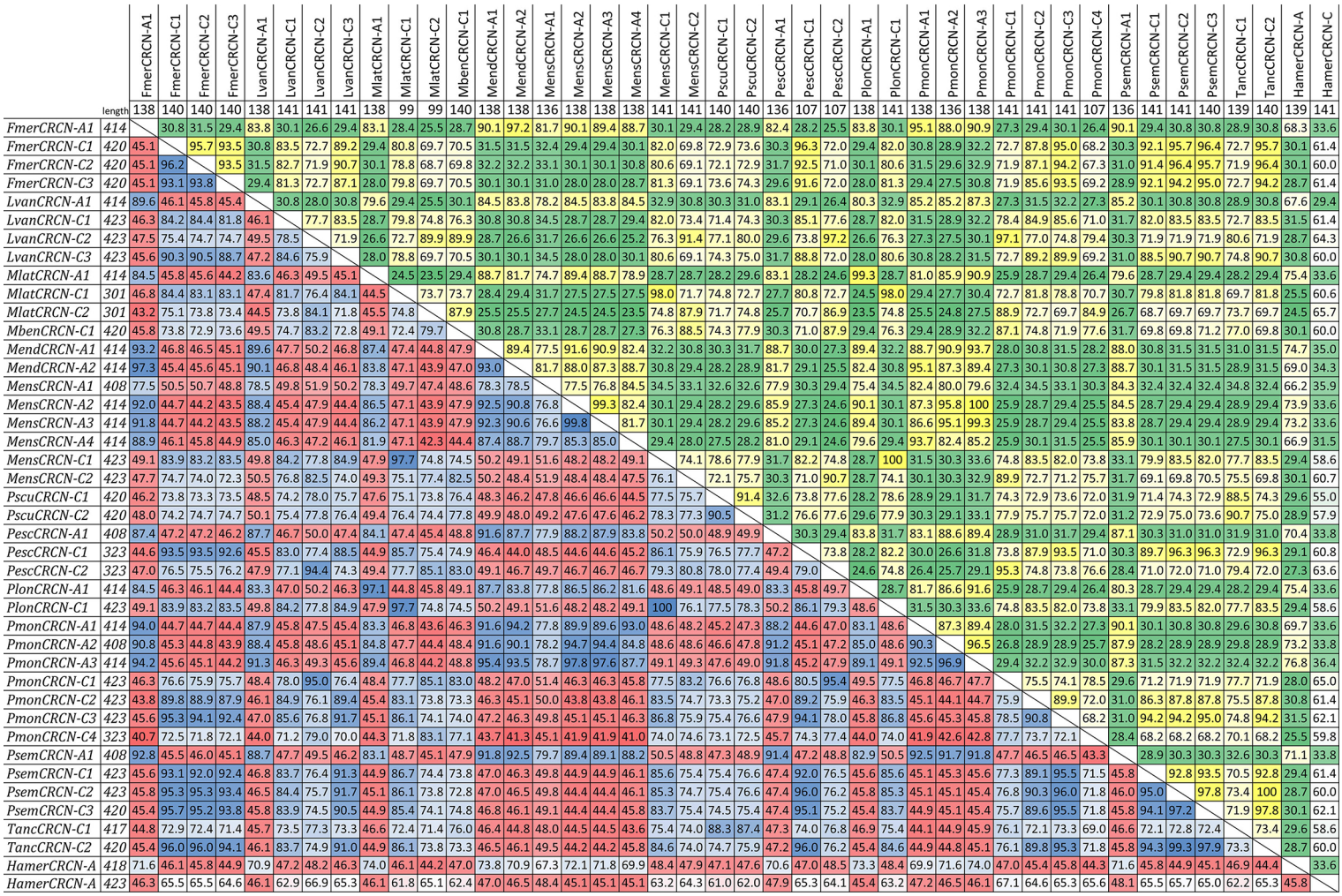
**Habash, J., Boggon, T. J., Raftery, J., Chayen, N. E., Zagalsky, P. F. and Helliwell, J. R.** (2003). Apocrustacyanin C-1 crystals grown in space and on earth using vapour-diffusion geometry: protein structure refinements and electron-density map comparisons. *Acta Crystallogr. Sec. D Biol. Crystallogr.* **59**, 1117-1123.

**Habash, J., Helliwell, J. R., Raftery, J., Cianci, M., Rizkallah, P. J., Chayen, N. E., Nneji, G. A. and Zagalsky, P. F.** (2004). The structure and refinement of apocrustacyanin C2 to 1.3 Å resolution and the search for differences between this protein and the homologous apoproteins A1 and C1. *Acta Crystallogr. Sec. D Biol. Crystallogr.* **60**, 493-498.

**Helliwell, J. R.** (2010). The structural chemistry and structural biology of colouration in marine Crustacea. *Crystallogr. Rev.* **16**, 231-242.

**Hughes, A. L.** (1994). The evolution of functionally novel proteins after gene duplication. *Proc. R. Soc. B Biol. Sci.* **256**, 119-124.
**Keen, J. N., Caceres, I., Eliopoulos, E. E., Zagalsky, P. F. and Findlay, J. B. C.** (1991a). Complete sequence and model for the A2 subunit of the carotenoid pigment complex, crustacyanin. *Eur. J. Biochem.* **197**, 407-417.
**Keen, J. N., Caceres, I., Eliopoulos, E. E., Zagalsky, P. F. and Findlay, J. B. C.** (1991b). Complete sequence and model for the C1 subunit of the carotenoprotein, crustacyanin, and model for the dimer, beta-crustacyanin, formed from the C1 and A2 subunits with astaxanthin. *Eur. J. Biochem.* **202**, 31-40.
**Kenyon, R. A., Ellis, N., Donovan, A. G., van der Velde, T. D., Fry, G., Tonks, M., Cheers, S. and Dennis, D.** (2015). *An integrated monitoring program for the Northern Prawn Fishery 2012-2015*. AFMA 2011/0811 Final Report. Brisbane: CSIRO Oceans and Atmosphere.
**Leu, J. H., Chang, C. C., Wu, J. L., Hsu, C. W., Hirono, I., Aoki, T., Juan, H. F., Lo, C. F., Kou, G. H. and Huang, H. C.** (2007). Comparative analysis of differentially expressed genes in normal and white spot syndrome virus infected Penaeus monodon. *BMC Genomics* **8**, 120.
**Nowak, M. A., Boerlijst, M. C., Cooke, J. and Smith, J. M.** (1997). Evolution of genetic redundancy. *Nature* **388**, 167-171.
**Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. and Ferrin, T. E.** (2004). UCSF chimera - a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605-1612.
**Promwikorn, W., Kirirat, P. and Thaweethamsewee, P.** (2004). Index of moult staging in the black tiger shrimp (*Penaeus monodon*). *Songklanakarin J. Sci. Technol.* **26**, 765-772.
**Quarmby, R., Nordens, D. A., Zagalsky, P. F., Ceccaldi, H. J. and Daumas, R.** (1977). Studies on the quaternary structure of the lobster exoskeleton carotenoprotein, crustacyanin. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **56**, 55-61.
**Robinson, N. A., Gopikrishna, G., Baranski, M., Katneni, V. K., Shekhar, M. S., Shanmugakarthik, J., Jothivel, S., Gopal, C., Ravichandran, P., Gitterle, T. et al.** (2014). QTL for white spot syndrome virus resistance and the sex-determining locus in the Indian black tiger shrimp (*Penaeus monodon*). *BMC Genomics* **15**, 731.
**Sellars, M. J., Rao, M., Arnold, S. J., Wade, N. M. Cowley, J. A.** (2011). *Penaeus monodon* is protected against gill-associated virus by muscle injection but not oral delivery of bacterially expressed dsRNAs. *Dis. Aquat. Org.* **95**, 19-30.
**Soding, J., Biegert, A. and Lupas, A. N.** (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **33**, W244-W248.
**Strambi, A. and Durbeej, B.** (2009). Excited-state modeling of the astaxanthin dimer predicts a minor contribution from exciton coupling to the bathochromic shift in crustacyanin. *J. Phys. Chem. B* **113**, 5311-5317.
**Wade, N. M., Tollenaere, A., Hall, M. R. and Degnan, B. M.** (2009). Evolution of a novel carotenoid-binding protein responsible for crustacean shell color. *Mol. Biol. Evol.* **26**, 1851-1864.
**Wade, N. M., Anderson, M., Sellars, M. J., Tume, R. K., Preston, N. P. and Glencross, B. D.** (2012). Mechanisms of colour adaptation in the prawn *Penaeus monodon*. *J. Exp. Biol.* **215**, 343-350.
**Wade, N. M., Paulo, C., Goodall, J., Fischer, M., Poole, S. and Glencross, B. D.** (2014). Quantitative methods to measure pigmentation variation in farmed Giant Tiger Prawns, *Penaeus monodon*, and the effects of different harvest methods on cooked colour. *Aquaculture* **433**, 513-519.
**Wade, N. M., Budd, A., Irvin, S. and Glencross, B. D.** (2015). The combined effects of diet, environment and genetics on pigmentation in the giant tiger prawn, *Penaeus monodon*. *Aquaculture* **449**, 78-86.
**Yang, F., Wang, M.-R., Ma, Y.-G., Ma, W.-M. and Yang, W.-J.** (2011). Prawn lipocalin: characterization of a color shift induced by gene knockdown and ligand binding assay. *J. Exp. Zool. A Ecol. Genet. Physiol.* **315A**, 562-571.
**Zagalsky, P. F.** (2003). beta-Crustacyanin, the blue-purple carotenoprotein of lobster carapace: consideration of the bathochromic shift of the protein-bound astaxanthin. *Acta Crystallogr. Sec. D Biol. Crystallogr.* **59**, 1529-1531.
**Zagalsky, P. F. and Cheesman, D. F.** (1963). Purification and properties of Crustacyanin. *Biochem. J.* **89**, P21-P35.
**Zagalsky, P. F. and Tidmarsh, M.-L.** (1985). Differences in the carapace astaxanthin proteins, crustacyanins, of the lobsters, *Homarus americanus* and *Homarus gammarus* (L.). *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **80**, 599-601.

## Supplementary Figure 1



| | | FmerCRCN-A1 | FmerCRCN-C1 | FmerCRCN-C2 | FmerCRCN-C3 | LvanCRCN-A1 | LvanCRCN-C1 | LvanCRCN-C2 | LvanCRCN-C3 | MlatCRCN-A1 | MlatCRCN-C1 | MlatCRCN-C2 | MbenCRCN-C1 | MendCRCN-A1 | MendCRCN-A2 | MensCRCN-A1 | MensCRCN-A2 | MensCRCN-A3 | MensCRCN-A4 | MensCRCN-C1 | MensCRCN-C2 | PscuCRCN-C1 | PscuCRCN-C2 | PescCRCN-A1 | PescCRCN-C1 | PescCRCN-C2 | PlonCRCN-A1 | PlonCRCN-C1 | PmonCRCN-A1 | PmonCRCN-A2 | PmonCRCN-A3 | PmonCRCN-C1 | PmonCRCN-C2 | PmonCRCN-C3 | PmonCRCN-C4 | PsemCRCN-A1 | PsemCRCN-C1 | PsemCRCN-C2 | PsemCRCN-C3 | TancCRCN-C1 | TancCRCN-C2 | HamerCRCN-A | HamerCRCN-C |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | length | 138 | 140 | 140 | 140 | 138 | 141 | 141 | 141 | 138 | 99 | 99 | 140 | 138 | 138 | 136 | 138 | 138 | 138 | 141 | 141 | 140 | 140 | 136 | 107 | 107 | 138 | 141 | 138 | 136 | 138 | 141 | 141 | 141 | 107 | 136 | 141 | 141 | 140 | 139 | 140 | 139 | 141 |
| *FmerCRCN-A1* | *414* | | 30.8 | 31.5 | 29.4 | 83.8 | 30.1 | 26.6 | 29.4 | 83.1 | 28.4 | 25.5 | 28.7 | 90.1 | 97.2 | 81.7 | 90.1 | 89.4 | 88.7 | 30.1 | 29.4 | 28.2 | 28.9 | 82.4 | 28.2 | 25.5 | 83.8 | 30.1 | 95.1 | 88.0 | 90.9 | 27.3 | 29.4 | 30.1 | 26.4 | 90.1 | 29.4 | 30.8 | 30.8 | 28.9 | 30.8 | 68.3 | 33.6 |
| *FmerCRCN-C1* | *420* | 45.1 | | 95.7 | 93.5 | 30.8 | 83.5 | 72.7 | 89.2 | 29.4 | 80.8 | 69.7 | 70.5 | 31.5 | 31.5 | 32.4 | 29.4 | 29.4 | 30.1 | 82.0 | 69.8 | 72.9 | 73.6 | 30.3 | 96.3 | 72.0 | 29.4 | 82.0 | 30.8 | 28.9 | 32.2 | 72.7 | 87.8 | 95.0 | 68.2 | 30.3 | 92.1 | 95.7 | 96.4 | 72.7 | 95.7 | 30.1 | 61.4 |
| *FmerCRCN-C2* | *420* | 45.1 | 96.2 | | 93.5 | 31.5 | 82.7 | 71.9 | 90.7 | 30.1 | 78.8 | 68.7 | 69.8 | 32.2 | 32.2 | 33.1 | 30.1 | 30.1 | 30.8 | 80.6 | 69.1 | 72.1 | 72.9 | 31.7 | 92.5 | 71.0 | 30.1 | 80.6 | 31.5 | 29.6 | 32.9 | 71.9 | 87.1 | 94.2 | 67.3 | 31.0 | 91.4 | 96.4 | 95.7 | 71.9 | 96.4 | 30.1 | 60.0 |
| *FmerCRCN-C3* | *420* | 45.1 | 93.1 | 93.8 | | 29.4 | 81.3 | 72.7 | 87.1 | 28.0 | 79.8 | 69.7 | 70.5 | 30.1 | 30.1 | 31.0 | 28.0 | 28.0 | 28.7 | 81.3 | 69.1 | 73.6 | 74.3 | 29.6 | 91.6 | 72.0 | 28.0 | 81.3 | 29.4 | 27.5 | 30.8 | 71.9 | 85.6 | 93.5 | 69.2 | 28.9 | 92.1 | 94.2 | 95.0 | 72.7 | 94.2 | 28.7 | 61.4 |
| *LvanCRCN-A1* | *414* | 89.6 | 46.1 | 45.8 | 45.4 | | 30.8 | 28.0 | 30.8 | 79.6 | 29.4 | 25.5 | 30.1 | 84.5 | 83.8 | 78.2 | 84.5 | 83.8 | 84.5 | 32.9 | 30.8 | 30.3 | 31.0 | 83.1 | 29.1 | 26.4 | 80.3 | 32.9 | 85.2 | 85.2 | 87.3 | 27.3 | 31.5 | 32.2 | 27.3 | 85.2 | 30.1 | 30.8 | 30.8 | 28.9 | 30.8 | 67.6 | 29.4 |
| *LvanCRCN-C1* | *423* | 46.3 | 84.2 | 84.4 | 81.8 | 46.1 | | 77.7 | 83.5 | 28.7 | 79.8 | 74.8 | 76.3 | 30.8 | 30.8 | 34.5 | 28.7 | 28.7 | 29.4 | 82.0 | 73.4 | 71.4 | 74.3 | 30.3 | 85.1 | 77.6 | 28.7 | 82.0 | 31.5 | 28.9 | 32.2 | 78.4 | 84.9 | 85.6 | 71.0 | 31.7 | 82.0 | 83.5 | 83.5 | 72.7 | 83.5 | 31.5 | 61.4 |
| *LvanCRCN-C2* | *423* | 47.5 | 75.4 | 74.7 | 74.7 | 49.5 | 78.5 | | 71.9 | 26.6 | 72.7 | 89.9 | 89.9 | 28.7 | 26.6 | 31.7 | 26.6 | 26.6 | 25.2 | 76.3 | 91.4 | 77.1 | 80.0 | 29.6 | 73.8 | 97.2 | 26.6 | 76.3 | 27.3 | 27.5 | 30.1 | 97.1 | 77.0 | 74.8 | 79.4 | 30.3 | 71.9 | 71.9 | 71.9 | 80.6 | 71.9 | 28.7 | 64.3 |
| *LvanCRCN-C3* | *423* | 45.6 | 90.3 | 90.5 | 88.7 | 47.2 | 84.6 | 75.9 | | 28.0 | 78.8 | 69.7 | 70.5 | 30.1 | 30.1 | 34.5 | 28.0 | 28.0 | 30.1 | 80.6 | 69.1 | 74.3 | 75.0 | 31.7 | 88.8 | 72.0 | 28.0 | 80.6 | 30.8 | 28.2 | 31.5 | 72.7 | 89.2 | 89.9 | 69.2 | 31.0 | 88.5 | 90.7 | 90.7 | 74.8 | 90.7 | 30.8 | 60.0 |
| *MlatCRCN-A1* | *414* | 84.5 | 45.8 | 45.6 | 44.2 | 83.6 | 46.3 | 49.5 | 45.1 | | 24.5 | 23.5 | 29.4 | 88.7 | 81.7 | 74.7 | 89.4 | 88.7 | 78.9 | 28.7 | 28.7 | 28.2 | 29.6 | 83.1 | 28.2 | 24.6 | 99.3 | 28.7 | 81.0 | 85.9 | 90.9 | 25.9 | 28.7 | 29.4 | 26.4 | 79.6 | 28.7 | 29.4 | 29.4 | 28.2 | 29.4 | 75.4 | 33.6 |
| *MlatCRCN-C1* | *301* | 46.8 | 84.4 | 83.1 | 83.1 | 47.4 | 81.7 | 76.4 | 84.1 | 44.5 | | 73.7 | 73.7 | 28.4 | 29.4 | 31.7 | 27.5 | 27.5 | 27.5 | 98.0 | 71.7 | 74.8 | 72.7 | 27.7 | 80.8 | 72.7 | 24.5 | 98.0 | 29.4 | 27.7 | 30.4 | 72.7 | 81.8 | 78.8 | 70.7 | 30.7 | 79.8 | 81.8 | 81.8 | 69.7 | 81.8 | 25.5 | 60.6 |
| *MlatCRCN-C2* | *301* | 43.2 | 75.1 | 73.8 | 73.4 | 44.5 | 73.8 | 84.1 | 71.8 | 45.5 | 74.8 | | 87.9 | 25.5 | 25.5 | 27.7 | 24.5 | 24.5 | 23.5 | 74.8 | 87.9 | 71.7 | 74.8 | 25.7 | 70.7 | 86.9 | 23.5 | 74.8 | 25.5 | 24.8 | 27.5 | 88.9 | 72.7 | 69.7 | 84.9 | 26.7 | 68.7 | 69.7 | 69.7 | 73.7 | 69.7 | 24.5 | 65.7 |
| *MbenCRCN-C1* | *420* | 45.8 | 73.8 | 72.9 | 73.6 | 49.5 | 74.7 | 83.2 | 72.8 | 49.1 | 72.4 | 79.7 | | 30.8 | 28.7 | 33.1 | 28.7 | 28.7 | 27.3 | 76.3 | 88.5 | 74.3 | 77.9 | 30.3 | 71.0 | 87.9 | 29.4 | 76.3 | 29.4 | 28.9 | 32.2 | 87.1 | 74.8 | 71.9 | 77.6 | 31.7 | 69.8 | 69.8 | 71.2 | 77.0 | 69.8 | 30.1 | 60.0 |
| *MendCRCN-A1* | *414* | 93.2 | 46.8 | 46.5 | 45.1 | 89.6 | 47.7 | 50.2 | 46.8 | 87.4 | 47.4 | 44.8 | 47.9 | | 89.4 | 77.5 | 91.6 | 90.9 | 82.4 | 32.2 | 30.8 | 30.3 | 31.7 | 88.7 | 30.0 | 27.3 | 89.4 | 32.2 | 88.7 | 90.9 | 93.7 | 28.0 | 30.8 | 31.5 | 28.2 | 88.0 | 30.8 | 31.5 | 31.5 | 31.0 | 31.5 | 74.7 | 35.0 |
| *MendCRCN-A2* | *414* | 97.3 | 45.4 | 45.6 | 45.1 | 90.1 | 46.8 | 48.4 | 46.1 | 83.8 | 47.1 | 43.9 | 47.0 | 93.0 | | 81.7 | 88.0 | 87.3 | 88.7 | 30.8 | 29.4 | 28.2 | 28.9 | 81.7 | 29.1 | 25.5 | 82.4 | 30.8 | 95.1 | 87.3 | 89.4 | 27.3 | 30.1 | 30.8 | 27.3 | 88.7 | 30.1 | 31.5 | 31.5 | 28.9 | 31.5 | 69.0 | 34.3 |
| *MensCRCN-A1* | *408* | 77.5 | 50.5 | 50.7 | 48.8 | 78.5 | 49.8 | 51.9 | 50.2 | 78.3 | 49.7 | 47.4 | 48.6 | 78.3 | 78.5 | | 77.5 | 76.8 | 84.5 | 34.5 | 33.1 | 32.6 | 32.6 | 77.9 | 30.3 | 29.4 | 75.4 | 34.5 | 82.4 | 80.0 | 79.6 | 32.4 | 34.5 | 33.1 | 30.3 | 84.3 | 32.4 | 32.4 | 32.4 | 34.8 | 32.4 | 66.2 | 35.9 |
| *MensCRCN-A2* | *414* | 92.0 | 44.7 | 44.2 | 43.5 | 88.4 | 45.4 | 47.9 | 44.4 | 86.5 | 47.1 | 43.9 | 47.9 | 92.5 | 90.8 | 76.8 | | 99.3 | 82.4 | 30.1 | 29.4 | 28.2 | 29.6 | 85.9 | 27.3 | 24.6 | 90.1 | 30.1 | 87.3 | 95.8 | 100 | 25.9 | 28.7 | 29.4 | 25.5 | 84.5 | 28.7 | 29.4 | 29.4 | 28.9 | 29.4 | 73.9 | 33.6 |
| *MensCRCN-A3* | *414* | 91.8 | 44.7 | 44.2 | 43.5 | 88.2 | 45.4 | 47.9 | 44.4 | 86.2 | 47.1 | 43.9 | 47.9 | 92.3 | 90.6 | 76.6 | 99.8 | | 81.7 | 30.1 | 29.4 | 28.2 | 29.6 | 85.2 | 27.3 | 24.6 | 89.4 | 30.1 | 86.6 | 95.1 | 99.3 | 25.9 | 28.7 | 29.4 | 25.5 | 83.8 | 28.7 | 29.4 | 29.4 | 28.9 | 29.4 | 73.2 | 33.6 |
| *MensCRCN-A4* | *414* | 88.9 | 46.1 | 45.8 | 44.9 | 85.0 | 46.3 | 47.2 | 46.1 | 81.9 | 47.1 | 42.3 | 44.4 | 87.4 | 88.7 | 79.7 | 85.3 | 85.0 | | 29.4 | 28.0 | 27.5 | 28.2 | 81.0 | 29.1 | 24.6 | 79.6 | 29.4 | 93.7 | 82.4 | 85.2 | 25.9 | 30.1 | 31.5 | 25.5 | 85.9 | 30.1 | 30.1 | 30.1 | 27.5 | 30.1 | 66.9 | 31.5 |
| *MensCRCN-C1* | *423* | 49.1 | 83.9 | 83.2 | 83.5 | 49.8 | 84.2 | 77.8 | 84.9 | 47.9 | 97.7 | 74.8 | 74.5 | 50.2 | 49.1 | 51.6 | 48.2 | 48.2 | 49.1 | | 74.1 | 78.6 | 77.9 | 31.7 | 82.2 | 74.8 | 28.7 | 100 | 31.5 | 30.3 | 33.6 | 74.8 | 83.5 | 82.0 | 73.8 | 33.1 | 79.9 | 83.5 | 82.0 | 77.7 | 83.5 | 29.4 | 58.6 |
| *MensCRCN-C2* | *423* | 47.7 | 74.7 | 74.0 | 72.3 | 50.5 | 76.8 | 82.5 | 74.0 | 49.3 | 75.1 | 77.4 | 82.5 | 50.2 | 48.4 | 51.9 | 48.4 | 48.4 | 47.5 | 76.1 | | 72.1 | 75.7 | 30.3 | 71.0 | 90.7 | 28.7 | 74.1 | 30.1 | 30.3 | 32.9 | 89.9 | 72.7 | 71.2 | 75.7 | 31.7 | 69.1 | 69.8 | 70.5 | 75.5 | 69.8 | 30.1 | 60.7 |
| *PscuCRCN-C1* | *420* | 46.2 | 73.8 | 73.3 | 73.5 | 48.5 | 74.2 | 78.0 | 75.7 | 47.6 | 75.1 | 73.8 | 76.4 | 48.3 | 46.2 | 47.8 | 46.6 | 46.6 | 44.5 | 77.5 | 75.7 | | 91.4 | 32.6 | 73.8 | 77.6 | 28.2 | 78.6 | 28.9 | 29.1 | 31.7 | 74.3 | 72.9 | 73.6 | 72.0 | 31.9 | 71.4 | 74.3 | 72.9 | 88.5 | 74.3 | 29.6 | 55.0 |
| *PscuCRCN-C2* | *420* | 48.0 | 74.2 | 74.7 | 74.7 | 50.1 | 75.4 | 77.8 | 76.4 | 49.4 | 76.4 | 74.4 | 77.8 | 49.9 | 48.0 | 49.2 | 47.6 | 47.6 | 46.2 | 78.3 | 77.3 | 90.5 | | 31.2 | 76.6 | 77.6 | 29.6 | 77.9 | 30.3 | 29.1 | 33.1 | 77.9 | 75.7 | 75.7 | 72.0 | 31.2 | 72.9 | 75.0 | 73.6 | 90.7 | 75.0 | 28.9 | 57.9 |
| *PescCRCN-A1* | *408* | 87.4 | 47.2 | 47.2 | 46.2 | 87.7 | 46.7 | 50.0 | 47.4 | 84.1 | 47.4 | 45.4 | 48.8 | 91.6 | 87.7 | 77.9 | 88.2 | 87.9 | 83.8 | 50.2 | 50.0 | 48.9 | 49.9 | | 30.3 | 29.4 | 83.8 | 31.7 | 83.1 | 88.6 | 89.4 | 28.9 | 31.0 | 31.7 | 29.4 | 87.1 | 30.3 | 31.0 | 31.0 | 31.9 | 31.0 | 70.4 | 33.8 |
| *PescCRCN-C1* | *323* | 44.6 | 93.5 | 93.5 | 92.6 | 45.5 | 83.0 | 77.4 | 88.5 | 44.9 | 85.7 | 75.4 | 74.9 | 46.4 | 44.0 | 48.5 | 44.6 | 44.6 | 45.2 | 86.1 | 75.9 | 76.5 | 77.7 | 47.2 | | 73.8 | 28.2 | 82.2 | 30.0 | 26.6 | 31.8 | 73.8 | 87.9 | 93.5 | 71.0 | 30.3 | 89.7 | 96.3 | 96.3 | 72.9 | 96.3 | 29.1 | 60.8 |
| *PescCRCN-C2* | *323* | 47.0 | 76.5 | 75.5 | 76.2 | 47.9 | 77.1 | 94.4 | 74.3 | 49.4 | 77.7 | 85.1 | 83.0 | 49.1 | 46.7 | 49.7 | 46.7 | 46.7 | 46.7 | 79.3 | 80.8 | 78.0 | 77.4 | 49.4 | 79.0 | | 24.6 | 74.8 | 26.4 | 25.7 | 29.1 | 95.3 | 74.8 | 73.8 | 76.6 | 28.4 | 71.0 | 72.0 | 71.0 | 79.4 | 72.0 | 27.3 | 63.6 |
| *PlonCRCN-A1* | *414* | 84.5 | 46.3 | 46.1 | 44.4 | 83.3 | 47.0 | 50.2 | 46.3 | 97.1 | 44.8 | 45.8 | 49.1 | 87.7 | 83.8 | 77.8 | 86.5 | 86.2 | 81.6 | 48.6 | 49.1 | 48.5 | 49.0 | 83.3 | 45.8 | 49.7 | | 28.7 | 81.7 | 86.6 | 91.6 | 25.9 | 28.7 | 29.4 | 26.4 | 80.3 | 28.7 | 29.4 | 29.4 | 28.2 | 29.4 | 75.4 | 33.6 |
| *PlonCRCN-C1* | *423* | 49.1 | 83.9 | 83.2 | 83.5 | 49.8 | 84.2 | 77.8 | 84.9 | 47.9 | 97.7 | 74.8 | 74.5 | 50.2 | 49.1 | 51.6 | 48.2 | 48.2 | 49.1 | 100 | 76.1 | 77.5 | 78.3 | 50.2 | 86.1 | 79.3 | 48.6 | | 31.5 | 30.3 | 33.6 | 74.8 | 83.5 | 82.0 | 73.8 | 33.1 | 79.9 | 83.5 | 82.0 | 77.7 | 83.5 | 29.4 | 58.6 |
| *PmonCRCN-A1* | *414* | 94.0 | 44.7 | 44.7 | 44.4 | 87.9 | 45.8 | 47.5 | 45.4 | 83.3 | 46.8 | 43.6 | 46.3 | 91.6 | 94.2 | 77.8 | 89.9 | 89.6 | 93.0 | 48.6 | 48.2 | 45.2 | 47.3 | 88.2 | 44.6 | 47.0 | 83.1 | 48.6 | | 87.3 | 89.4 | 28.0 | 31.5 | 32.2 | 27.3 | 90.1 | 30.1 | 30.8 | 30.8 | 29.6 | 30.8 | 69.7 | 33.6 |
| *PmonCRCN-A2* | *408* | 90.8 | 45.3 | 44.8 | 43.9 | 88.4 | 45.8 | 48.6 | 45.1 | 84.8 | 47.7 | 44.4 | 48.4 | 91.6 | 90.1 | 78.2 | 94.7 | 94.4 | 84.8 | 48.6 | 48.6 | 46.6 | 47.8 | 91.2 | 45.1 | 47.2 | 85.0 | 48.6 | 90.3 | | 96.5 | 26.8 | 28.9 | 28.9 | 25.7 | 87.9 | 28.2 | 28.9 | 28.9 | 29.8 | 28.9 | 73.2 | 33.8 |
| *PmonCRCN-A3* | *414* | 94.2 | 45.6 | 45.1 | 44.2 | 91.3 | 46.3 | 49.3 | 45.6 | 89.4 | 46.8 | 44.2 | 48.8 | 95.4 | 93.5 | 78.7 | 97.8 | 97.6 | 87.7 | 49.1 | 49.3 | 47.6 | 49.0 | 91.8 | 45.2 | 47.9 | 89.1 | 49.1 | 92.5 | 96.9 | | 29.4 | 32.2 | 32.9 | 30.0 | 87.3 | 31.5 | 32.2 | 32.2 | 32.4 | 32.2 | 76.8 | 36.4 |
| *PmonCRCN-C1* | *423* | 46.3 | 76.6 | 75.9 | 75.7 | 48.4 | 78.0 | 95.0 | 76.4 | 48.4 | 77.7 | 85.1 | 83.0 | 48.2 | 47.0 | 51.4 | 46.3 | 46.3 | 45.8 | 77.5 | 83.2 | 76.6 | 76.8 | 48.6 | 80.5 | 95.4 | 49.5 | 77.5 | 46.8 | 46.7 | 47.7 | | 75.5 | 74.1 | 78.5 | 29.6 | 71.2 | 71.9 | 71.9 | 77.7 | 71.9 | 28.0 | 65.0 |
| *PmonCRCN-C2* | *423* | 43.8 | 89.8 | 88.9 | 87.9 | 46.1 | 84.9 | 76.1 | 89.4 | 45.4 | 83.1 | 73.8 | 73.3 | 46.3 | 45.1 | 50.0 | 43.8 | 43.8 | 46.1 | 83.5 | 74.7 | 73.3 | 75.2 | 47.0 | 89.2 | 75.9 | 46.3 | 83.5 | 45.1 | 44.1 | 44.7 | 75.9 | | 89.9 | 72.0 | 31.0 | 86.3 | 87.8 | 87.8 | 75.5 | 87.8 | 30.8 | 61.4 |
| *PmonCRCN-C3* | *423* | 45.6 | 95.3 | 94.1 | 92.4 | 47.0 | 85.6 | 76.8 | 91.7 | 45.1 | 86.1 | 74.1 | 74.0 | 47.2 | 46.3 | 49.8 | 45.1 | 45.1 | 46.3 | 86.8 | 75.9 | 75.4 | 76.6 | 47.9 | 94.1 | 78.0 | 45.8 | 86.8 | 45.6 | 45.3 | 45.8 | 78.5 | 90.8 | | 68.2 | 31.0 | 94.2 | 94.2 | 95.0 | 74.8 | 94.2 | 31.5 | 62.1 |
| *PmonCRCN-C4* | *323* | 40.7 | 72.5 | 71.8 | 72.1 | 44.0 | 71.2 | 79.0 | 70.0 | 44.3 | 71.8 | 83.1 | 77.1 | 43.7 | 41.3 | 45.1 | 41.9 | 41.9 | 41.0 | 74.0 | 74.6 | 73.1 | 72.5 | 45.7 | 74.3 | 77.4 | 44.0 | 74.0 | 41.9 | 42.6 | 42.8 | 77.7 | 73.7 | 72.1 | | 28.4 | 68.2 | 68.2 | 68.2 | 70.1 | 68.2 | 25.5 | 59.8 |
| *PsemCRCN-A1* | *408* | 92.8 | 45.5 | 46.0 | 45.1 | 88.7 | 47.7 | 49.5 | 46.2 | 83.1 | 48.7 | 45.1 | 47.9 | 91.8 | 92.5 | 79.7 | 89.4 | 89.1 | 88.2 | 50.5 | 48.8 | 47.3 | 48.9 | 91.4 | 47.2 | 48.8 | 82.9 | 50.5 | 92.5 | 91.7 | 91.8 | 47.7 | 46.5 | 46.5 | 43.3 | | 28.9 | 30.3 | 30.3 | 32.6 | 30.3 | 71.1 | 33.8 |
| *PsemCRCN-C1* | *423* | 45.6 | 93.1 | 92.0 | 92.4 | 46.8 | 83.7 | 76.4 | 91.3 | 44.9 | 86.7 | 74.4 | 73.8 | 47.0 | 46.3 | 49.8 | 44.9 | 44.9 | 46.1 | 85.6 | 75.4 | 75.4 | 76.6 | 47.4 | 92.0 | 76.5 | 45.6 | 85.6 | 45.1 | 45.3 | 45.6 | 77.3 | 89.1 | 95.5 | 71.5 | 45.8 | | 92.8 | 93.5 | 70.5 | 92.8 | 29.4 | 61.4 |
| *PsemCRCN-C2* | *423* | 45.8 | 95.3 | 95.3 | 93.4 | 46.5 | 84.4 | 75.7 | 91.7 | 45.1 | 86.1 | 73.8 | 72.8 | 47.0 | 46.5 | 49.8 | 44.4 | 44.4 | 46.1 | 85.3 | 74.7 | 75.4 | 76.6 | 47.4 | 96.0 | 76.2 | 45.8 | 85.3 | 45.1 | 45.1 | 45.4 | 76.8 | 90.3 | 96.0 | 71.8 | 46.0 | 95.0 | | 97.8 | 73.4 | 100 | 28.7 | 60.0 |
| *PsemCRCN-C3* | *420* | 45.4 | 95.7 | 95.2 | 93.8 | 45.8 | 83.9 | 74.5 | 90.5 | 44.9 | 85.4 | 74.1 | 74.8 | 46.8 | 46.1 | 49.1 | 44.4 | 44.4 | 45.8 | 83.7 | 75.4 | 74.5 | 75.4 | 47.4 | 95.1 | 75.2 | 45.4 | 83.7 | 44.9 | 45.1 | 45.4 | 75.7 | 89.6 | 95.5 | 71.8 | 45.8 | 94.1 | 97.2 | | 71.9 | 97.8 | 30.1 | 62.1 |
| *TancCRCN-C1* | *417* | 44.8 | 72.9 | 72.4 | 71.4 | 45.7 | 73.5 | 77.3 | 73.3 | 46.6 | 72.4 | 71.4 | 76.0 | 46.4 | 44.8 | 48.0 | 44.5 | 44.5 | 43.6 | 75.4 | 74.0 | 88.3 | 87.4 | 47.3 | 74.0 | 76.8 | 46.9 | 75.4 | 44.1 | 44.9 | 45.9 | 76.1 | 72.1 | 73.3 | 69.0 | 46.6 | 72.1 | 72.8 | 72.4 | | 73.4 | 29.6 | 58.6 |
| *TancCRCN-C2* | *420* | 45.4 | 96.0 | 96.0 | 94.1 | 46.1 | 83.7 | 74.9 | 91.0 | 44.9 | 86.1 | 73.8 | 73.3 | 46.5 | 46.1 | 49.5 | 44.2 | 44.2 | 45.8 | 84.6 | 74.0 | 74.7 | 75.9 | 47.2 | 96.0 | 76.2 | 45.4 | 84.6 | 44.9 | 44.8 | 45.1 | 76.1 | 89.8 | 95.3 | 71.8 | 45.8 | 94.3 | 99.3 | 97.9 | 73.3 | | 28.7 | 60.0 |
| *HamerCRCN-A* | *418* | 71.6 | 46.1 | 45.8 | 44.9 | 70.9 | 47.2 | 48.2 | 46.3 | 74.0 | 46.1 | 44.2 | 47.0 | 73.8 | 70.9 | 67.3 | 72.1 | 71.8 | 69.9 | 48.4 | 47.9 | 47.1 | 47.6 | 70.6 | 47.0 | 48.5 | 73.3 | 48.4 | 69.9 | 71.6 | 74.0 | 47.0 | 45.4 | 45.8 | 44.3 | 71.6 | 45.8 | 44.9 | 45.1 | 46.9 | 44.4 | | 33.6 |
| *HamerCRCN-A* | *423* | 46.3 | 65.5 | 65.5 | 64.6 | 46.1 | 62.9 | 66.9 | 65.3 | 46.1 | 61.8 | 65.1 | 62.4 | 47.0 | 46.5 | 48.4 | 45.1 | 45.1 | 45.1 | 63.2 | 64.3 | 61.0 | 62.0 | 47.9 | 65.3 | 64.1 | 45.4 | 63.2 | 47.2 | 46.5 | 46.1 | 67.1 | 64.6 | 65.3 | 65.6 | 48.1 | 65.5 | 65.7 | 65.0 | 62.2 | 65.3 | 45.8 | |

**Fig. S1. Pairwise comparisons of each isoform identified in this study.** Below the diagonal line is nucleotide similarity (%) and above the line is amino acid similarity (%). Nucleotide similarity is highlighted as three colours: high (blue), intermediate (white) or low (red) while protein similarity is highlighted as: high (yellow), intermediate (white), low (green).

**Supplementary Figure 2**

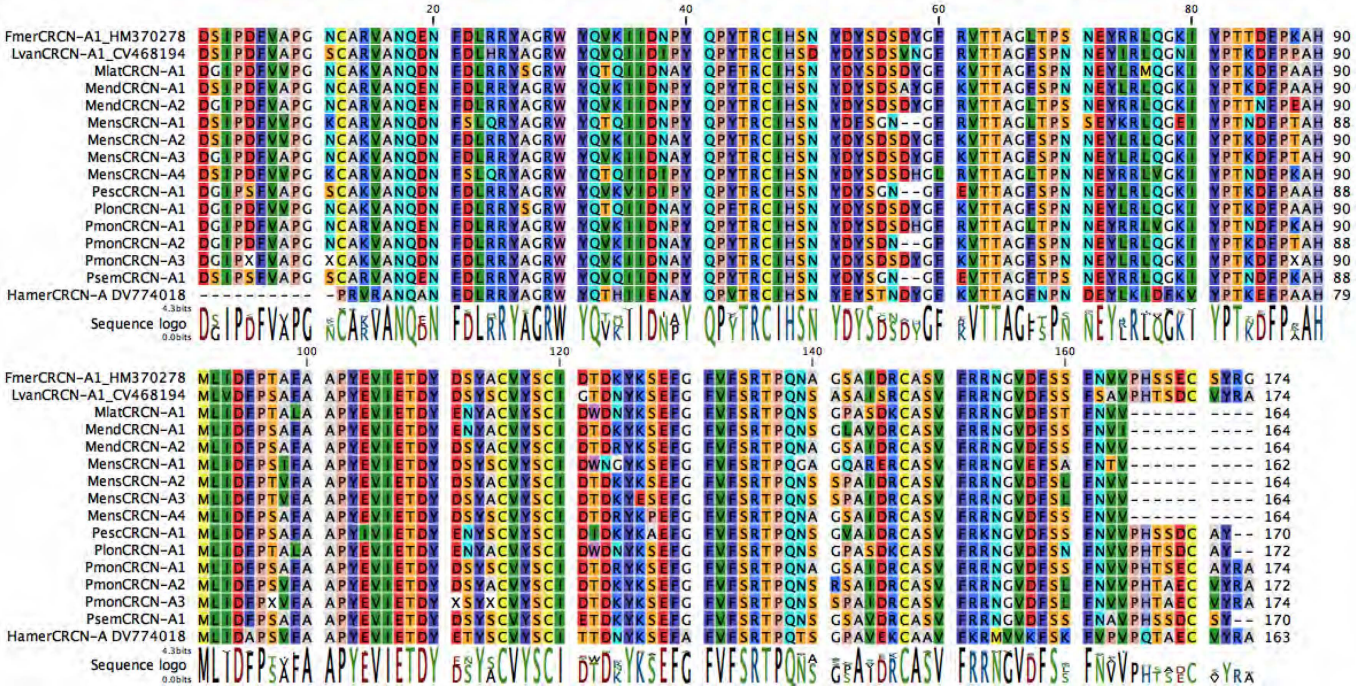



**Fig. S2. Amino acid alignment of identified *P. monodon* CRCN-A isoforms with the clawed lobster CRCN-A sequence.** Only coding sequence is shown, and the known signal peptide at the beginning of the protein has been removed. Gene abbreviations are shown in Table 3, and Genbank accession numbers shown for existing CRCN sequences. Each amino acid is displayed as single letter code and background coloured using rasmol colors. The sequence logo is a visualization tool of conserved regions, with larger sized letters indicative of the level of conservation of that residue across species.

## Supplementary Figure 3

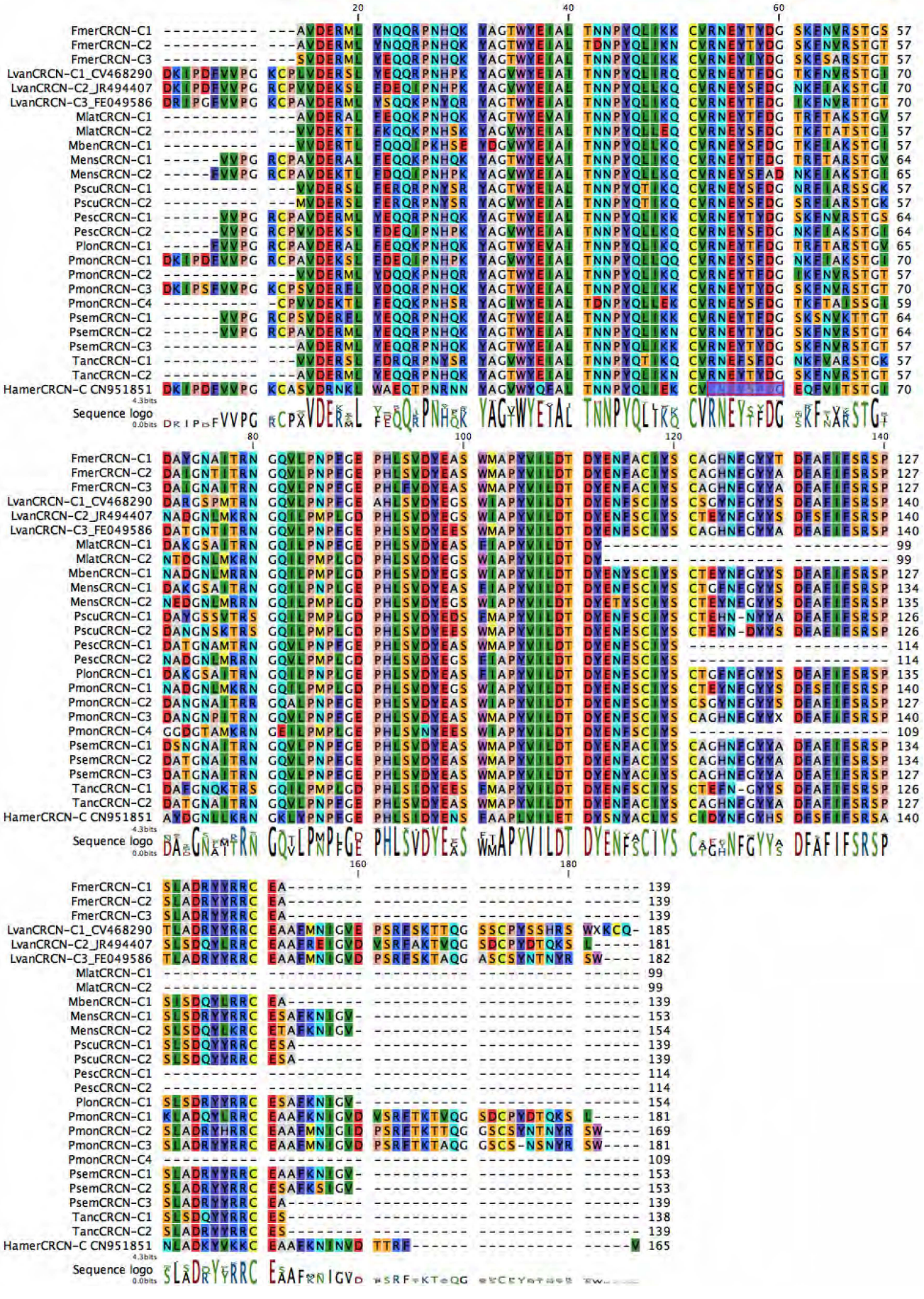


**Fig. S3. Amino acid alignment of identified *P. monodon* CRCN-C isoforms with the clawed lobster CRCN-C sequence.** Only coding sequence is shown, and the known signal peptide at the beginning of the protein has been removed. Gene abbreviations are shown in Table 3, and Genbank accession numbers shown for existing CRCN sequences. Each amino acid is displayed as single letter code and background coloured using rasmol colors. The sequence logo is a visualisation tool of conserved regions, with larger sized letters indicative of the level of conservation of that residue across species.

**Table S1. Average uncooked RGB colour values of prawns injected with different dsRNA constructs and measured over time (days).** Increasing values are shown as darker red, green or blue colours, respectively. Summary statistics from 2-way ANOVA analysis demonstrate the effect of time, injection or the interaction of both. Significant P values are marked in bold for the comparison of values over time (below) or across treatments (right).

Average R Value

| Injection | Luc | CRCN-A | CRCN-C1 | CRCN-C2 |
|---|---|---|---|---|
| Day 0 | 64.8 | 58.3 | 64.7 | 54.8 |
| Day 2 | 53.7 | 85.8 | 79.0 | 97.1 |
| Day 4 | 59.7 | 100.6 | 86.4 | 113.6 |
| Day 7 | 56.2 | 91.4 | 79.4 | 119.1 |
| *P value (Time)* | | | | |
| *0 vs 2* | 0.11 | **0.00** | 0.04 | **0.00** |
| *0 vs 4* | 0.46 | **0.00** | **0.00** | **0.00** |
| *0 vs 7* | 0.21 | **0.00** | 0.04 | **0.00** |

*P value (Injection)*

| *Luc vs CRCN-A* | *Luc vs CRCN-C1* | *Luc vs CRCN-C2* |
|---|---|---|
| 0.35 | 0.99 | 0.15 |
| **0.00** | **0.00** | **0.00** |
| **0.00** | **0.00** | **0.00** |
| **0.00** | **0.00** | **0.00** |

Summary Statistics

| | *F* | *P value* |
|---|---|---|
| Time (T) | 29.7 | **0.000** |
| Injection (I) | 42.6 | **0.000** |
| T x I | 8.6 | **0.000** |

Average G Value

| Injection | Luc | CRCN-A | CRCN-C1 | CRCN-C2 |
|---|---|---|---|---|
| Day 0 | 70.3 | 63.0 | 69.0 | 64.2 |
| Day 2 | 65.9 | 70.1 | 74.4 | 74.6 |
| Day 4 | 70.3 | 76.8 | 78.4 | 87.6 |
| Day 7 | 72.6 | 74.2 | 75.2 | 92.3 |
| *P value (Time)* | | | | |
| *0 vs 2* | 0.33 | **0.11** | 0.22 | **0.02** |
| *0 vs 4* | 0.99 | **0.00** | **0.04** | **0.00** |
| *0 vs 7* | 0.60 | **0.01** | 0.17 | **0.00** |

*P value (Injection)*

| *Luc vs CRCN-A* | *Luc vs CRCN-C1* | *Luc vs CRCN-C2* |
|---|---|---|
| 0.10 | 0.77 | 0.17 |
| 0.34 | 0.06 | 0.05 |
| 0.14 | 0.07 | **0.00** |
| 0.71 | 0.56 | **0.00** |

Summary Statistics

| | *F* | *P value* |
|---|---|---|
| Time (T) | 14.0 | **0.000** |
| Injection (I) | 8.1 | **0.000** |
| T x I | 3.0 | **0.010** |

Average B Value

| Injection | Luc | CRCN-A | CRCN-C1 | CRCN-C2 |
|---|---|---|---|---|
| Day 0 | 55.4 | 48.1 | 53.7 | 51.4 |
| Day 2 | 54.5 | 46.6 | 56.0 | 49.1 |
| Day 4 | 58.5 | 53.0 | 58.6 | 60.8 |
| Day 7 | 63.5 | 52.6 | 57.0 | 65.8 |
| *P value (Time)* | | | | |
| *0 vs 2* | 0.81 | 0.65 | 0.49 | 0.50 |
| *0 vs 4* | 0.36 | 0.15 | 0.15 | **0.01** |
| *0 vs 7* | 0.02 | 0.18 | 0.33 | **0.00** |

*P value (Injection)*

| *Luc vs CRCN-A* | *Luc vs CRCN-C1* | *Luc vs CRCN-C2* |
|---|---|---|
| 0.03 | 0.61 | 0.24 |
| **0.02** | 0.66 | 0.11 |
| 0.11 | 0.97 | 0.48 |
| **0.00** | 0.06 | 0.50 |

Summary Statistics

| | *F* | *P value* |
|---|---|---|
| Time (T) | 11.9 | **0.000** |
| Injection (I) | 9.1 | **0.000** |
| T x I | 1.6 | 0.149 |