Wiley
Aquaculture Research
Volume 2024, Article ID 4118218, 12 pages
https://doi.org/10.1155/2024/4118218

WILEY

Research Article

Profiling Circular RNA Expression Across 11 Tissues in the Marine Gastropod *Rapana venosa*

Mingwei Hong and Zixia Huang

School of Biology and Environmental Science, University College Dublin, Belfield, Dublin 4, Dublin, Ireland

Correspondence should be addressed to Zixia Huang; zixia.huang@ucd.ie

Received 22 June 2024; Accepted 3 September 2024

Academic Editor: Christyn Bailey

The carnivorous marine gastropod, *Rapana venosa*, also known as the veined rapa whelk, has recently gained considerable attention due to its dual significance as both a valuable commercial seafood resource and a biological invader with widespread ecological implications. Significant efforts have been made to investigate the molecular pathways governing the specific adaptations in *R. venosa*. However, the regulatory mechanisms of these pathways remain largely unclear. Circular RNA (circRNA), a newly-recognized class of non-coding molecules, plays crucial roles in post-transcriptional regulation, impacting various fundamental bioprocesses including development, cell cycle, immunity and disease pathogenesis. In this study, we conducted a comprehensive genome wide profiling of circRNAs across 11 tissues of *R. venosa*. We identified a total of 1214 circRNA genes across tissues, with 640 regarded as high-confidence candidates. circRNAs displayed overall low expression levels, diverse isoform types and dynamic expression patterns across various tissues. Our comparative analyses revealed a few circRNA genes with a great diversity of isoforms and abundant expression in the salivary gland, suggesting potential roles in the feeding process of *R. venosa*. Furthermore, Gene Ontology (GO) enrichment analysis indicated the potential involvement of commonly expressed circRNAs in fundamental cellular processes, including cell division, amide biosynthesis and cellular response to hormone stimuli. Our study describes a preliminary examination of circRNAs across various tissues of *R. venosa*, offering a foundation for future research into the molecular mechanisms that influence the biology, ecology and behaviour of marine gastropods.

Keywords: circRNA profiling; gene regulation; marine gastropod; *Rapana venosa*; RNA-seq

1. Introduction

The carnivorous *Rapana venosa*, commonly known as veined rapa whelk, is a species of large predatory marine gastropods within the family Muricidae [1, 2]. Native to the coastal seas of China, Japan and Korea, this molluscan species was introduced into many other regions worldwide, including the Black Sea, through ship transportation [3, 4]. The most prominent characteristics of *R. venosa* include its distinctive spiral shell adorned with ridges and its substantial size, which distinguishes it as one of the largest marine gastropod species. *R. venosa* has garnered considerable attention because of its ecological impact on aquatic ecosystems [5], coupled with its commercial significance as a valuable seafood resource in fisheries and aquaculture [6, 7]. In particular, the veined rapa whelk is regarded as a notorious biological invader due to its remarkable ecological adaptability, characterised by high fecundity, rapid growth rate, early sexual maturation and wide tolerance to various environmental stressors [8–11]. Their voracious predatory behaviour, which includes feeding on a variety of benthic organisms, such as bivalves, other gastropods and crustaceans, can result in significant alterations to the composition of native species and disruptions in food webs in local aquatic ecosystems [5, 12]. Hence, understanding the biology of *R. venosa* is crucial for formulating effective management strategies to mitigate its impact on fisheries and prevent ecological invasions within local biosystems.

Owing to its ecological invasiveness and commercial significance, substantial research efforts have been dedicated to understanding various aspects of *R. venosa*, including its

developmental biology [13–15], population dynamics [3, 16], ecological impact [5, 17] and aquaculture [18, 19] in the past few decades. The advent of next-generation sequencing technology has profoundly propelled the research into the genomic and genetic basis underlying the biology, ecology and behaviour of *R. venosa*. For instance, Song et al. [20] have recently published the chromosome-level genome assembly of *R. venosa*, providing valuable resources for studying its genomic architecture and evolutionary trajectory. Several studies have catalogued the transcriptomes and proteomes of *R. venosa* and explored the molecular changes associated with larval development and growth, particularly focusing on the biphasic transitions like settlement and metamorphosis [21–24]. Moreover, recent investigations into metagenomics and metabolomics [25, 26], alongside the examination of microRNA–mRNA gene networks in *R. venosa* [27], have significantly contributed to our understanding of its developmental processes, adaptive responses to environmental stressors and physiological mechanisms. Despite advancements in our understanding, the regulatory basis underlying these fundamental processes and biological pathways in the veined rapa whelk remains largely elusive.

Circular RNA (circRNA) represents a unique class of single-stranded non-coding RNA molecules that have recently attracted large attention for their vital roles in post-transcriptional regulation. Unlike their linear parental mRNA, circRNAs are distinguished by their covalently closed loop structures formed through an event known as backsplicing, where a downstream splice donor site joins with an upstream splice acceptor site [28]. The circularisation process grants circRNAs greater resistance to RNA exonuclease degradation when compared to their linear counterparts, resulting in their enhanced stability. Initially misidentified as transcriptional errors, circRNAs have since been recognised for their diverse regulatory functions by acting as microRNA sponges, interacting with RNA-binding proteins and being translated themselves. Numerous studies have elucidated the regulatory roles of circRNAs in gene expression and their involvement in many fundamental pathways such as cell proliferation, cellular homeostasis, immunity and disease pathogenesis [29, 30]. These discoveries of circRNAs highlight an additional dimension of complexity in the regulation of biological systems. While circRNAs have been relatively well characterised in a few model organisms, the vast majority of species remain largely unexplored in circRNA research [31, 32]. Within marine invertebrates, genome wide circRNA profiling has been conducted only in a limited number of species. For instance, a previous study performed circRNA profiling and investigated their regulatory roles in determining the body colouration of leopard corals [33]. Additionally, Farhadi et al. [34] constructed a circRNA–microRNA–mRNA network associated with amino acid, carbohydrate and glycogen metabolism in the mud crab *Scylla serrata*, while Ibrahim et al. [35] predicted circRNAs and explored their functional roles underlying the physiological response to salinity stress in the oyster *Crassostrea hongkongensis*. However, circRNAs in *R. venosa*, and marine gastropods more broadly, have not been characterised, leaving their regulatory roles underlying the distinctive adaptations of the veined rapa whelk largely unknown.

In this study, we conducted a comprehensive profiling of circRNA expression across 11 tissues of the marine gastropod *R. venosa*. We investigated circRNA isoform diversity and expression levels across both biological replicates and tissues. Through genome wide comparative analyses, we identified circRNAs that are commonly expressed across all tissues, as well as those that are tissue-specific. Additionally, we discussed the functional significance of these circRNAs in relation to *R. venosa*. This study provides an initial exploration of circRNAs across multiple tissues of *R. venosa*, offering a basis for future research into the molecular mechanisms underlying the biology, ecology and behaviour of this species.

2. Methods

2.1. RNA-Seq Samples of R. venosa. Thirty-three RNA-seq samples of *R. venosa* were obtained from the Sequence Read Archive (SRA) in the National Centre for Biotechnology Information (NCBI) database under the accession number PRJNA855327. These samples, consisting of 100 bp, paired-end reads, were sequenced from 11 tissues, with each three biological replicates. The tissues encompassed in this study are as follows: accessory salivary gland (ASG), foot (Fo), gill (Bi), intestine (In), liver (Li), mantle (Ma), Osphradium (Os), salivary gland (Si), siphon (Sip), stomach (St) and tentacle (Te). Raw RNA-seq data were initially subject to quality assessment. Sequencing adaptors were examined and regions of low quality (defined as less than Q20) were trimmed using Cutadapt (v4.4) [36]. Following these criteria, all reads were of high quality and retained for subsequent circRNA prediction. Further information regarding the 33 RNA-seq samples is provided in Supporting Information 1: Table S1.

2.2. circRNA Prediction. The prediction of circRNAs was carried out using the CIRI2 pipeline (v1.1.0) [37], which identifies RNA-seq reads transcribed from back-spliced junctions (BSJ) of a gene, a hallmark of circRNA presence. The genome and annotation files of *R. venosa* (GCA_028751875.1) [20] were obtained to run CIRI2. Given that the protein-coding sequences were predicted based solely on gene models, we compared them against the UniprotKB database using BLASTx (v2.15.0) [38], with an E value cut-off of 10^{-5}. Sequences with the best hit in the database (E value $< 10^{-5}$) were subsequently annotated with their corresponding gene names.

To run CIRI2, the genome of *R. venosa* was indexed using the Burrows–Wheeler Aligner (BWA, v0.7.17-r1188) [39]. Subsequently, RNA-seq reads from each sample were individually mapped to the genome using BWA with the default parameter settings. With the genome annotation, CIRI2 predicted circRNAs from the mapping files (in SAM format) with the default parameters. The output from CIRI2 comprised a text file containing the genomic coordinates of predicted circRNAs, the read count at BSJs supporting these predictions, and the read count of their respective parental linear genes.

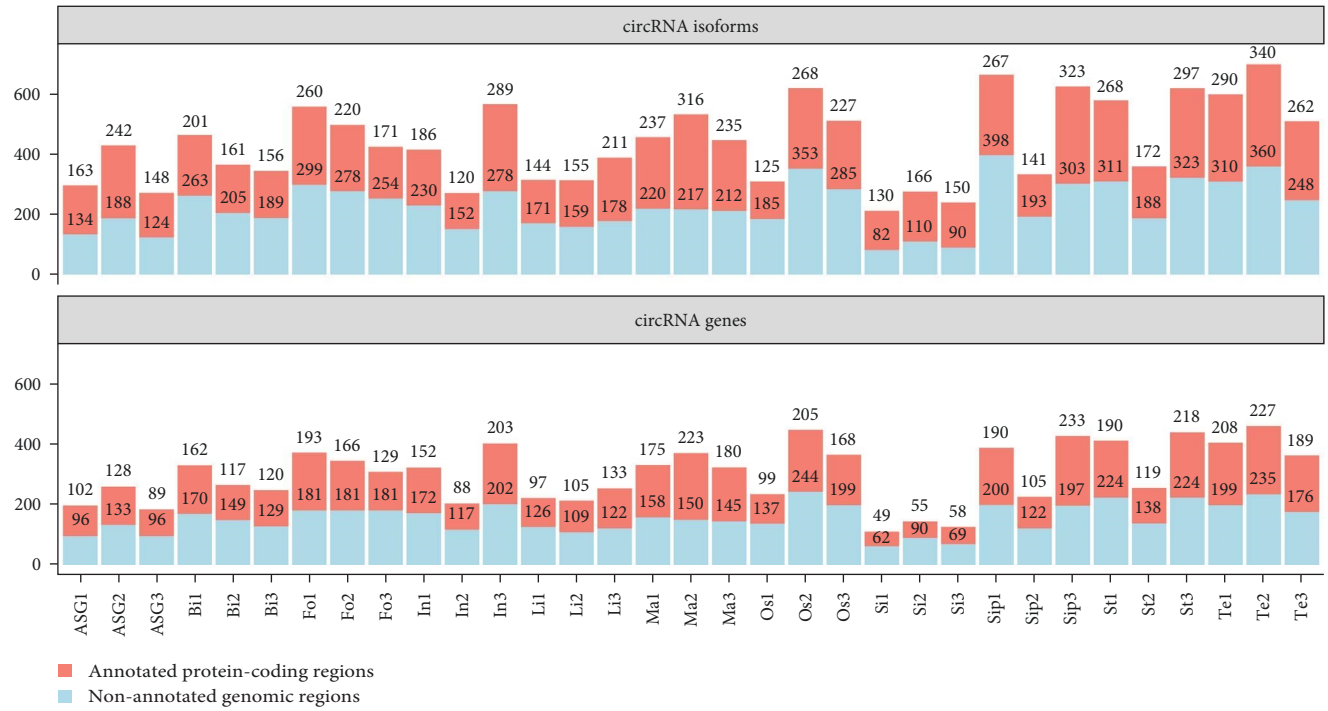


FIGURE 1: The number of circRNAs predicted in 33 samples. The upper panel illustrates the number of identified circRNAs at the isoform level, while the lower panel depicts the count of identified circRNA genes from which one or more isoforms originated. These circRNA genes/isoforms were classified into two groups based on their genomic origins: annotated protein-coding regions or non-annotated genomic regions.

Given that CIRI2 outputs circRNA predictions at the isoform level, to facilitate comparisons across samples and tissues, we obtained gene-level predictions by recognising the genomic coordinates of circRNA isoforms originating from the same parental gene using Bedtools (v2.30.0) [40]. Therefore, a circRNA gene is defined as the specific genomic locus or region from which one or more circRNA isoforms are transcribed through back-splicing events. To ensure the reliability of our data analyses, we initially excluded circRNA genes derived from non-annotated genomic regions. Subsequently, we retained only those circRNA genes identified in at least two out of 33 samples for further downstream analyses. This filtering strategy resulted in the identification of 640 high-confidence circRNA genes across 11 tissues.

2.3. Analyses of circRNAs Across Tissues. We employed the count of reads mapped to the BSJ as a measure for circRNA expression. To aggregate circRNA isoform expression at the gene level, we summed the expression of isoforms derived from the same circRNA gene using the R package *tximport* (v3.18) [41]. This allows for a comprehensive assessment of expression levels for the genomic loci (circRNA genes) that transcribe multiple isoforms, enabling comparisons of circRNA expression patterns across samples and tissues despite the inherent heterogeneity of isoform diversity. To gain an overview of circRNA gene expression across tissues, principal component analysis (PCA) was conducted using the R package *prcomp*. Prior to PCA, circRNA gene raw counts were normalised by the trimmed mean of *M* values (TMM) method, followed by a $\log_2$ transformation ($\log_2$ (TMM + 1)). In addition, to compare the expression patterns between circRNA genes and their linear counterparts, PCA was also performed on the expression levels of 640 corresponding linear parental genes, using the same method as described above. Differential circRNA gene expression analyses were conducted between each pair of tissues (55 pairs) using the R package *edgeR* (v3.36.0) [42]. Next, circRNA genes shared across multiple tissues as well as tissue-specific circRNA genes were identified using the R package *UpsetR* (v1.4.0) [43]. Gene Ontology (GO) enrichment analysis was performed on circRNA genes expressed in at least nine out of 11 tissues using Metascape (v3.5) [44], with the parental genes of 640 circRNAs as the background.

2.4. Statistical Analyses. The statistical analyses employed in this study, including Spearman's correlation test and PCA, were performed in R (v4.1.1) [45]. *P* values were corrected by multiple tests using FDR where applicable. Statistical tests with corrected $P < 0.05$ were considered significant.

3. Results

3.1. Identification of circRNA Genes in 33 RNA-Seq Samples. In this study, we profiled circRNAs expressed in 33 RNA-seq samples from 11 tissues of *R. venosa*, with three biological replicates per tissue (Supporting Information 1: Table S1). Altogether, 7.17 billion paired-ends reads were sequenced, averaging 217.2 ± 71.8 million reads per sample. On average, our prediction yielded 438 ± 132 circRNAs across 33 samples, corresponding to 303 ± 98 unique circRNA genes (Figure 1). Notably, a significant positive correlation was observed between

sample sequencing depth and the number of predicted circRNA ($r = 0.757$; $P = 3.35 \times 10^{-7}$; Spearman's correlation test). The Te exhibited the highest number of predicted circRNAs, contrasted by the lowest number observed in the Si (Figure 1). This discrepancy may be attributed to the disparate sequencing depths of these two tissues (Supporting Information 1: Table S1). Of all circRNAs predicted across samples, 49.22% were detected within annotated protein-coding gene loci, while 50.78% were found in non-annotated genomic regions. Since the functional roles of circRNAs originating from non-annotated genomic regions remain uncharacterised, we opted to exclude them from subsequent analyses.

3.2. circRNA Isoform Diversity Across Samples and Tissues. Overall, we detected 1214 unique circRNA genes originating from annotated protein-coding regions across all 33 RNA-Seq samples, with a considerable proportion expressed solely in a single sample. To ensure the robustness of our data analysis, we retained only those circRNA genes expressed in at least two out of 33 samples. As a result, 640 circRNA genes of high confidence were kept for downstream analyses, of which, only 429 (67.0%) showed homology to the gene entries in the UniprotKB database according to our annotation (See Methods). Typically, parental genes with a higher number of exons have the potential to generate a greater variety of circRNA isoforms; however, we found no significant correlation between the number of circRNA isoforms and the exon count of their parental genes (Figure 2; $p > 0.05$; Spearman's correlation test). The vast majority of circRNA genes yielded single circRNA isoforms across samples, with an average percentage of 78.4% (Supporting Information 4: Figure S1). On average, each circRNA gene produced 1.58 ± 0.48 circRNA isoforms across samples. We noted that the Si has the highest number of circRNA isoforms per gene (Si1: 2.76, Si2: 3.3 and Si3: 2.86), despite having the fewest circRNA genes identified in this tissue. By analysing the top circRNA genes with high isoform diversity (≥5 isoform types), we found that each biological sample displayed a distinct list of circRNA genes within the tissue. Only a small number of these circRNAs genes consistently generated five or more isoform types across all three biological replicates, with the exception of the Si (Supporting Information 4: Figure S2). The Fo, Os and Sip had no circRNA genes consistently having five or more isoform types across all biological replicates, while the Si showed six out of seven circRNA genes meeting this criterion (Supporting Information 4: Figure S2).

3.3. Overview of circRNA Gene Expression in 11 Tissues. To facilitate comparisons across samples and tissues, we analysed circRNA expression at the gene level (see Methods, Supporting Information 2: Table S2). Across all tissues, an average of $18.1\% \pm 2.9\%$ of circRNA genes were consistently identified in all three biological replicates, with a presence of many circRNA genes expressed uniquely in each sample within the tissue (Supporting Information 4: Figure S3). Based on the expression levels of 640 circRNA genes, PCA revealed two main clusters: one comprising the liver, intestine and stomach and the other consisting of the remaining eight tissues grouped together (Figure 3A). Interestingly, the PCA of their linear counterparts displayed a similar pattern (Figure 3B, Supporting Information 2: Table S2). In each tissue, the majority of circRNA genes exhibited low expression levels (Figure 4A), with an average of $79.9\% \pm 6.9\%$ of circRNA genes showing normalised expression values above one across samples (see Methods). In addition, pairwise differential circRNA expression analyses revealed that only a few circRNA genes showed significant differentially expressed between tissues (false positive rate (FDR) < 0.05), except the comparisons between the Si and the other 10 tissues (Figure 4B). This result may stem from the relatively low number of identified circRNAs in the Si (Figure 1).

3.4. Common and Distinctive circRNA Gene Expression Across Tissues. Next, we identified circRNA gene expression shared among tissues and those unique to each tissue. Sixteen circRNA genes were found to be expressed in all 11 tissues, while a relatively small number (1–11) was identified as tissue-specific circRNA genes, varying according to tissue types (Figure 5; Supporting Information 3: Table S3). The expression levels and isoform diversity of 16 common circRNA genes were observed to vary across samples (Figures 6A,B). Remarkably, half of these circRNAs originated from parental genes whose functions remain elusive. Next, we performed a GO enrichment analysis using the parental genes of circRNAs ($n = 36$) expressed in at least nine out of 11 tissues. The significant enriched GO terms (FDR < 0.05) include cysteine-type endopeptidase activity (GO:0004197), cell division (GO:0051301), amide biosynthesis process (GO:0043604), regulation of Wnt signalling pathway (GO:0030111), ATP-dependent activity (GO:0140157), cell adhesion molecular binding (GO:0050839) and cellular response to hormone stimuli (GO:0032870), among others (Figure 6C). Both the Ma and Si displayed the highest number of tissue-specific circRNA genes, with 11 identified, whereas only one was found in the Bi and Sip, respectively (Supporting Information 3: Table S3).

4. Discussion

circRNAs have emerged as a novel and intriguing class of RNA species with diverse functions and potential implications in many cellular processes and diseases. Despite their importance, circRNAs remain poorly characterised among marine invertebrates, leaving their roles largely elusive. In this study, we described a landscape of circRNA expression patterns across 11 tissues of the marine gastropod *R. venosa*, providing a basis for further exploration of their functional roles in marine invertebrates.

In total, we detected 1214 circRNA genes derived from annotated protein-coding genes across 11 tissues of *R. venosa*, with 640 of these identified as high-confidence candidates. The number (1214) is comparable to that reported in the sea cucumber *Apostichopus japonicus* (1087) [46]. However, one study utilising circRNA-seq identified a significant greater number of circRNAs in the oyster *C. hongkongensis* [35]. This indicates that this new technology offers superior circRNA enrichment for sequencing, likely leading to underestimation of circRNA numbers with conventional RNA-seq

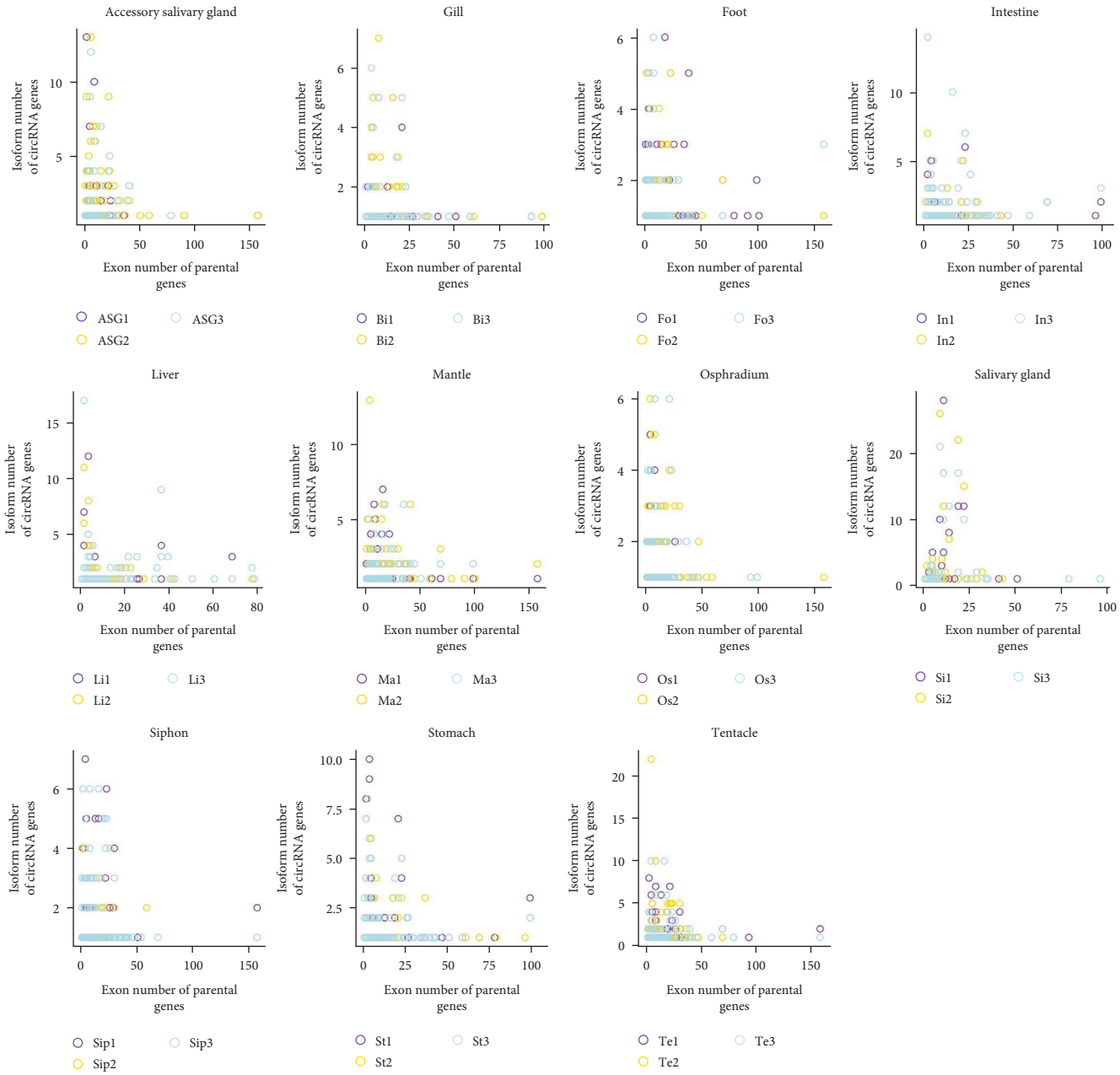


FIGURE 2: The correlation between the number of circRNA isoforms and the exon number of their parental protein-coding genes. Only the expressed circRNA genes out of the 640 identified were included in the plot for each sample. Spearman's correlation coefficient was calculated for each sample.

approaches. We identified a varying number of circRNAs across samples, a trend that exhibits a significant positive correlation with sample sequencing depth ($P<0.05$). This emphasises the requirement of in-depth RNA-seq for capturing a comprehensive array of circRNAs. In addition to circRNAs derived from annotated protein-coding genes, we noticed that approximately half of the circRNA genes were located within non-annotated genomic regions, primarily within intergenic regions (Figure 1). This observation aligns with previous findings of intergenic circRNAs [47], which have been extensively documented across various taxa including marine invertebrates [33], fish [48], birds and mammals [32]. However, we cannot rule out the possibility that some of these circRNAs originated from non-annotated protein-coding genes in the genome or transcriptional noises.

We noticed a diversity of circRNA isoforms expressed from the same genes among biological replicates within most of tissues we investigated. Surprisingly, we also found no significant correlation between the number of circRNA isoforms and the exon count of parental gene (Figure 2). These phenomena indicate the complex and dynamic nature of circRNA biogenesis and suggest that factors beyond the genomic architecture of parental genes likely play essential roles in shaping circRNA diversity within biological systems.

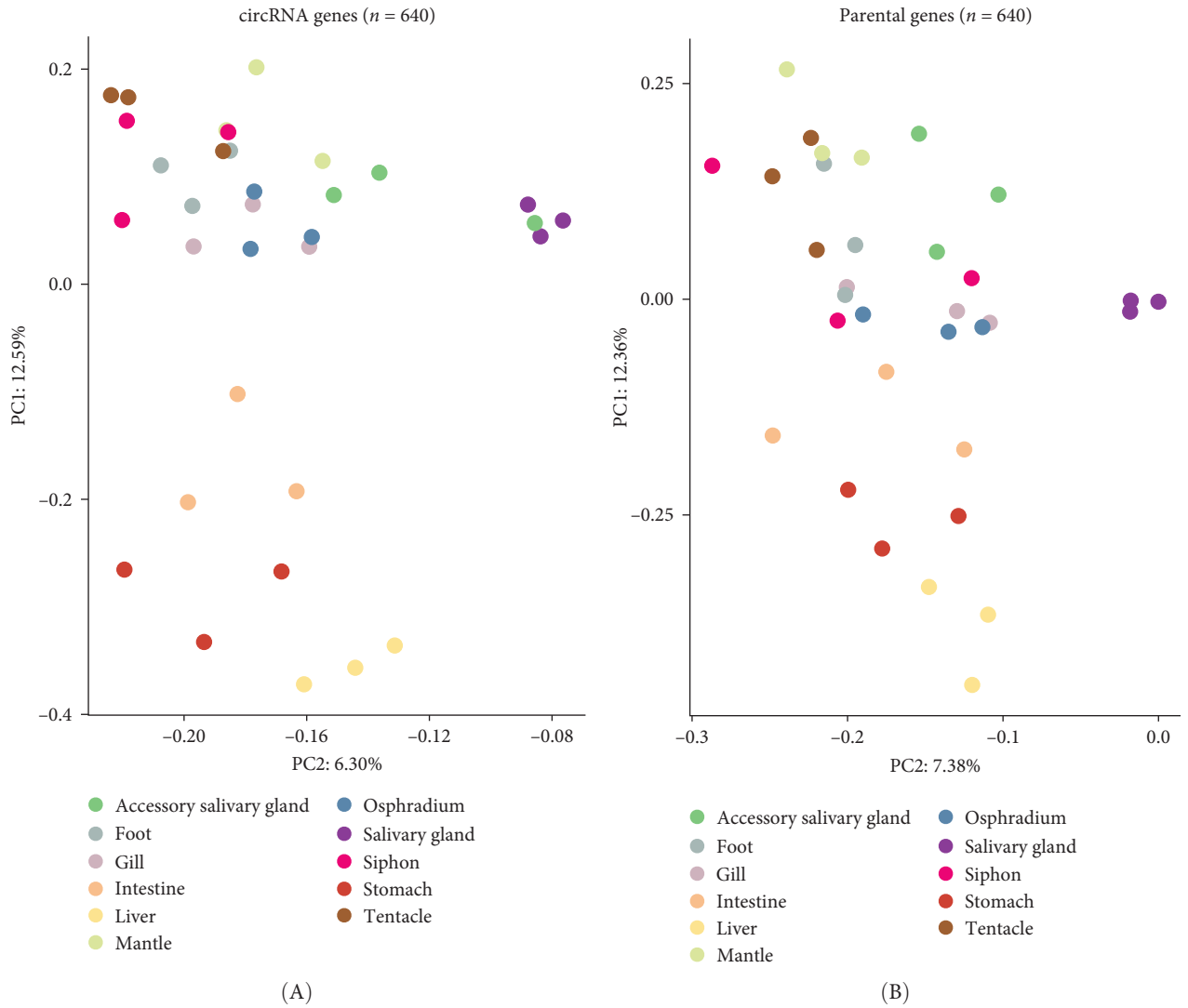
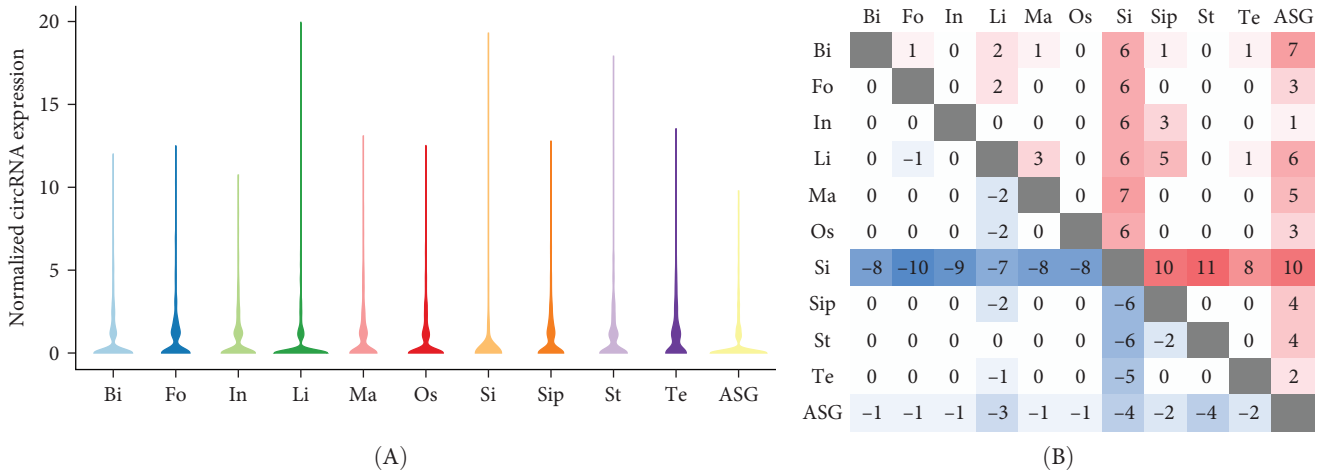


FIGURE 3: The principal component analysis based on the expression levels of (A) 640 circRNA genes and (B) their corresponding parental linear genes, respectively.



	Bi	Fo	In	Li	Ma	Os	Si	Sip	St	Te	ASG
Bi		1	0	2	1	0	6	1	0	1	7
Fo	0		0	2	0	0	6	0	0	0	3
In	0	0		0	0	0	6	3	0	0	1
Li	0	−1	0		3	0	6	5	0	1	6
Ma	0	0	0	−2		0	7	0	0	0	5
Os	0	0	0	−2	0		6	0	0	0	3
Si	−8	−10	−9	−7	−8	−8		10	11	8	10
Sip	0	0	0	−2	0	0	−6		0	0	4
St	0	0	0	0	0	0	−6	−2		0	4
Te	0	0	0	−1	0	0	−5	0	0		2
ASG	−1	−1	−1	−3	−1	−1	−4	−2	−4	−2	

FIGURE 4: The expression levels of 640 circRNA genes across 11 tissues. (A) The distribution of normalised expression levels of these circRNA genes across the tissues. (B) The number of differentially expressed circRNA genes between all possible pairs of tissues, with the upregulated circRNA genes indicated in the upper-right triangle and the downregulated circRNA genes in the lower-left triangle.

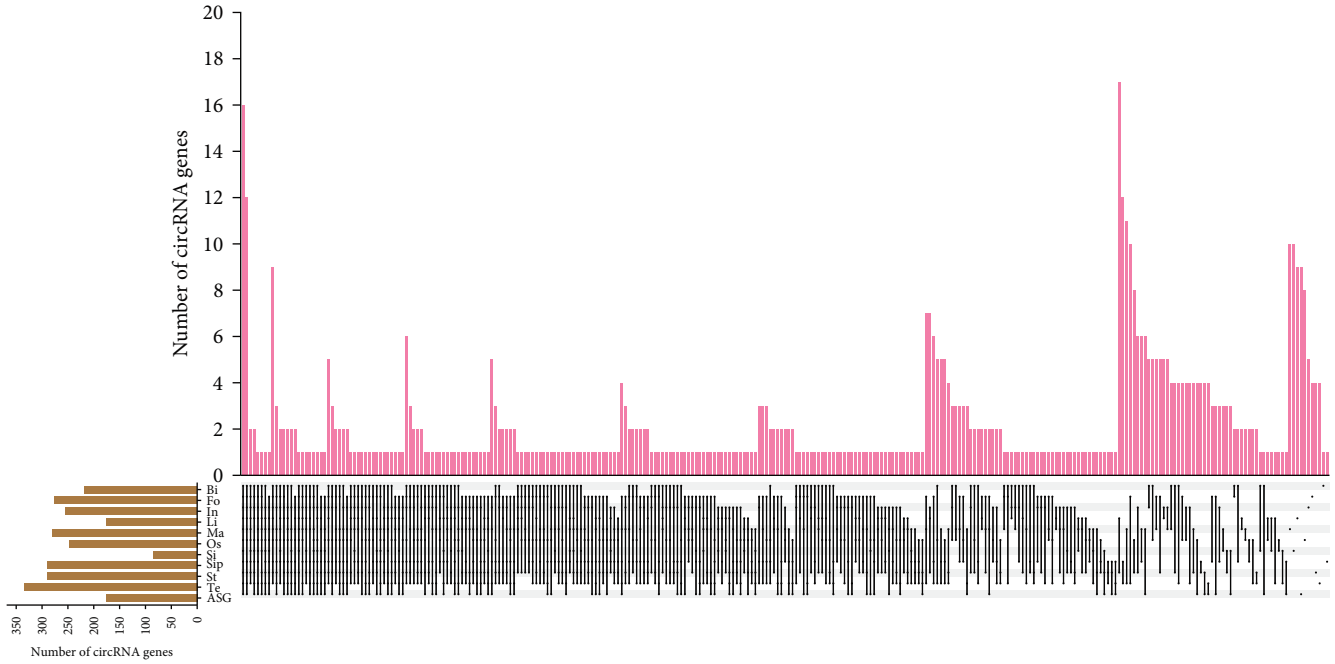
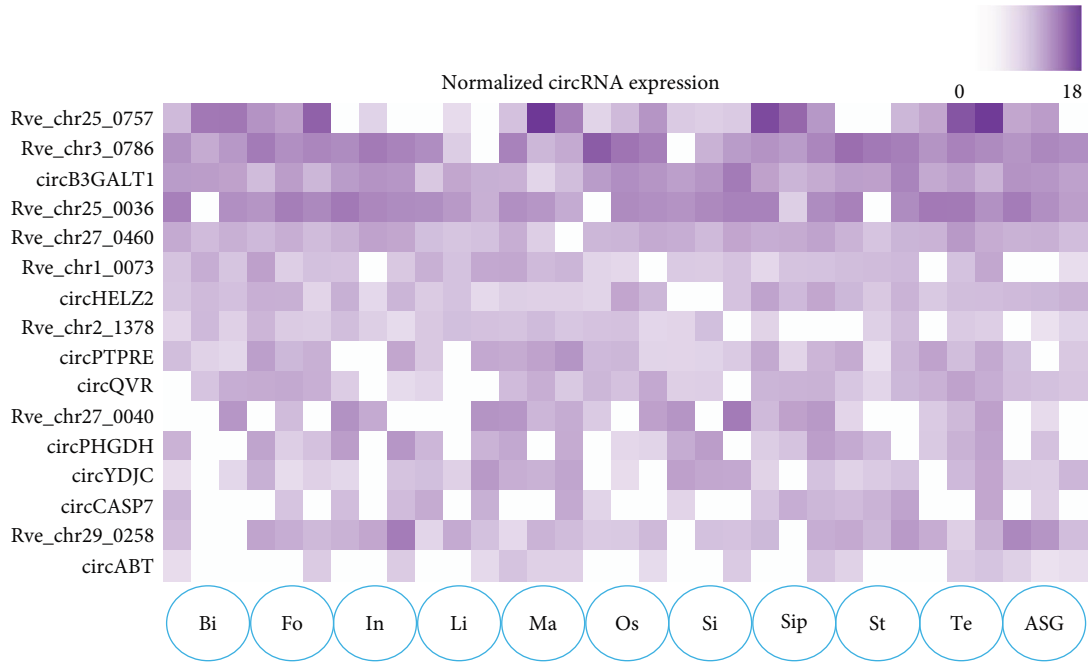


FIGURE 5: The UpsetR plot showing the intersection of circRNA genes ($n = 640$) across 11 tissues. The lower left panel depicts the number of circRNA genes identified in each tissue. In the main panel, the lines connecting dots (tissues), together with the bar on the top, illustrate the number of circRNA genes commonly identified across the corresponding tissues.

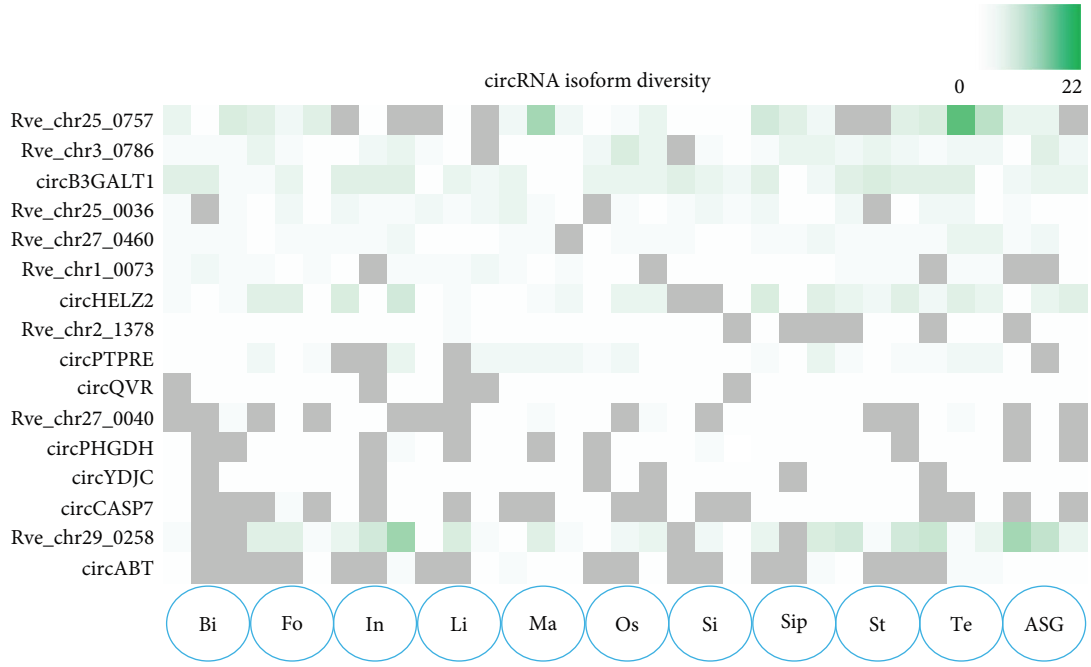
However, in contrast to other tissues, the Si exhibited six circRNA genes that consistently produced a large number of isoforms across biological replicates. Notably, all these circRNA genes displayed high expression levels exclusively in the Si. Specifically, the parental genes of two exhibited weak homology to the genes (*nas-14* and *nas-38*) encoding zinc metalloprotease enzymes in *Caenorhabditis elegans*. One notable regulatory mechanism of circRNAs involves the possession of microRNA binding sites shared with their parental genes, allowing them to act as microRNA sponges and thereby, maintaining their parental gene expression [30]. Indeed, the parental genes of these two circRNA genes also demonstrated elevated expression levels in the Si of *R. venosa*. Previous studies have reported high abundance of zinc metalloprotease enzymes in the Si of ticks [49] and snake venom [50], where they are implicated in interfering with host peptide signalling cascades, inflammatory response and host matrix remodelling during the feeding process.

A considerable portion of circRNAs were lowly expressed across tissues (Figure 3), aligning with the findings of numerous previous studies profiling circRNAs in different taxa [51]. Together with the observed heterogeneity in isoform diversity across samples, our results demonstrate that many circRNAs may have not been subject to substantial selective constraints over time, indicating a minimal regulatory impact on biological processes. However, we identified five circRNA genes consistently exhibiting upregulation (FDR < 0.05) in the Si compared to the other 10 tissues. Noticeably, among these upregulated circRNA genes are the two associated with parental genes encoding zinc metalloprotease enzymes aforementioned, as well as one related to *SIFaR*, a neuropeptide SIFamide receptor identified in *Drosophila melanogaster*. In fruit flies, *SIFaR* has been demonstrated to be involved in courtship behaviour, sleep regulation and hunger-driven behaviour [52]. Interestingly, the homologous gene of *SIFaR* has also been reported to be highly expressed in the Si of the black-legged tick [53]. However, the functional relevance of this gene specifically to the Si remains unclear.

We only identified 16 circRNA genes (2.5%) that were commonly expressed across all 11 tissues (Figure 6A). These circRNA genes typically displayed higher expression levels and have likely been to selection over a long evolutionary period, potentially playing crucial roles in regulating biological processes. As circRNAs primarily function as microRNA sponges, their high expression levels usually lead to the maintenance of their parental gene expression. Among these commonly expressed circRNA genes, *circB3GALT1* exhibited heightened expression levels across all the tissues (Figure 6A). Its parental gene *B3GALT1* encodes β-1,3-galactosyltransferase 1 enzyme, which plays an essential role in the biosynthesis of complex glycoproteins and glycolipids. *B3GALT1* is involved in diverse cellular functions essential for normal physiology, including glycosylation, extracellular matrix formation, cell signalling, immune response and development [54]. Likewise, *circHELZ2* was highly expressed across tissues and its parental gene *HELZ2* is involved in multiple cellular processes, such as transcriptional regulation, RNA processing and chromatin remodelling [55]. Both of them are indispensable for the functionality of a wide range of cell types. However, it is noteworthy that half of these 16 commonly expressed circRNA genes lacked homologues in the UniprotKB database, highlighting the need for further characterisation of these genes.



(A)



(B)

FIGURE 6: Continued.

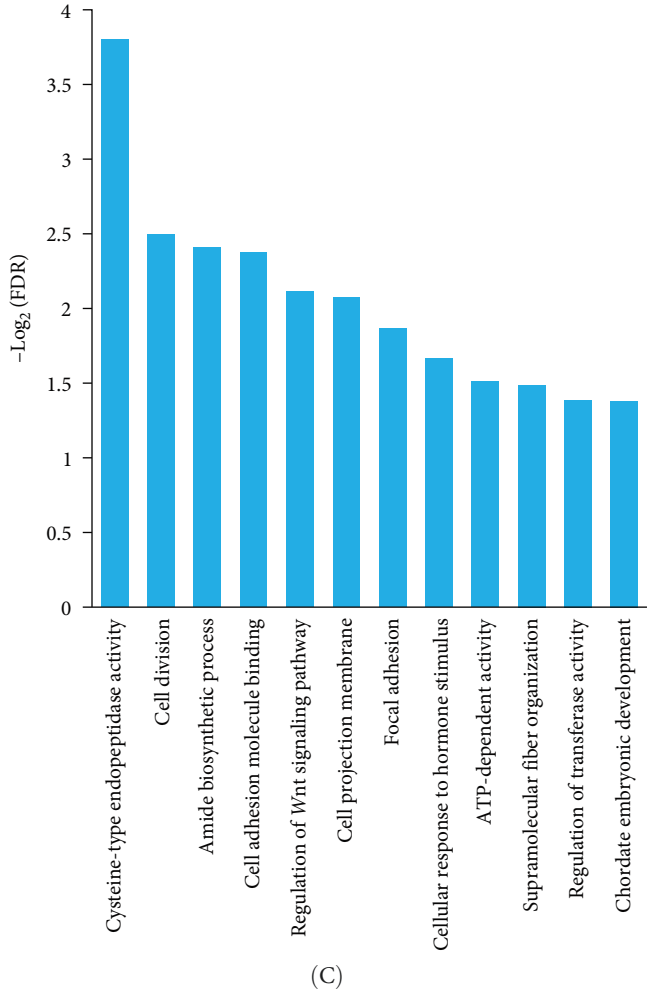


FIGURE 6: circRNA commonly expressed across tissues. (A) The heatmap showing the normalised expression levels of 16 commonly expressed circRNA genes across 11 tissues. (B) The heatmap illustrating the number of isoforms of 16 commonly expressed circRNA genes across 11 tissues. The grey colour indicates that a circRNA gene was not expressed. (C) The enrichment Gene Ontology pathways based on the circRNA genes expressed in at least nine out of 11 tissues.

We performed a GO enrichment analysis on the parental genes of circRNAs with annotations and expressed in at least nine out of 11 tissues (Figure 6C). The enriched terms encompass various fundamental biological processes, such as cell division, amide biosynthetic process, ATP-dependent activity and cellular response to hormone stimulus, all of which are essential for the functioning of diverse cell types. Simultaneously, we also detected a number of tissue-specific circRNA gene candidates, most of which exhibited low expression levels in respective tissues. Since the sequencing depth plays a crucial role in determining the number of identified circRNAs, it is important to note that tissue-specific circRNA genes may be discovered to be shared among different tissues as the sequencing depth increases.

This study describes the circRNA expression patterns across 11 tissues of the marine gastropod *R. venosa*, revealing an unexplored facet of transcriptional regulation in this species. Due to the incomplete characterisation of protein-coding genes in *R. venosa*, the functional relevance of identified circRNAs to tissue-specific pathways remains elusive. Future integration of paired microRNA datasets will help predict microRNA binding sites on both circRNAs and their parental genes, improving our understanding of their regulatory mechanisms and tissue specificity in *R. venosa*.

A preprint has previously been published [56].

Data Availability Statement

The raw RNA-Seq data are available via the National Centre for Biotechnology Information (NCBI) under the BioProject PRJNA855327. The genome of *R. venosa* was downloaded from NCBI via the accession number GCA_028751875.1.

Ethics Statement

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Author Contributions

Z.H. devised the study and wrote the first draft. M.H. and Z.H. analysed the data and interpreted the results. Both authors edited and approved the final manuscript.

Funding

This study is supported by the Irish Research Council Laureate Award bursary grant (IRCLA/2022/3212) awarded to Z.H.

Acknowledgments

We acknowledge the Irish Centre for High-End Computing and the UCD Sonic HPC cluster for the provision of computational facilities and support. This study is supported by the Irish Research Council Laureate Award bursary grant (IRCLA/2022/3212) awarded to Z.H.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Supporting Information 1. Table S1: The information of 33 RNA-seq samples used in this study.

Supporting Information 2. Table S2: The expression count of 640 high-confidence circRNA genes and their corresponding parental genes.

Supporting Information 3. Table S3: Tissue-specific circRNA genes.

Supporting Information 4. Figure S1: The distribution of isoform number of circRNA genes across 11 tissues. For each tissue, the percentage of circRNA genes with single isoforms and multiple isoforms was averaged by the three biological replicates. Figure S2: The overlap of circRNA genes with five or more isoform types across biological replicates for each tissue. Figure S3: The overlap of circRNA genes found across biological replicates for each tissue.

References

[1] W. F. Ponder and D. R. Lindberg, "Towards a Phylogeny of Gastropod Molluscs: An Analysis Using Morphological Characters," *Zoological Journal of the Linnean Society* 119, no. 2 (1997): 83–265.

[2] G. J. Vermeij and S. J. Carlson, "The Muricid Gastropod Subfamily Rapaninae: Phylogeny and Ecological History," *Paleobiology* 26, no. 1 (2000): 19–46.

[3] I. P. Bondarev, "Dynamics of *Rapana venosa* (Valenciennes, 1846)(Gastropoda: Muricidae) Population in the Black Sea," *International Journal of Marine Science* 4, no. 3 (2014).

[4] D.-X. Xue, J. Graves, A. Carranza, et al., "Successful Worldwide Invasion of the Veined Rapa Whelk, *Rapana venosa*, Despite a Dramatic Genetic Bottleneck," *Biological Invasions* 20, no. 11 (2018): 3297–3314.

[5] S. Moncheva, J. Namiesnik, R. Apak, et al., "*Rapana venosa* as a Bioindicator of Environmental Pollution," *Chemistry and Ecology* 27, no. 1 (2011): 31–41.

[6] M. Dağtekin, S. Candemir, G. Erik, and G. B. Mısır, "Determinants of Economic Efficiency: A Case Study of Rapa Whelk (*Rapana venosa*) Fisheries in the South Black Sea," *Ege Journal of Fisheries and Aquatic Sciences* 39, no. 1 (2022).

[7] D. Savini and A. Occhipinti-Ambrogi, "Consumption Rates and Prey Preference of the Invasive Gastropod *Rapana venosa* in the Northern Adriatic Sea," *Helgoland Marine Research* 60, no. 2 (2006): 153–159.

[8] J. M. Harding, "Growth and Development of Veined Rapa Whelk *Rapana venosa* Veligers," *Journal of Shellfish Research* 25, no. 3 (2006): 941–946.

[9] J. M. Harding, R. Mann, and C. W. Kilduff, "The Effects of Female Size on Fecundity in a Large Marine Gastropod *Rapana venosa* (Muricidae)," *Journal of Shellfish Research* 26, no. 1 (2007): 33–42.

[10] J. M. Harding, R. Mann, P. Moeller, and M. S. Hsia, "Mortality of the Veined Rapa Whelk, *Rapana venosa*, in Relation to a Bloom of Alexandrium Monilatum in the York River, United States," *Journal of Shellfish Research* 28, no. 2 (2009): 363–367.

[11] R. Mann and J. M. Harding, "Salinity Tolerance of Larval *Rapana venosa*: Implications for Dispersal and Establishment of an Invading Predatory Gastropod on the North American Atlantic Coast," *The Biological Bulletin* 204, no. 1 (2003): 96–103.

[12] N. Hu, F. Wang, T. Zhang, H. Song, Z.-L. Yu, and D.-P. Liu, "Prey Selection and Foraging Behavior of the Whelk *Rapana venosa*," *Marine Biology* 163, no. 11 (2016): 1–12.

[13] S. Ban, T. Zhang, H. Pan, Y. Pan, P. Wang, and D. Xue, "Effects of Temperature and Salinity on the Development of Embryos and Larvae of the Veined Rapa Whelk *Rapana venosa* (Valenciennes, 1846)," *Chinese Journal of Oceanology and Limnology* 32, no. 4 (2014): 773–782.

[14] M.-J. Yang, H. Song, Z.-L. Yu, et al., "Expression and Activity of Critical Digestive Enzymes During Early Larval Development of the Veined Rapa Whelk, *Rapana venosa* (Valenciennes, 1846)," *Aquaculture* 519 (2020): 734722.

[15] Z. Yang, H. Yu, R. Yu, and Q. Li, "Induced Metamorphosis in Larvae of the Veined Rapa Whelk *Rapana venosa* Using Chemical Cues," *Marine Biology Research* 11, no. 10 (2015): 1085–1092.

[16] D. Savini, M. Castellazzi, M. Favruzzo, and A. Occhipinti-Ambrogi, "The Alien Mollusc *Rapana venosa* (Valenciennes, 1846; Gastropoda, Muricidae) in the Northern Adriatic Sea: Population Structure and Shell Morphology," *Chemistry and Ecology* 20, no. sup1 (2004): 411–424.

[17] D. A. Giberto, C. S. Bremec, L. Schejter, A. Schiariti, H. Mianzan, and E. M. Acha, "The Invasive Rapa Whelk *Rapana venosa* (Valenciennes 1846): Status and Potential Ecological Impacts in the Río De LA Plata Estuary, Argentina-Uruguay," *Journal of Shellfish Research* 25, no. 3 (2006): 919–924.

[18] T. Sahin and S. Ergün, "Incorporation of Rapa Whelk (*Rapana venosa*) Meal in Diets for Rainbow Trout (*Oncorhynchus mykiss*) Fry," *Aquaculture Research* 52, no. 2 (2021): 678–692.

[19] T. Sahin, S. Yılmaz, M. Gürkan, and S. Ergün, "Effects of *Rapana venosa* Meal-Supplemented Diets on Reproduction, Histopathology and Some Blood Parameters of Rainbow Trout (*Oncorhynchus mykiss*) Broodstock," *Aquaculture Research* 52, no. 10 (2021): 4897–4910.

[20] H. Song, Z. Li, M. Yang, et al., "Chromosome-Level Genome Assembly of the Caenogastropod Snail *Rapana venosa*," *Scientific Data* 10, no. 1 (2023): 539.

[21] H. Song, H.-Y. Wang, and T. Zhang, "Comprehensive and Quantitative Proteomic Analysis of Metamorphosis-Related Proteins in the Veined Rapa Whelk, *Rapana venosa*,"

International Journal of Molecular Sciences 17, no. 6 (2016): 924.

[22] H. Song, Z.-L. Yu, L.-N. Sun, Y. Gao, T. Zhang, and H.-Y. Wang, "*De novo* Transcriptome Sequencing and Analysis of *Rapana venosa* From Six Different Developmental Stages Using Hi-seq 2500," *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* 17 (2016): 48–57.

[23] H. Song, Z.-L. Yu, L.-N. Sun, D.-X. Xue, T. Zhang, and H.-Y. Wang, "Transcriptomic Analysis of Differentially Expressed Genes during Larval Development of *Rapana venosa* by Digital Gene Expression Profiling," *G3: Genes, Genomes, Genetics* 6, no. 7 (2016): 2181–2193.

[24] Z. Zeng, C. Jiang, Q. Tan, B. Tang, and Z. Huang, "Larvae of a Marine Gastropod and a Marine Bivalve Share Common Gene Expression Signatures During Metamorphic Competence," *Marine Biology* 169, no. 9 (2022): 117.

[25] H. Song, L.-Y. Sun, Z.-L. Yu, et al., "Metabolomic Analysis of Competent Larvae and Juvenile Veined Rapa Whelks (*Rapana venosa*)," *Marine Biology* 163, no. 6 (2016): 1–8.

[26] M.-J. Yang, H. Song, J. Feng, et al., "Symbiotic Microbiome and Metabolism Profiles Reveal the Effects of Induction by Oysters on the Metamorphosis of the Carnivorous Gastropod *Rapana venosa*," *Computational and Structural Biotechnology Journal* 20 (2022): 1–14.

[27] H. Song, L. Qi, T. Zhang, and H.-Y. Wang, "Understanding microRNA Regulation Involved in the Metamorphosis of the Veined Rapa Whelk (*Rapana venosa*)," *G3: Genes, Genomes, Genetics* 7, no. 12 (2017): 3999–4008.

[28] L. S. Kristensen, M. S. Andersen, L. V. W. Stagsted, K. K. Ebbesen, T. B. Hansen, and J. Kjems, "The Biogenesis, Biology and Characterization of Circular RNAs," *Nature Reviews Genetics* 20, no. 11 (2019): 675–691.

[29] L. Verduci, E. Tarcitano, S. Strano, Y. Yarden, and G. Blandino, "CircRNAs: Role in Human Diseases and Potential use as Biomarkers," *Cell Death & Disease* 12, no. 5 (2021): 468.

[30] W.-Y. Zhou, Z.-R. Cai, J. Liu, D.-S. Wang, H.-Q. Ju, and R.-H. Xu, "Circular RNA: Metabolism, Functions and Interactions With Proteins," *Molecular Cancer* 19, no. 1 (2020): 1–19.

[31] X. Chen, P. Han, T. Zhou, X. Guo, X. Song, and Y. Li, "CircRNADb: A Comprehensive Database for Human Circular RNAs with Protein-Coding Annotations," *Scientific Reports* 6, no. 1 (2016): 34985.

[32] W. Wu, P. Ji, and F. Zhao, "CircAtlas: An Integrated Resource of One Million Highly Accurate Circular RNAs From 1070 Vertebrate Transcriptomes," *Genome Biology* 21, no. 1 (2020): 1–14.

[33] R. Hao, X. Zhu, C. Tian, Y. Huang, G. Li, and C. Zhu, "Identification and Characterization of circRNAs From Different Body Color Leopard Coral Grouper (*Plectropomus leopardus*)," *Frontiers in Marine Science* 10 (2023): 1201726.

[34] A. Farhadi, L. Lv, J. Song, et al., "Whole-Transcriptome RNA Sequencing Revealed the Roles of Chitin-Related Genes in the Eyestalk Abnormality of a Novel Mud Crab Hybrid (*Scylla serrata*♀× *S. paramamosain*♂)," *International Journal of Biological Macromolecules* 208 (2022): 611–626.

[35] S. Ibrahim, C. Yang, C. Yue, et al., "Whole Transcriptome Analysis Reveals the Global Molecular Responses of mRNAs, lncRNAs, miRNAs, circRNAs, and Their ceRNA Networks to Salinity Stress in Hong Kong Oysters, *Crassostrea hongkongensis*," *Marine Biotechnology* 25, no. 4 (2023): 624–641.

[36] M. Martin, "Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads," *EMBnet Journal* 17, no. 1 (2011): 10–12.

[37] Y. Gao, J. Wang, and F. Zhao, "CIRI: An Efficient and Unbiased Algorithm for *de novo* Circular RNA Identification," *Genome Biology* 16, no. 1 (2015): 1–16.

[38] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, "Basic Local Alignment Search Tool," *Journal of Molecular Biology* 215, no. 3 (1990): 403–410.

[39] H. Li and R. Durbin, "Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform," *Bioinformatics* 25, no. 14 (2009): 1754–1760.

[40] A. R. Quinlan and I. M. Hall, "BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features," *Bioinformatics* 26, no. 6 (2010): 841-842.

[41] C. Soneson, M. I. Love, and M. D. Robinson, "Differential Analyses for RNA-Seq: Transcript-Level Estimates Improve Gene-Level Inferences," *F1000Research* 4 (2015).

[42] M. D. Robinson, D. J. McCarthy, and G. K. Smyth, "EdgeR: A Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data," *Bioinformatics* 26, no. 1 (2010): 139-140.

[43] J. R. Conway, A. Lex, and N. Gehlenborg, "UpSetR: An R Package for the Visualization of Intersecting Sets and Their Properties," *Bioinformatics* 33, no. 18 (2017): 2938–2940.

[44] Y. Zhou, B. Zhou, L. Pache, et al., "Metascape Provides a Biologist-Oriented Resource for the Analysis of Systems-Level Datasets," *Nature Communications* 10, no. 1 (2019): 1523.

[45] R. C. Team, "R: A Language and Environment for Statistical Computing," (R Foundation for Statistical Computing, Vienna, Austria (2014).

[46] X. Zhao, X. Duan, W. Zhang, M. Guo, and C. Li, "Genome-Wide Identification of Circular RNAs Revealed the Dominant Intergenic Region Circularization Model in *Apostichopus japonicus*," *Frontiers in Genetics* 10 (2019): 461046.

[47] Y. Zhang, X.-O. Zhang, T. Chen, et al., "Circular Intronic Long Noncoding RNAs," *Molecular Cell* 51, no. 6 (2013): 792–806.

[48] Y.-J. Cai, W. Huang, L.-Y. Zhu, et al., "Identification of circRNAs and circRNA-mRNA Network of *Epinephelus coioides* During Singapore Grouper Iridovirus Infection," *Fish & Shellfish Immunology* 142 (2023): 109113.

[49] J. Perner, D. Helm, P. Haberkant, et al., "The Central Role of Salivary Metalloproteases in Host Acquired Resistance to Tick Feeding," *Frontiers in Cellular and Infection Microbiology* 10 (2020): 563349.

[50] C.d F. P. Teixeira, C. M. Fernandes, J. P. Zuliani, and S. F. Zamuner, "Inflammatory Effects of Snake Venom Metalloproteinases," *Memórias Do Instituto Oswaldo Cruz* 100, no. suppl 1 (2005): 181–184.

[51] W. Wu, J. Zhang, X. Cao, Z. Cai, and F. Zhao, "Exploring the Cellular Landscape of Circular RNAs Using Full-Length Single-Cell RNA Sequencing," *Nature Communications* 13, no. 1 (2022): 3242.

[52] C. Martelli, U. Pech, S. Kobbenbring, et al., "SIFamide Translates Hunger Signals Into Appetitive and Feeding Behavior in *Drosophila*," *Cell Reports* 20, no. 2 (2017): 464–478.

[53] L. Šimo, D.Ž. Itňan, and Y. Park, "Two Novel Neuropeptides in Innervation of the Salivary Glands of the Black-Legged Tick, *Ixodes scapularis*: Myoinhibitory Peptide and SIFamide," *Journal of Comparative Neurology* 517, no. 5 (2009): 551–563.

[54] T. Hennet, "The Galactosyltransferase Family," *Cellular and Molecular Life Sciences (CMLS)* 59, no. 7 (2002): 1081–1095.
[55] A. Luqman-Fatah, Y. Watanabe, K. Uno, F. Ishikawa, J. V. Moran, and T. Miyoshi, "The Interferon Stimulated Gene-Encoded Protein HELZ2 Inhibits Human LINE-1 Retrotransposition and LINE-1 RNA-Mediated Type I Interferon Induction," *Nature Communications* 14, no. 1 (2023): 203.
[56] M. Hong and Z. Huang, "The Circular RNA Expression Landscape across Tissues Implicates Their Functional Roles in the Marine Gastropod *Rapana venosa*," *Research Square* (2024).