Marine Genomics 63 (2022) 100948

Regulation of gene expression during ontogeny of physiological function in the brackishwater amphipod *Gammarus chevreuxi*

C. McAndry, M. Collins, O. Tills, J.I. Spicer, M. Truebano *

Marine Biology and Ecology Research Centre, School of Biological and Marine Sciences, University of Plymouth, Drake Circus, Plymouth PL4 8AA, UK

**A R T I C L E   I N F O**

**Keywords:** Transcriptome, Heart, Crustacean, Comparative developmental physiology

**A B S T R A C T**

Embryonic development is a complex process involving the co-ordinated onset and integration of multiple morphological features and physiological functions. While the molecular basis of morphological development in embryos is relatively well known for traditional model species, the molecular underpinning of the development of physiological functions is not. Here, we used global gene expression profiling to investigate the transcriptional changes associated with the development of morphological and physiological function in the amphipod crustacean *Gammarus chevreuxi*. We compared the transcriptomes at three timepoints during the latter half of development, characterised by different stages of the development of heart form and function: 10 days post fertilisation (dpf, Early: no heart structure visible), 15 dpf (Middle: heart present but not fully functional), and 18 dpf (Late: regular heartbeat). Gene expression profiles differed markedly between developmental stages, likely representing a change in the activity of different processes throughout the latter period of *G. chevreuxi* embryonic development. Differentially expressed genes belonged to one of three distinct clusters based on their expression patterns across development. One of these clusters, which included key genes relating to cardiac contractile machinery and calcium handling, displayed a pattern of sequential up-regulation throughout the developmental period studied. Further analyses of these transcripts could reveal genes that may influence the onset of a regular heartbeat. We also identified morphological and physiological processes that may occur alongside heart development, such as development of digestive caeca and the cuticle. Elucidating the mechanisms underpinning morphological and physiological development of non-model organisms will support improved understanding of conserved mechanisms, addressing the current phylogenetic gap between relatively well known model species.

1. **Introduction**

Development is a dynamic process involving the ontogeny of form (morphology) and function (physiology), and their interaction, leading to a functioning biological system. The onset of function requires the presence of all necessary morphological structures, and the concurrent development and onset of physiological regulatory mechanisms needed to drive the system (Adolph, 1968; Burggren and Warburton, 2005; Spicer and Burggren, 2003). The molecular mechanisms that govern such processes during development are complex and multifaceted. Recent advances in “omic” techniques have fuelled the investigation of such mechanisms in a number of taxa, first in model species such as the mouse (He et al., 2020), zebrafish (Mathavan et al., 2005; Per et al., 2011), sea urchin (Sodergren et al., 2006) and fruit fly, *Drosophila* (Graveley et al., 2011), and progressively more in emerging marine invertebrate models, e.g. the amphipod *Parhyale hawaiensis* (Averof and Patel, 1997; Sun and Patel, 2019) and various molluscs (e.g. Heyland et al., 2011; Samadi and Steiner, 2010). Through this proliferation, it has become apparent that many mechanisms may be conserved across taxa, such as those governing development of form through early cell fate and morphological patterning (Iijima et al., 2006; Jaramillo et al., 2016; McGinnis et al., 1984; Serano et al., 2016) and heart development (Bodmer, 1995; Bodmer and Venkatesh, 1998; Cripps and Olson, 2002; Hami et al., 2011; Lo and Frasch, 2003; Medioni et al., 2009; Reim and Frasch, 2010; Szeto et al., 2002; Vogler and Bodmer, 2015). For example, a key class of transcription factors, the homeobox genes, are involved in the morphological development of many anatomical features across taxa, playing vital roles in segmentation of arthropods, via *Hox* genes (e.g. Serano et al., 2016), as well as in cardiac morphological patterning through the transcription factor *tinman* and its downstream cascade in both arthropods and vertebrates (Bodmer and Frasch, 2010; Vogler and Bodmer, 2015). The development of physiological function,
however, has received comparatively little attention, particularly in the lower vertebrates and invertebrates, despite its importance to organism survival (Burggren, 2004; Spicer et al., 2018; Spicer and Burggren, 2003; Spicer and Gaston, 1999). In particular, studies investigating the molecular mechanisms that underpin the onset of physiological function during development are scarce, especially in non-model species. To date, such studies have focused primarily on processes such as osmoregulation (Truebano et al., 2020), digestion (Dai et al., 2009), the function of oxygen transporting proteins (Durstewitz and Terwilliger, 1997; Havird et al., 2016; Wang et al., 2015) and cardiac function (Tyser et al., 2016; Tyser and Srinivas, 2020).

The development and onset of cardiac function, the physiological system long been considered the first to begin to function across many animal groups (Adolph, 1968; Komuro and Izumo, 1993; Romney and Reiber, 2013), has been investigated in a variety of organisms including commonly studied vertebrate (Komuro and Izumo, 1993) and non-vertebrate (Bodmer and Frasch, 2016; Hu et al., 2009) models, as well as, increasingly, in non-model aquatic vertebrates (Miller et al., 2011), and invertebrates (McMahon et al., 1997; Reiber and Harper, 2001; Spicer, 2002; Spicer, 2001; Spicer, 1994). There are a number of reviews on the development of circulatory function across the invertebrates (McMahon et al., 1997; Mill, 1972). They clearly demonstrate the extraordinary diversity in circulatory systems and pumping mechanisms employed across this wide taxonomic group. In most cases within the Arthropoda, hemolymph is circulated around the body in an “open” circulatory system by a muscular heart. Its morphology can range from complex globular hearts in decapod crustaceans to seemingly simpler tube-like hearts in amphipods and insects (McMahon et al., 1997). The first circulatory systems are believed to be peristaltic tubes driving extracellular fluid into an open vascular system (Bettex et al., 2014). The tubular heart of crustaceans is thought to be plesiomorphic, with early forms possessing a uniform segmented tubular heart running from head to tail (Wilkens, 1999). Interestingly, in some arthropods, e.g. some cirripedes, ostracods and copepods, the heart is absent, and circulation is provided by contraction of cirri and gut muscles (Maynard, 1961; McGaw and Reiber, 2015).

McGaw and Reiber (2015), McMahon et al. (1997), and Spicer (2001) reviewed the patterns of heart rate, stroke volume and cardiac output throughout the development of crustacean hearts. Since then, we have increased our understanding of the mechanisms underlying these patterns using traditional model organisms. For example, calcium handling mechanisms have been shown to play a vital role in the development and maintenance of regular cardiac myocyte contractions in vertebrate (e.g. Koushik et al., 2001; Langenbacher et al., 2005; Tyser et al., 2016; Tyser and Srinivas, 2020) and invertebrate (Sanyal et al., 2006) models. However, analogous studies of non-model invertebrates are lacking. Many mechanisms of early cardiac morphological development appear to be conserved across taxa (Bodmer, 1995; Bodmer and Venkatesh, 1998), despite considerable differences in complexity of cardiac systems. This generalisation remains largely untested, especially in the case of physiological ontogeny, preventing the bridging of the current phylogenetic gap between the common model species, and our understanding of evolution (Serano et al., 2016; Sun and Patel, 2019). Further understanding of the mechanisms governing the phenotype of physiological systems, such as cardiac activity, in a wide variety of developing animals is essential if we are to predict how and when selection pressures may operate on a developing individual (Adolph, 1968; Burggren, 2018; Burggren and Warburton, 2005) and so probe Darwinian fitness (Mueller et al., 2015; Spicer et al., 2018).

Consequently, the aim of this study was to investigate some of the molecular mechanisms underpinning physiological development, concentrating on the development of cardiovascular structure and function in one of the better characterised non-model invertebrate groups, the crustaceans (McMahon, 2001; Spicer, 2022). We achieved this by characterising the transcriptome level changes associated with the onset of cardiac function in the latter half of embryonic development (10–18 days post fertilisation – dpf; average development time ~ 20 days (pers. obs.)) of the direct developing amphipod, Gammarus chevreuxi. During this period (around 15 dpf), the embryo first develops a heart, identifiable visually by the presence of an irregular beat. In G. chevreuxi, the heart is a tubular structure located in the pericardiac sinus, dorsal to the hepato-pancreatic tubes ( Sexton, 1928; Spicer, 2001). This beat becomes regular by 18 dpf. G. chevreuxi is an ideal model to address our aim as its development is relatively well understood ( Sexton, 1928; Truebano et al., 2020), is ecologically-important in brackish waters (Cogne et al., 2019), lab-hardy ( Sexton, 1928), and its transcriptome has recently been sequenced (Collins et al., 2019; Truebano et al., 2016).

2. Materials and methods

2.1. Sample preparation and sequencing

Amphipod collection and maintenance was carried out exactly as detailed in Truebano et al. (2016). The transcriptome of G. chevreuxi was sequenced at three developmental stages in the latter half of embryonic development from 10 to 18 days post fertilisation (dpf); this period is characterised by the development and onset of heart function. A detailed description of sample preparation, sequencing and assembly is presented in Truebano et al. (2016). Briefly, eggs were removed from the female brood pouches of a stock population of G. chevreuxi after 10, 15 and 18 dpf, delineating each of three developmental stages: “Early” – no visible heart structure present; “Middle” – irregular heart beat present; and “Late” – regular heart beat present (Fig. 1a–c). Therefore the development of heart form is visible between Early and Middle stages, and the onset of heart function (i.e. regular heart beat) occurs between Middle and Late stages.

Total RNA was isolated from three replicate pools (n = 170 individuals per pool) for each of the three stages. TruSeq RNA libraries (Illumina, San Diego, USA) were synthesised and sequenced using 100 base paired-end sequencing (HiSeq 2000, Illumina, San Diego, USA) at the NBAF GenePool genomics facility, University of Edinburgh. Sequencing produced 52.6 M paired-end reads.

2.2. Transcriptome assembly and annotation

Reads were assembled using Trinity v 20.140.412 (Grabherr et al., 2011) producing 40,391 transcripts assigned to 26,828 genes (Truebano et al., 2016). The assembled transcriptome was annotated using Trinotate v 3.2.0 (Bryant et al., 2017). BLAST searches (blastx and blastp) were performed against the Uniprot/Swissprot database using an E-value threshold of 1 × 10^-8 (Table A1). Transcripts also received annotations from the Gene Ontology (GO) and Kyoto Encyclopaedia of Genes and Genomes (KEGG) databases. Approximately 23% of total transcripts were putatively annotated (blastx). Previous work on this species achieved high levels of annotation using the transcriptome as a reference (Collins et al., 2017; Truebano et al., 2016). While there is a genome available for the related amphipod Parhyale hawaiensis, the limitations of annotating from a fairly distant relative provide no substantial advantage over the method we implemented.

In addition, we performed BLAST searches (blastx and blastp) against Drosophila melanogaster in the Uniprot/Swissprot database using an E-value threshold of 1 × 10^-8 (Table A2), enabling us to compare our non-model invertebrate to a well-characterised invertebrate model and investigate potential similarities in mechanisms of heart development.

2.3. Gene expression analyses

Transcript expression levels were quantified using Kallisto v 0.44.0 (Bray et al., 2016). Low count genes were removed (FPKM < 1). Counts data were imported into R v 3.6.3 (R Core Team, 2020; Bioconductor Team, 2020) and summarised to gene level using tximport v 1.16.1 (Soneson et al., 2018).
et al., 2016). Patterns in global gene expression across developmental stages were visualised using PCA of variance stabilised counts using the plotPCA function from DESeq2 v 1.26.0 (Love et al., 2014). Genes were correlated ($P < 0.01$) to PC1 and PC2 using dimdesc (in FactoMineR v 2.3; Le et al., 2008).

DESeq2 v 1.26.0 (Love et al., 2014) was used to perform differential gene expression analysis, with developmental stage as a factor, and the three stages as levels (Early, Middle, Late). Log2 fold changes were shrunk using “ashr” to adjust for bias in gene count (Stephens, 2016; Stephens et al., 2020), and pairwise comparisons between developmental stages were carried out to identify changes associated with morphological development (Early vs. Middle) and physiological development (Middle vs. Late). Significantly differentially expressed genes (DEGs) were identified ($P_{adj} < 0.01$, log fold change $>1$ or $<-1$) from pairwise comparisons and utilised in subsequent clustering analyses. Hierarchical clustering was performed on Euclidean distances of variance stabilised counts using the hclust function from stats v 4.0.3. The elbow method (Thorndike, 1953) was used to identify the optimum number of clusters to represent the data. Expression of gene clusters were then visualised using pheatmap v 1.0.12 (Kolde, 2019).

Additionally, we searched the transcriptome for genes that were annotated in the Drosophila-specific BLAST search that have been identified as important in cardiac morphogenesis in Drosophila (Ahmad, 2017; Vogler and Bodmer, 2015; for more detail on search terms used and results, see Appendix B).

### 2.4. Functional enrichment analysis

To identify the broader function of genes correlated to principal components, genes correlated to PC1 and PC2 were assigned GOSlim annotations using GSEABase v 1.52.1 (Morgan et al., 2020). To explore how different patterns of gene expression relate to the onset of vital morphological and physiological function in the organism, GO enrichment analysis was performed on clustered DEGs using goseq v 1.38.0 (Young et al., 2010). GO terms were considered enriched if $P_{adj} < 0.01$. Similarly, KEGG enrichment analysis was carried out on DEGs within each cluster using clusterProfiler v 3.18.0 (Yu et al., 2012) in order to identify significantly enriched pathways ($P_{adj} < 0.01$). Given that the developmental stages chosen are typified by the development of heart form and function, the expression of gene clusters which displayed enrichment of cardiac-related pathways was explored using pheatmap v 1.0.12 (Kolde, 2019).

### 2.5. Data deposition

Raw data are available from the European Nucleotide Archive under the BioProject accession: PRJEB11518.

### 3. Results

#### 3.1. Transcriptomic changes with development

Principal Component Analysis (PCA) of all genes ($n = 26,828$) revealed a separation between developmental stages along PC1 (13,397 correlated genes, $P < 0.01$), accounting for 78% of the total variance (Fig. 2). Genes correlated with PC1 had a broad range of functions including roles in biosynthetic processes and nucleobase-containing compound metabolism (Tables C1–2). There was greater separation between the Early and Middle stages along PC1, compared to the Middle and Late stages. PC2 (1279 correlated genes, $P < 0.01$) accounted for 14% of the total variance, and though displaying similar broad scale annotations to PC1 (Fig. 2, Tables C3–4), this component separated the Middle developmental stage from Early and Late. The separation between the different life stages and large proportion of variance explained with PCA highlights significant differences in global gene expression profiles between the three stages of development.

#### 3.2. Differential gene expression and cluster analysis

A total of 6231 unique differentially expressed genes (DEGs) were identified between the Early and Middle stages of development, of which approximately 61% were up-regulated. Fewer genes ($n = 954$) were uniquely differentially expressed between Middle and Late development, of which approximately 89% were up-regulated. There were 2510 commonly regulated genes (Fig. 3; Appendix D).

DEGs were assigned to three clusters based upon their patterns of expression across development (Fig. 4, Table E1). Cluster A (4,206 DEGs) displayed relatively low expression in the Early stage, which then markedly increased towards the Middle stage and remained high into Late development. There were 142 enriched GO terms ($P_{adj} < 0.01$) in this cluster, representing a broad range of processes including terms relating to: signalling and communication; membrane components and processes; ion binding and transport; actin binding; and chitin and amino acid metabolism (Tables E2–3). There were two KEGG pathways enriched ($P_{adj} < 0.01$) within this cluster: glycolysis/gluconeogenesis, and synaptic vesicle cycle (Tables E4–5).

Cluster B (2,597 DEGs) displayed high expression during Early development, was markedly down-regulated towards the Middle stage.
3.3. Heart-related gene expression

There were nine DEGs from Cluster C corresponding to enriched cardiac-related KEGG pathways, which all displayed up-regulation throughout development (Fig. 5; Table F1). They represent two functional groups, either being involved in the calcium (Ca\(^{2+}\)) ion handling in cardiac muscle cells or are fibres involved in the structural machinery of muscle contraction. Calcium ion handling genes included Ryanodine receptor (RYR2), Sodium/calcium exchanger 2 (NCX), Calcium-transporting ATPase sarcoplasmic/endoplasmic reticulum type (ATP2A a.k.a. SERCA2a), Plasma membrane calcium-transporting ATPase 3 (ATP2B) and Protein kinase C, brain isozyme (PRKCA a.k.a. PKCo). Muscle fibre genes included Titin (TTN), Actin-2 (ACTB_G1), Delta-sarcoglycan (SGCD) and Myosin heavy chain, muscle (MYH6_7). While all genes were up-regulated to some degree throughout the period of development, the former group of genes displayed a slightly different pattern of differential expression to the latter. RyR2, NCX and SERCA2a were only significantly up-regulated between Middle and Late stages (P_{adj} < 0.01). PKCo was only significantly up-regulated between Early and Middle. ATP2B as well as TTN, ACTB_G1, SGCD and MYH6_7 were all significantly up-regulated in both pairwise comparisons.

4. Discussion

The main aim of this study was to investigate some of the molecular mechanisms underpinning physiological development in Gammarus chevreuxi, associated with the ontogeny of cardiac structure and function. Gene expression profiles of the three developmental stages differed markedly, likely representing a change in the activity of different processes throughout the latter period of G. chevreuxi embryonic development. Differentially expressed genes were assigned to three clusters based on their expression patterns across development. Transcripts related to cardiac contractile machinery and calcium handling were sequentially up-regulated throughout the developmental period studied, reaching the highest levels of expression in the Late stage embryos. Some genes putatively identified as being involved in cardiac morphogenesis in the model arthropod Drosophila melanogaster, were also expressed in the embryonic stages of G. chevreuxi examined here. This may suggest universality of some mechanisms across taxa, with amphipod analogues of identified Drosophila genes controlling aspects of early cardiac morphogenesis and processes directing onset of cardiac function during the period studied.

4.1. Gene expression profiles differ between stages in latter half of development

The marked contrast in the gene expression profiles of each developmental stage could represent a change in the activity of different processes throughout the latter period of G. chevreuxi embryonic development. There was a greater difference between the gene
expression profiles in the Early and Middle stages than there were between Middle and Late, potentially indicating that a greater range of biological processes were activated between the Early and Middle stages studied. Moreover, alongside transcriptional regulation, differences in protein production, structure, and function that direct development could be induced by post-transcriptional and translational regulation further downstream. For example, miRNA is known to play an important role in post-transcriptional modification of gene expression in a variety of processes in Drosophila, ranging from tissue growth to endocrinology and the development and activity of the central nervous system (Carthew et al., 2017). Studies in mice have demonstrated its role in cardiac development (reviewed in Braga et al., 2021; Ouyang and Wei, 2021) including regulation of cardiomyocyte proliferation (Liu et al., 2008). In our transcriptome, fewer genes regulated transcriptionally between Middle and Late could therefore represent a greater degree of translational regulation. Alternatively, it is possible that this observed effect is due to the longer time period between the Early and Middle stages (5 days) compared to the Middle and Late stages (3 days). However, the timing of these periods was essential to achieve our aim; comparing these stages allowed us to discern potential transcriptomic evidence to support morphological and physiological development observed during the latter half of G. chevreuxi embryonic development, as well as the development of other similar species.

4.2. Onset of heart function

While there is no direct description of the developmental stages of G. chevreuxi embryos, the embryonic development of the congeneric Gammarus pulex (Mc Cahon and Pascoe, 1988) and the emerging model talitrid amphipod Parhyale hawaiensis (Browne et al., 2005) have been comprehensively described. The relative timing of heart development

**Fig. 4.** (a) Heatmap (VST transformed gene counts) of differentially expressed genes (P<sub>adj</sub> < 0.01, log<sub>2</sub> FC > 1 or < −1) through three stages of development (Early, E1 – E3; Middle, M1 – M3; Late, L1 – L3). Genes clustered into three groups (A, B, C) by means of “complete” hierarchical clustering with Euclidean distance. (b - d) For each cluster, number of genes and mean scaled counts are presented. Black lines represent expression of individual genes and coloured lines represent the mean expression pattern of each cluster. (e - g) The enriched KEGG pathways (P<sub>adj</sub> < 0.01) for each cluster are shown.

**Fig. 5.** Heatmap (VST transformed counts) of differentially expressed Cluster C genes (P<sub>adj</sub> < 0.01, log<sub>2</sub> fold change > 1 or < −1) mapped to cardiac-related enriched KEGG pathways (ko05414, ko05410, ko05412, ko04262), across three stages of development (Early, E1 – E3; Middle, M1 – M3; Late, L1 – L3). Genes clustered by means of complete hierarchical clustering with Euclidean distance.
appears to be highly comparable across these species (Fig. G1), with the heart first becoming visible around three quarters of the way through development (*G. chevreuxi* ~ 75%, *G. pulex* ~ 74%, *P. hawaiensis* ~ 72%; Browne et al., 2005; Mc Cahon and Pascoe, 1988), and the heart beat becoming strong and regular soon after (*G. chevreuxi* ~ 90%, *P. hawaiensis* ~ 86.4%; Browne et al., 2005; Mc Cahon and Pascoe, 1988). Despite considerable differences between habitats and life histories across these amphipod species, the highly similar timing of cardiac development suggests that processes in the latter half of development, and potentially the mechanisms underpinning them, could be similar. There is also extensive understanding of the embryonic development in more phylogenetically distant model invertebrates such as *D. melanogaster* (Hartenstein and Campos-Ortega, 1985), in which the heart is formed in a mid-dorsal position approximately 58% through development (Hartenstein and Campos-Ortega, 1985; Markow et al., 2009). This is relatively earlier in development compared to amphipod species, suggesting that the order of processes in development could be different in these two taxonomic groups (heterochrony), cardiac development could take less time in *D. melanogaster*, or a functioning heart could be necessary earlier in development in the fruit fly (Burggren, 2004).

Changes in the expression of cardiac-related genes were observed in the development of *G. chevreuxi* embryos between 10 dpf and 18 dpf, closely corresponding to the development of heart form (15 dpf) and function (18 dpf) during this period. In order for contraction to occur within heart tissue, two requirements must be met: the formation of the contractile machinery, the sarcomere; and the coordinated movement of Ca$^{2+}$ ions by channels to alter the cytosolic Ca$^{2+}$ concentration. Up-regulation of sarcomeric proteins ACTB, MYH6, 7, TN and SGCDD throughout development represents the increase in production of contractile machinery in muscle cells, including cardiac muscle. Actin and Myosin form the thin and thick filaments respectively in the sarcomere. Delta Sarcoglycan and the giant connecting protein Titin play roles in localisation (Cheng et al., 2006; Shi et al., 2004) and force transmission (Campbell, 2013; Campbell et al., 2008; Chopra et al., 2018) between elements of the sarcomere and sarcolemma. Together, these genes ensure that the machinery needed for contraction is in place. The expression of these genes throughout the period of late development of *G. chevreuxi* provides transcriptomic evidence for the formation and localisation of the sarcomere during cardiac myogenesis during this time.

The second requirement centres on calcium transport; a major driver of muscle contraction, which plays an essential role in regulating cardiac cycles (Cartwright et al., 2005; Ebert et al., 2005; Langenbacher et al., 2005). The calcium transport protein genes NCX, RYR2, SERCA2a, ATP2B and PKCa were up-regulated throughout development. NCX is a reversible transporter involved in both influx and efflux of Ca$^{2+}$ across the sarcolemma, hence it has a role in both depolarisation and repolarisation of cardiac muscle cells (Blausstein and Lederer, 1999; Brini and Carafoli, 2011). RYR2, a cardiac-specific ryanodine receptor, controls the large release of Ca$^{2+}$ ions from the sarcoplasmic reticulum during depolarisation (Stokes and Wagenknecht, 2000). ATP2B (a type of PMCA – plasma membrane calcium pump) and SERCA2a are both involved in the restoration of basal Ca$^{2+}$ levels after, through extrusion of calcium ions out of the cell or sequestration into the sarcoplasmic reticulum respectively (Altshuler et al., 2012; Brini and Carafoli, 2011; Campbell et al., 2008; Ebert et al., 2005; Langenbacher et al., 2005; MacLennan, 1970; Stokes and Wagenknecht, 2000). PKCa has demonstrated a role in the activity of the SERCA2a pump in mice, via the dephosphorylation of a pump inhibitory protein, hence also playing a role in the regulation of contractility (Braz et al., 2004). The up-regulation of these calcium transport protein genes is indicative of increasing control over cardiac muscle contraction throughout the latter half of development of *G. chevreuxi*, and suggests universality of some of the mechanisms of cardiac function. The increase in expression of these calcium handling genes towards the Late stage of development in *G. chevreuxi* is concomitant with the period when the heart begins to function with more regularity and rhythmicity, suggesting this regulation may play a key role in creating and maintaining a regular heartbeat between 15 dpf and 18 dpf.

The temporal regulation of calcium handling gene expression could elucidate the sequence of key cardiac development processes during the latter half of *G. chevreuxi* development. While there was a broad pattern of up-regulation in cardiac genes throughout the developmental stages studied, RYR2, NCX and SERCA2a were not significantly up-regulated from Early to Middle stages, while ACTB, MYH6, 7, TN, SGCDD and ATP2B were all significantly up-regulated across both pairwise comparisons. PKCa was only significantly up-regulated in the Early to Middle stages. Before any heart contraction can occur, the proteins required for contractile machinery must develop; this was shown to occur in a region and time-dependent manner in mice (Tyser et al., 2016; Tyser and Srinivas, 2020). As cardiac development continues, the expression of contractile proteins continues to increase, creating more sarcomere band structures within the tissue and increasing contraction. The earlier up-regulation of contractile machinery compared to calcium transport machinery from the Early to Middle stages in our data could suggest that sarcomere formation plays a greater role in the formation of heart morphology that occurs in this time frame, whereas the development of heart function requires greater control of calcium handling. In addition, the sarcoplasmic reticulum plays an increasingly important role in calcium handling as heart development continues, but seems to only have limited involvement in the early stages of onset of cardiac function in mice (Rapila et al., 2008; Sasse et al., 2007; Tyser and Srinivas, 2020). Future targeted study of cardiac gene expression in other crustaceans, on a finer temporal scale and using specifically cardiac tissue, would enable us to separate the interaction of these two aspects of heart development: form and function.

Additionally, as transcriptional regulation of genes may not necessarily map directly onto the final protein product and function, further study of downstream translational regulation and its interaction with transcription would enhance our ability to infer the role of specific gene products in morphological and physiological cardiac development.

### 4.3. Other processes

Prior to the onset of a regular heart beat in *P. hawaiensis*, the extension of the digestive caeca occurs (Stage 23–25, ~57.6–72% through development). This extension is characterised by a series of peristaltic muscular contractions from rudiments projecting from the midgut (Browne et al., 2005). Similar peristalsis in the midgut has been observed to occur alongside a heart beat in *G. pulex* (17–22 dpf at 11°C, ~74–96% through development, Mc Cahon and Pascoe, 1988) and mytid shrimp *Metapenaeus ensis* (McMahon et al., 1997). Also known as the hepatopancreas, digestive caeca are major secretory organs that consist of a number of blind-ended tubes and are thought to play a part in nutrient adsorption and storage in crustaceans, including amphipods (Browne et al., 2005; Schmitz, 1992; Strus et al., 2019). In our *G. chevreuxi* transcriptome, pathways associated with pancreatic secretion are enriched within genes up-regulated throughout the latter half of development. Despite the absence of detailed staging for congeneric amphipods, this could represent molecular evidence for the process of digestive caeca extension alongside development of cardiac function in *G. chevreuxi*.

Cuticle production is an important process in arthropod development, and it occurs over a similar timeframe to heart development in *P. hawaiensis* (Stage 26, ~72% through development, Browne et al., 2005) and *D. melanogaster* (Stage 16, ~58.4% through development, Hartenstein and Campos-Ortega, 1985). In *G. chevreuxi*, 60 terms relating to cuticle production, including cuticle components, chitin metabolism processes and the extracellular matrix were enriched in genes upregulated as the heart developed in form and function. This could provide transcriptomic evidence of increased cuticle secretion throughout the latter half of *G. chevreuxi* embryo development, in a
similar time frame to the onset of heart function.

Interestingly our transcriptome contains some of the key genes identified as having important functions in cardiac morphogenesis in Drosophila, but few of these are differentially expressed between the late developmental stages analysed here (see Appendix B). Transcription factors such as tinman, identified as key components of cardiac development in Drosophila (Bodmer, 1993), as well as several vertebrates (e.g. mice - Lints et al., 1993) and invertebrates (e.g. the polychaete worm Platynereis dumerilii, Saudemont et al., 2008) are not present in our transcriptome. This paucity could reflect the timing of our sampling; the time points employed in this study were not targeted to capture the earliest cardiac cell differentiation, in which these pathways play the most vital role. While we demonstrate that some conservation of the mechanisms of cardiac morphogenesis identified in model species could play a role here, we note that there is still a considerable knowledge gap surrounding the mechanisms governing later physiological development and onset of function.

4.4. Conclusions

We have compared the gene expression profiles of three stages in the late embryonic development of G. chevreuxi, typified by the development of cardiac form and function. The three stages differed in their gene expression profiles, representing changes in activity in developmental processes during late embryonic development. Differentially expressed genes clustered into three expression patterns, with genes relating to cardiac contractile machinery being up-regulated broadly before the up-regulation of calcium handling genes, suggesting that this sequence is important in the early establishment of cardiac function in G. chevreuxi. We also identified expression changes supporting the occurrence of key morphological processes including pancreatic secretions and cuticle synthesis during the latter half of development. Additionally, we identified some key genes involved in cardiac development across taxa, though others were not present in our transcriptome. This may indicate an overrepresentation of studies focussing on early development, leading to relatively fewer genes with roles in later development being identified across taxa. Alternatively, it may suggest a lack of universality across key cardiac-related genes identified in model species and those expressed in more distant non-models. Future work should aim to improve understanding of the mechanisms governing developmental processes, especially regarding physiological ontogeny, across taxa and life histories to bridge the phylogenetic gaps in our understanding of the mechanisms governing later physiological development and onset of function.

Supplementary data to this article can be found online at https://doi.org/10.1101/j.margen.2020.100948.

CRediT authorship contribution statement

C. McAndry: Software, Formal analysis, Visualization, Data curation, Writing – review & editing. M. Collins: Methodology, Software, Formal analysis, Visualization, Writing – original draft, Supervision. O. Tills: Software, Data curation, Writing – review & editing. J.I. Spicer: Conceptualization, Methodology, Writing – review & editing, Funding acquisition. M. Truebano: Conceptualization, Methodology, Investigation, Resources, Writing – review & editing, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was funded by a NERC NBAF Pilot Scheme Grant (NBAF622) to JIS and MT. RNA sequencing and transcriptome assembly was carried out at the NBBF GenePool genomics facility, University of Edinburgh.

References


