

Communication

Transcriptomic Diversity in the Livers of South African Sardines Participating in the Annual Sardine Run

Arsalan Emami-Khoyi ¹, Rynhardt Le Roux ¹, Matthew G. Adair ¹, Daniela M. Monsanto ¹, Devon C. Main ¹, Shilpa P. Parbhu ¹, Claudia M. Schnelle ¹, Carl D. van der Lingen ^{2,3}, Bettine Jansen van Vuuren ¹ and Peter R. Teske ^{1,*}

¹ Centre for Ecological Genomics and Wildlife Conservation, Department of Zoology, University of Johannesburg, Auckland Park 2006, South Africa; ekarsalan@gmail.com (A.E.-K.); rynhardt07@gmail.com (R.L.R.); mattgadair@gmail.com (M.G.A.); dmonsanto119@gmail.com (D.M.M.); maindevon@gmail.com (D.C.M.); shilpa.parbhu@yahoo.com (S.P.P.); cmschnelle@gmail.com (C.M.S.); bettinevv@uj.ac.za (B.J.v.V.)

² Branch: Fisheries Management, Department of Environment, Forestry and Fisheries, Private Bag X2, Vlaeberg 8012, South Africa; carl.vanderlingen@gmail.com

³ Department of Biological Sciences and Marine Research Institute, University of Cape Town, Private Bag X3, Rondebosch 7700, South Africa

* Correspondence: pteske101@gmail.com

Abstract: During austral winter, the southern and eastern coastlines of South Africa witness one of the largest animal migrations on the planet, the KwaZulu-Natal sardine run. Hundreds of millions of temperate sardines, *Sardinops sagax*, form large shoals that migrate north-east towards the subtropical Indian Ocean. Recent studies have highlighted the role that genetic and environmental factors play in sardine run formation. In the present study, we used massively parallel sequencing to assemble and annotate the first reference transcriptome from the liver cells of South African sardines, and to investigate the functional content and transcriptomic diversity. A total of 1,310,530 transcripts with an N50 of 1578 bp were assembled de novo. Several genes and core biochemical pathways that modulate energy production, energy storage, digestion, secretory processes, immune responses, signaling, regulatory processes, and detoxification were identified. The functional content of the liver transcriptome from six individuals that participated in the 2019 sardine run demonstrated heterogeneous levels of variation. Data presented in the current study provide new insights into the complex function of the liver transcriptome in South African sardines.

Keywords: *Sardinops sagax*; sardine run; liver transcriptome; migration; RNA-seq; climate change



Citation: Emami-Khoyi, A.; Le Roux, R.; Adair, M.G.; Monsanto, D.M.; Main, D.C.; Parbhu, S.P.; Schnelle, C.M.; van der Lingen, C.D.; Jansen van Vuuren, B.; Teske, P.R.

Transcriptomic Diversity in the Livers of South African Sardines Participating in the Annual Sardine Run. *Genes* **2021**, *12*, 368. <https://doi.org/10.3390/genes12030368>

Academic Editor: Antonio Figueras

Received: 20 January 2021

Accepted: 22 February 2021

Published: 4 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The South African coastline is unique in that it is strongly influenced by two of the Earth's major ocean currents, the cold Benguela Current and the warm Agulhas Current. This distinctive dynamic support some of the richest marine biodiversity on the planet [1]. During austral winter, typically between May and early June, the coastlines of the Eastern Cape (EC) and KwaZulu-Natal (KZN) provinces in South Africa witness one of Earth's largest annual animal migrations, the sardine run. Over the course of this mass migration, hundreds of millions of individuals from the South African population of Pacific sardine, *Sardinops sagax*, form large shoals in the temperate waters of the EC, which then swim north-east along the eastern South African coastline towards the subtropical waters of KZN [2].

Marine species typically migrate to find favourable spawning conditions or to exploit food resources [3,4]. However, the exact dynamic that drives sardine run formation remains unknown. Some studies point to the complex roles that genetic and environmental factors play in forming this mass migration [2,5,6]. Individuals that participate in the sardine run

face significant challenges, with the large number of migrating fish attracting a myriad of predators. Furthermore, this event plays a crucial role in the livelihoods and economy of local communities [7].

Recent advances in DNA sequencing technologies and computational resources have made it possible to study the functional content of the genome in a way that was previously impossible. Massively parallel sequencing of the transcriptome from different tissue types has been used to address the biological questions of species from diverse terrestrial and marine environments [8–15].

To better understand the role that genetics plays in the life-history of South African sardines and the formation of the sardine run, extensive genomic resources for this species are needed. One means of doing so is to study the functional content and diversity of the South African sardine genome. Such data not only provide insights into the genetic structure and origin of the individuals that form shoals, but they also shed light on the physiological status of the sardines participating in the migration.

The functional content of the liver transcriptome in marine vertebrates is particularly informative. Not only do hepatic cells play an essential role in the detoxification of harmful metabolites and anthropogenic pollutants that are accumulating in marine ecosystems at an unprecedented rate, but they are also involved in energy production, energy storage, and in the initiation of various physiological or immune responses necessary to fight pathogens [11,16,17].

Migrating animals face a combination of energetic and physiological challenges that differ from those during the non-migrating stages of their life cycles. Several authors have highlighted the critical roles that biochemical pathways in the liver play in migratory species prior to or during migration events [18–20].

As a first step towards an in-depth understanding of South African sardine biology, we assembled and annotated the first reference transcriptome of liver cells for this species. We further investigated the functional content and genetic diversity among six individuals that participated in the 2019 sardine run.

2. Materials and Methods

2.1. Animal Ethics Statement

Animal ethics approval for specimen collection was obtained from the Faculty of Science Ethics Committee, University of Johannesburg (approval code:19022015).

2.2. Specimen Collection

Six sardines were obtained in KZN during the 2019 sardine run—three individuals from Scottburgh (30°17′12.1″S 30°45′11.5″E) and three from Umgababa (30°08′43.7″S 30°50′00.5″E). All specimens were purchased from artisanal fishers. They had been caught using a beach seine net <15 min prior to specimen collection and had been stored on ice. Approximately 1 mm³ of liver tissue was excised immediately upon purchase and was transferred to a 10:1 volume of *RNAlater* solution (Qiagen, Hilden, Germany). The liver samples were then stored at −80 °C until the total RNA was extracted within a week after collection.

2.3. Nucleic Acid Extraction, Genomic Library Preparation, and Sequencing

The total RNA was extracted from the liver tissue using a combination of Trizol and Qiagen's RNeasy Mini Kit. A total amount of 0.4 µg RNA per extraction was used as the template for the RNA library preparation. Complementary DNA (cDNA) libraries were generated using the NEBNext Ultra RNA Library Prep Kit for Illumina (New England Biolabs, MA, USA), following the manufacturer's recommendations. This was done by first purifying the mRNA from the pool of extracted RNA using poly-T oligo-attached magnetic beads, followed by mRNA fragmentation using divalent cations under an elevated temperature in a NEBNext First Strand Synthesis Reaction Buffer (5×). The first cDNA strand was synthesized using a random hexamer primer and a M-MuLV Reverse

Transcriptase (RNase H-) enzyme. Similarly, the synthesis of the second strand of cDNA was performed using DNA Polymerase I and RNase H-. The overhangs were blunt ended by using a combination of exonuclease and polymerase activities. After the adenylation of the 3' ends of the cDNA fragments, NEBNext Adaptors with a hairpin loop structure were ligated to each library, and the resulting cDNA fragments (~250–300 bp) were purified for an additional step using the AMPure XP system prior to being sequenced on an Illumina HiSeq 4000 platform (San Diego, CA, USA), using a 2 × 150 bp TruSeq PE Cluster Kit v3-cBot-HS following the manufacturer's instructions.

2.4. Transcriptome Assembly, Functional Annotation, and Variant Calling

To construct a comprehensive reference liver transcriptome for South African sardines, the data from the six individuals generated in the current study were combined with sequence data from 20 additional South African sardines' liver transcriptomes from another study (Teske et al., unpublished data), prior to quality filtering and transcriptome assembly.

Leading and trailing low-quality sequences, N bases, all base pairs with an average Phred score <25 in a 4-bp sliding window, and Illumina adapter contaminations were removed using Trimmomatic (v0.39) [21]. A reference transcriptome for the liver cells of the South African sardine was assembled de novo using Trinity v2.8.6 [22] by sequentially running the Inchworm, Chrysalis, and Butterfly modules, using their default settings. The most likely reading frame within each transcript was predicted in TransDecoder v5.5.0 [22]. All transcripts were functionally annotated using the Trinotate pipeline [23], with transcripts being searched against known proteins and core metabolic pathways deposited in the Swiss-Prot, NCBI non-redundant protein, and Kyoto Encyclopaedia of Genes and Genomes databases. In addition, and where possible, the protein domains for the assembled transcripts were separately predicted in HMMER v3.1 [24], based on similarity to the deposited protein domains in the Pfam database. Finally, transmembrane proteins and ribosomal RNA were identified using TmHMM v2 [25] and RNAMMER v1.2 [26], respectively.

The gene content of the liver transcriptome was evaluated using BUSCO v4.1.4 and the ray-finned fish core BUSCOs database (actinopterygii_odb10) in transcriptome mode [27].

The quality-filtered sequences of the six sardines collected during the 2019 sardine run were aligned against the assembled reference liver transcriptome using semi-global BWA-MEM aligners [28]. A combination of the SAMtools v1.9 mpileup command [29] and Varscan2 v2.3.7 [30] was used to identify variant sites (Single Nucleotide Polymorphic sites, SNPs). To improve the accuracy of variant calling in Varscan2, the default *p*-value threshold for a variant site was lowered to 0.001.

The functional content and intra-individual genomic variation of the liver transcriptome were investigated by creating two independent datasets. In the first dataset, the abundance for each transcript was estimated using RSEM v1.3.1 [31], and the top 1% of transcripts with the highest normalized level of expression (estimated as the number of transcripts per million mapped reads, TPM) were then selected for more detailed functional analyses. The second dataset consisted of 25% of the assembled transcripts with the highest number of polymorphic sites independent of gene size, as reported in the Varscan2 output files. The enrichment of the generated datasets for particular biological functions in terms of Gene Ontology categories (GO terms) [32] were investigated using the g:Profiler online server [33], with significance thresholds adjusted for multiple testing using g:SCS methods. In the enrichment analysis, the annotated genome of the zebrafish, *Danio rerio* (e101_eg48_p14_baf17f0), was selected as the taxonomically most closely related high-quality annotated gene set against which the statistical significance of the enrichment of the liver transcriptome was tested. Finally, a combination of the WEGO v2.0 online tool [34], REVIGO online server [35], and CirGO [36] were used to visualize the results.

3. Results and Discussion

The current study is the first to assemble de novo and annotate a reference transcriptome from the liver tissue of South African sardines.

The Trinity pipeline assembled raw sequences (Supplementary Table S1) into 1,310,530 transcripts with an N50 value of 1578 bp (Supplementary Table S2, Supplementary File sardine.fasta). The gene content analysis of the liver transcriptome identified 73.8% of the complete ray-finned fish BUSCOs within South African sardine liver transcripts. In addition, 8.1% of the predicted BUSCOs were fragmented and 18.1% were missing from the expression profile. Varscan2 identified a total of 68,145 variant sites (SNP density of 0.18 per kb) in the six sardines.

The functional content of the liver transcriptome in South African sardines is consistent with the essential role that hepatic cells play in energy production, energy storage, digestion, secretory, physiological or immune responses to stressors, and the detoxification of harmful chemical compounds [11,16,17].

The liver transcripts were heavily enriched for molecular functions such as binding, catalytic, regulatory, and transducer activity. Similarly, biological processes such as those involved in the cellular macromolecule metabolic processes, physiological responses to internal and external stimuli, immune responses, cell communication, and cell signaling were significantly overrepresented (Figures 1–3).

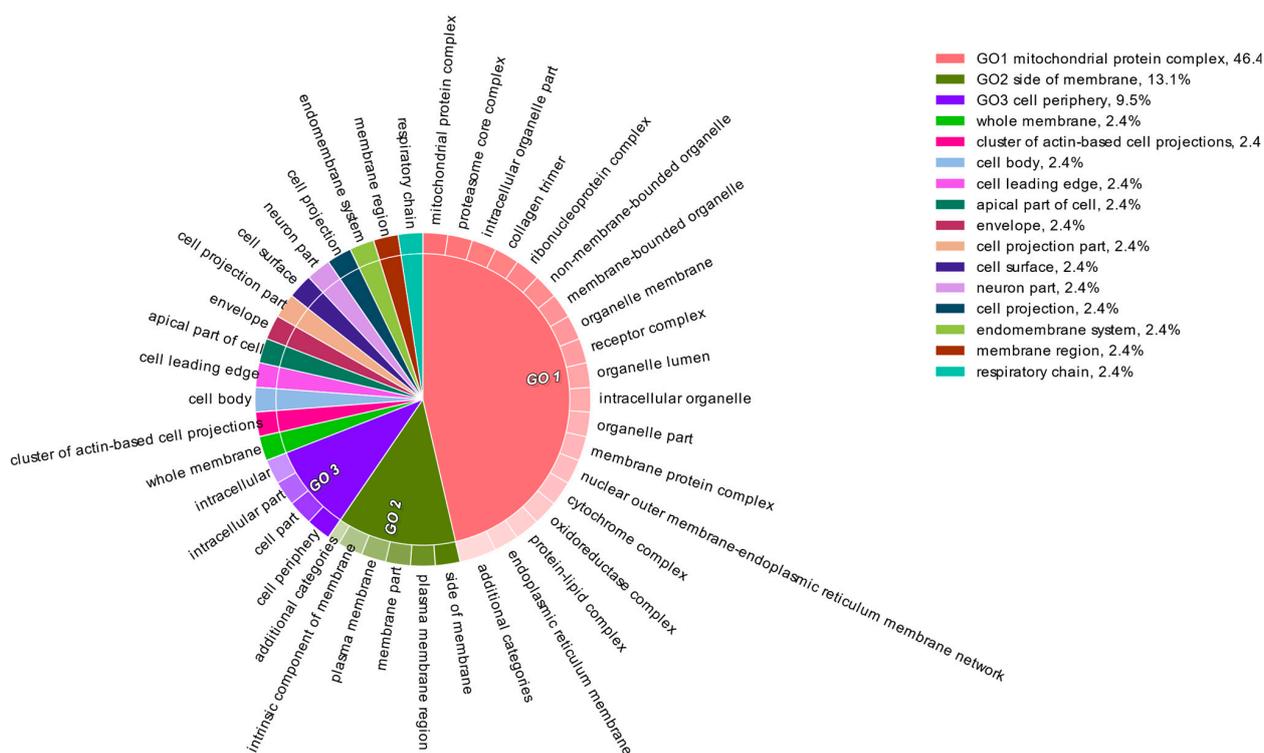


Figure 1. Functional content in terms of the Gene Ontology categorization of “cellular component” for the highly expressed transcripts from the liver transcriptome of South African sardines.

The KEGG pathway analysis of the South African sardine liver transcriptome highlighted that biochemical pathways that function as signal transducers ($n = 3812$), those involved in general metabolic pathways ($n = 3579$), in the immune system ($n = 1475$), and in the endocrine system ($n = 1473$) are active in the liver cells. Not surprisingly, a large number of metabolic pathways that are associated with the biodegradation of xenobiotic substances ($n = 210$), amino acid or lipid metabolism ($n = 1134$), the digestive system ($n = 442$), excretory system ($n = 121$), and environmental adaptation ($n = 289$) were also identified (Table 1).

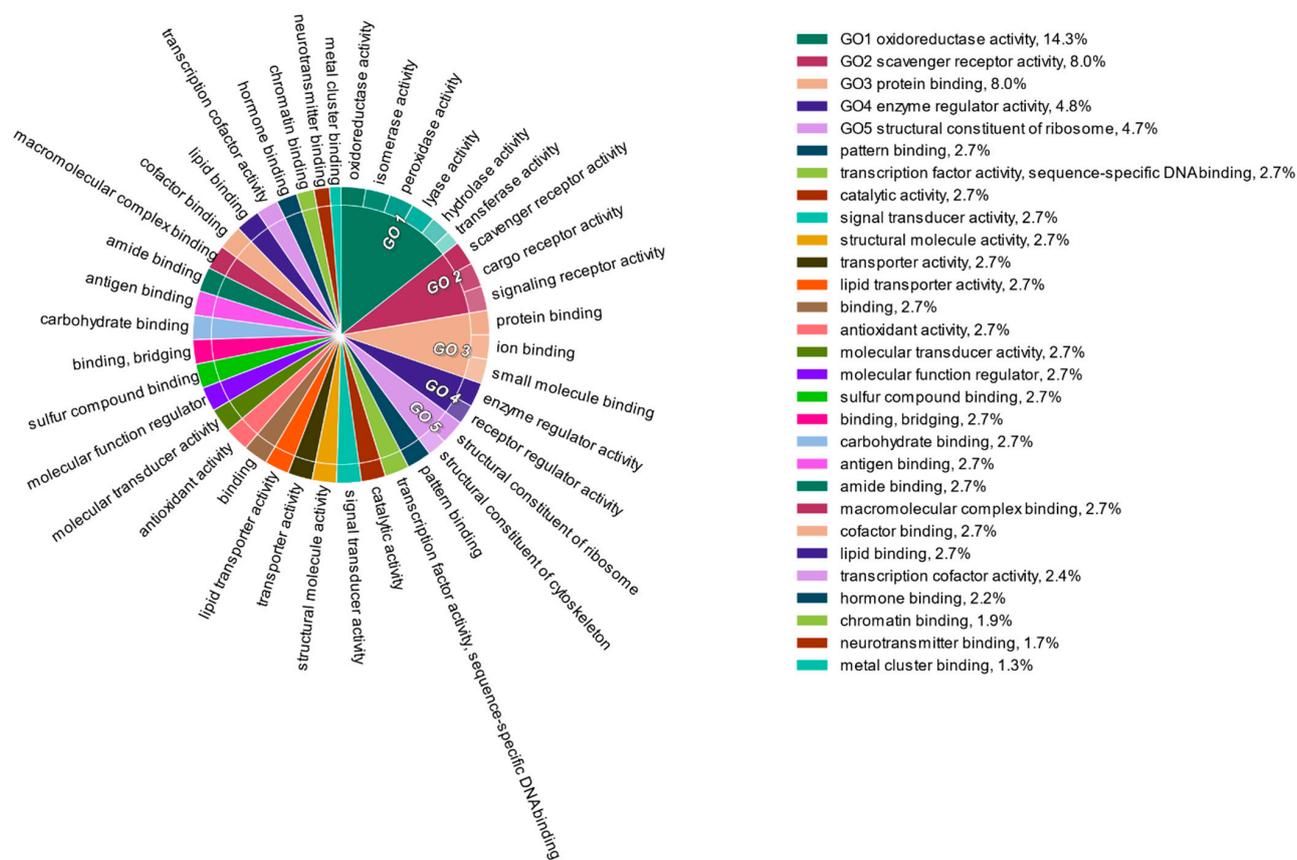


Figure 2. Functional content in terms of the Gene Ontology categorization of “molecular function” for the highly expressed transcripts from the liver transcriptomes of South African sardines.

In the dataset that consists of the 1% highly expressed genes in the South African sardine liver transcriptome (Supplementary Table S3), genes such as glutathione S-transferase (GTS) are a well-studied family of liver enzymes that play a critical role in the detoxification of electrophilic toxins [37]. High expression levels for these genes have been reported from both healthy liver cells and hepatic cells after exposure to toxic substances in different fish species, such as plaice (*Pleuronectes platessa*), flounder (*Platichthys flesus*), big-belly seahorse (*Hippocampus abdominalis*), and grass carp (*Ctenopharyngodon idella*) [38–44].

Similarly, a high expression of zona pellucida glycoprotein 3 (*zp3*) gene, which is involved in sperm–oocyte binding, was also reported from the liver cells of trout (*Oncorhynchus mykiss*), salmon (*Salmo salar*), and winter flounder (*Pseudo-pluronectes americanus*), as well as from the ovaries of zebrafish and common carp (*Cyprinus carpio*) [45,46]. Conner and Hughes [46] suggested that an ancient gene duplication event in teleost evolution resulted in two copies of these genes being independently active in ovaries and liver tissues, and that the expression level in each organ depends on a complex regulatory mechanism that involves cross-talks between tissue types and tissue-specific regulatory mechanisms. Among other highly expressed genes, the pleiotropic gene, X-box binding protein 1 (*XBPI*), triggers an endoplasmic stress response in zebrafish [47], and protein *SelK* (selenoprotein K) initiates innate immunological responses in common carp and rainbow trout (*Oncorhynchus mykiss*) [48]. The expression levels of *cirbp* (cold inducible RNA binding protein) and *csde1* (cold shock domain containing E1 genes) in the sardines’ livers are noteworthy, as both of these genes are involved in physiological responses to cold temperature. In the zebrafish, *cirbp* acts as a molecular chaperone that aids in unfolding RNA secondary structures, especially in colder water [49]. In other vertebrates, the upregulation of the *csde1* gene has been reported in response to the stress experienced during rapid temperature decrease [50] through complex nucleic acid-binding activities [51].

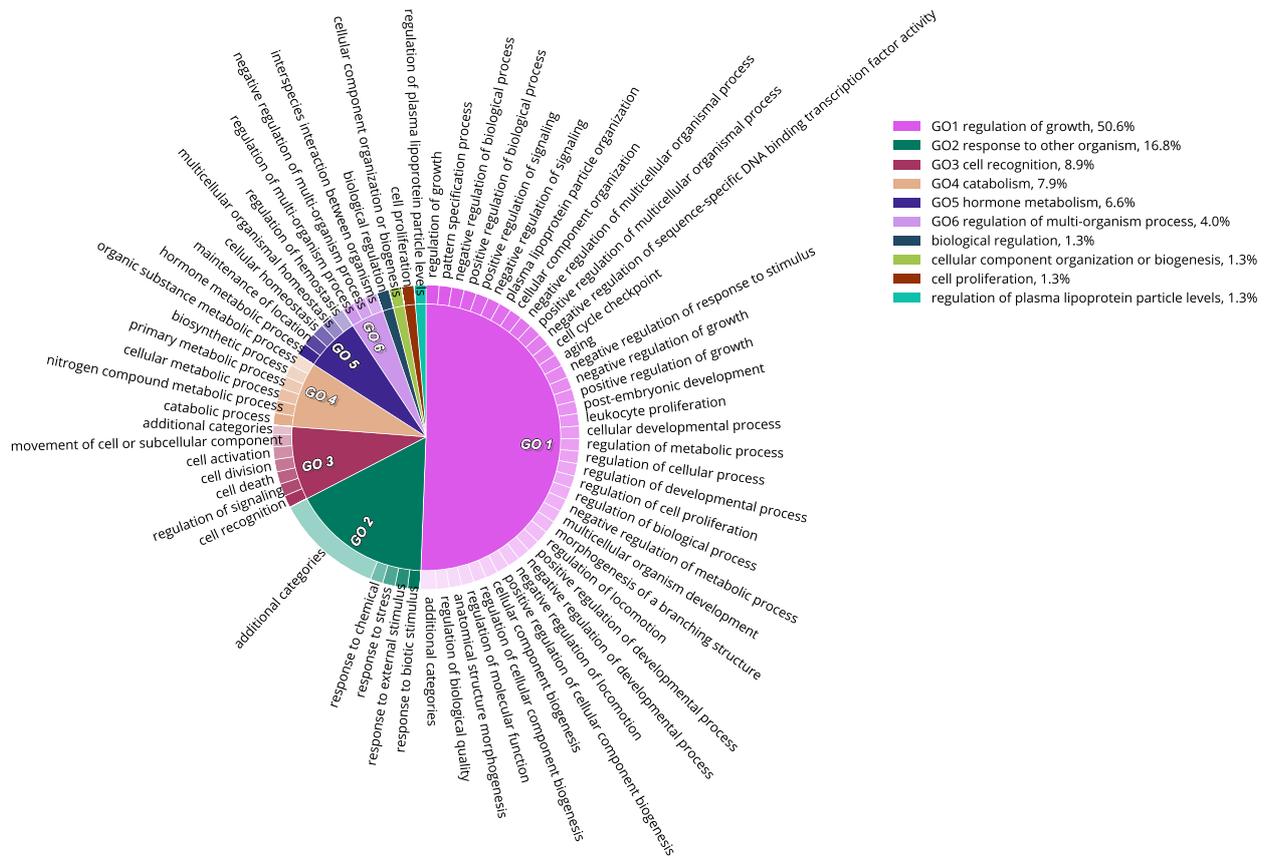


Figure 3. Functional content in terms of the Gene Ontology categorization of “biological process” for the highly expressed transcripts from the liver transcriptome of South African sardines.

A number of genes that are involved in oxygen transport (such as *Hbb*) and energy production pathways in the mitochondrion (such as *ATP5H*, *ATP5I*, *ATP5J*, *ATP5L*, *CYB5*, *CYC*, *HOGA1*, *MSRB2*, *UCP1*, *UCP2*, and *UCP3*) demonstrated comparatively high expression levels in the South African sardine liver transcriptome. While high expressions of the genes that are involved in energy production and those that modulate the physiological or immunological responses to the environmental stressors in the hepatic cells is not surprising, the extent to which the observed pattern may provide migrating fish with the sudden energetic and physiological requirements prior to or during migration events is yet to be fully understood, and this requires more detailed comparative studies between sardines that participate in the sardine run and those that do not. Sharma et al. [18] described enrichment in similar metabolic pathways within the liver transcriptome of migratory birds participating in seasonal migrations.

Among the highly polymorphic genes in the South African sardines’ liver transcriptome (Supplementary Table S4), the product of the arsenic (+3 oxidation state) methyltransferase gene (*AS3MT*) metabolizes naturally-occurring or xenobiotic arsenic compounds into the easily-excretable monomethylarsonic and dimethylarsonic products [52]. Similarly, glutathione S-transferase pi 1 (*GSTP1*) and glutathione S-transferase pi 2 (*GSTP2*) genes catalyse glutathione into electrophilic compounds, which are essential in the detoxification process [53]. Polymorphism in *GREB1* is notable, as the expression of this gene has been shown to synchronize the arrival of sexually mature or premature salmonid fish to their spawning grounds [54]. Other highly polymorphic genes are involved in a wide range of metabolic, regulatory, and catalytic activities.

Table 1. Description and number of occurrences of Kyoto Encyclopaedia of Genes and Genomes (KEGG) metabolic pathways predicted from the liver transcriptome of South African sardine.

Name of the Pathway	Occurrences
Signal transduction	3812
Global and overview maps	3579
Immune system	1475
Endocrine system	1473
Transport and catabolism	1424
Carbohydrate metabolism	774
Nervous system	750
Amino acid metabolism	647
Lipid metabolism	487
Cellular community—eukaryotes	468
Folding, sorting, and degradation	448
Digestive system	442
Translation	416
Energy metabolism	383
Glycan biosynthesis and metabolism	376
Metabolism of cofactors and vitamins	351
Replication and repair	299
Environmental adaptation	289
Development and regeneration	284
Circulatory system	249
Nucleotide metabolism	226
Xenobiotics biodegradation and metabolism	210
Cell motility	210
Transcription	181
Sensory system	164
Ageing	150
Metabolism of other amino acids	148
Biosynthesis of other secondary metabolites	125
Excretory system	121
Metabolism of terpenoids and polyketides	86

We suggest that many of the genes predicted from South African sardines' liver transcriptome perform similar functions as they do in other vertebrate species. Our results also indicate that the high expression level and polymorphism in the glutathione S-transferase (*GTS*) gene can make their products a potential biomarker that is both easy to detect and variable enough to reflect the health status of different individuals. However, more physiological and toxicological studies are needed in order to evaluate sensitivity and specificity of this gene as a suitable biomarker, similar to how it has been used as a biomarker in other species [55,56].

The functional content of highly expressed transcripts from the six individuals that participated in the sardine run showed a heterogeneous variation (Figures 4–8). The molecular component category showed minimal variation among six individuals. However, the variations in the two other gene ontology categories, molecular function and biological process, were comparatively more pronounced. Molecular functions, such as lipid transporter and binding activity, carbohydrate and carbohydrate derivative binding activity, and a series of biochemical pathways that are involved in various catalytic activities acting on the proteins and on other biomolecules, showed comparatively high intra-individual variations. Similarly, biological processes such as those that modulate the physiological and transcriptomic responses to biotic stimuli, immune effector processes, multicellular organism reproduction, developmental growth, cell motility, locomotion, and responses to other organisms varied considerably among individuals. The significance of the observed pattern in the functional content of the liver transcriptome among individual sardines, which were subjected to the same environmental stimuli and stressors during migration and capture,

requires more detailed study of sardine populations under controlled conditions, which lies beyond the scope of the current study.

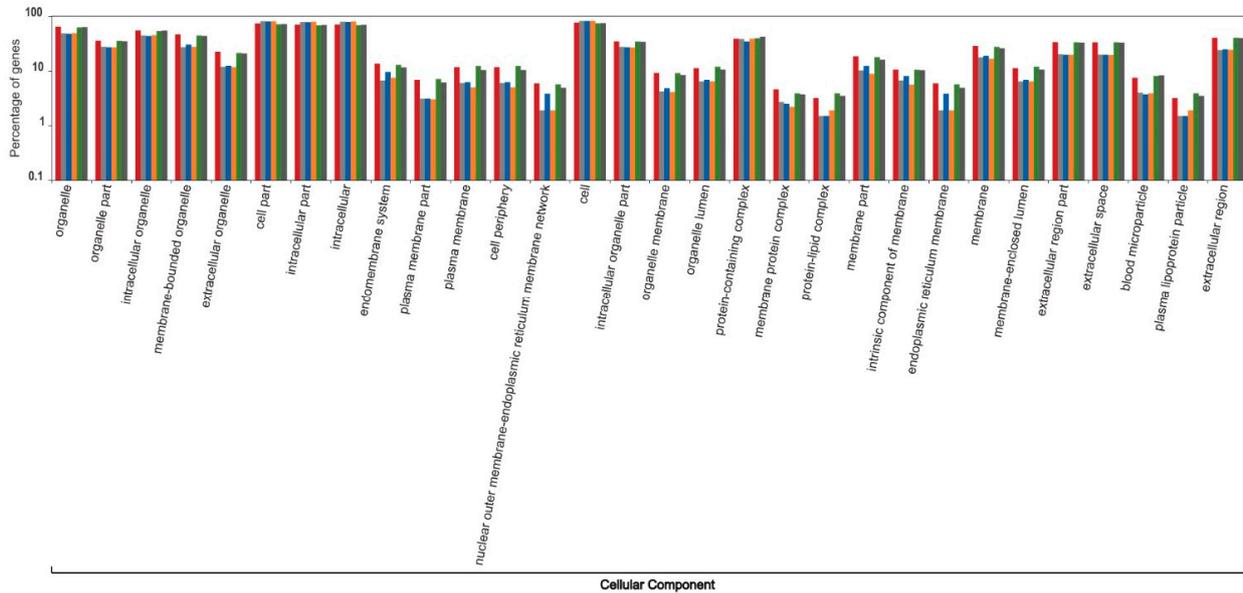


Figure 4. Intra-individual variation in the functional content of the top one percent of highly expressed transcripts within the liver transcriptome of South African sardines in terms of the Gene Ontology categorization “cellular component”, shown for six individuals that participated in the 2019 sardine run. Each individual is represented by a unique colour. Values on the Y axis are log transformed.

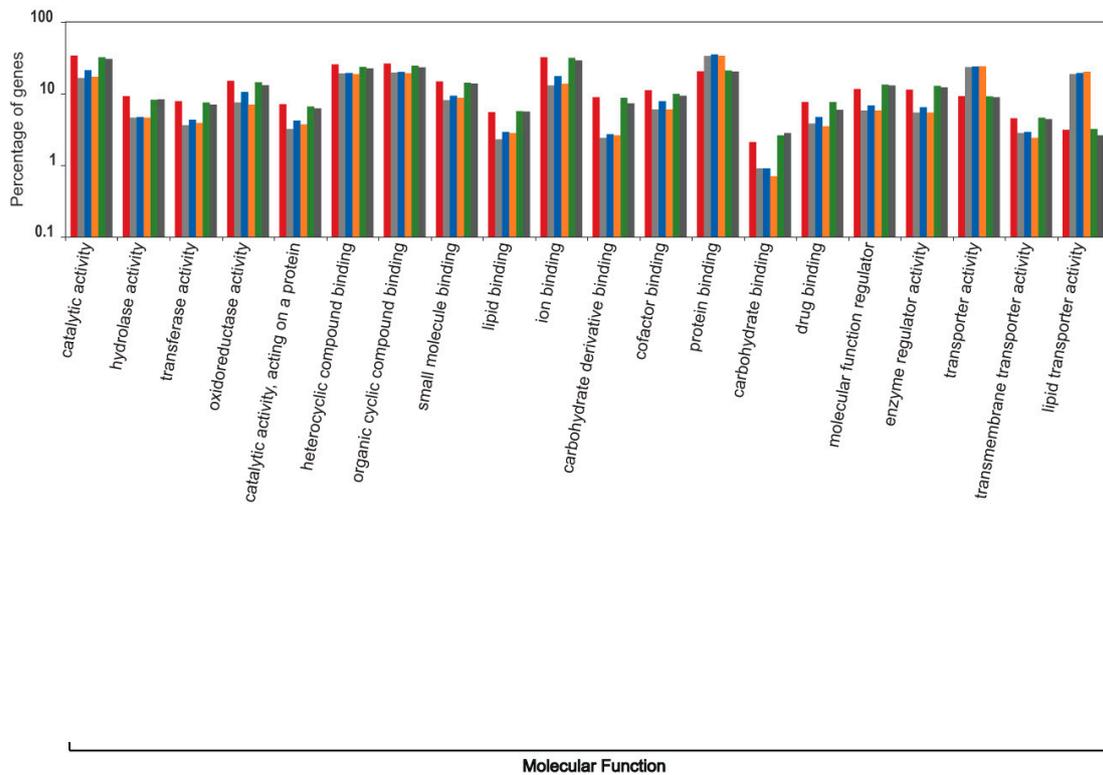


Figure 5. Intra-individual variation in the functional content of the top one percent of highly expressed transcripts within the liver transcriptome of South African sardines in terms of the Gene Ontology categorization “molecular function”, shown for six individuals that participated in the 2019 sardine run. Each individual is represented by a unique colour. Values on the Y axis are log transformed.

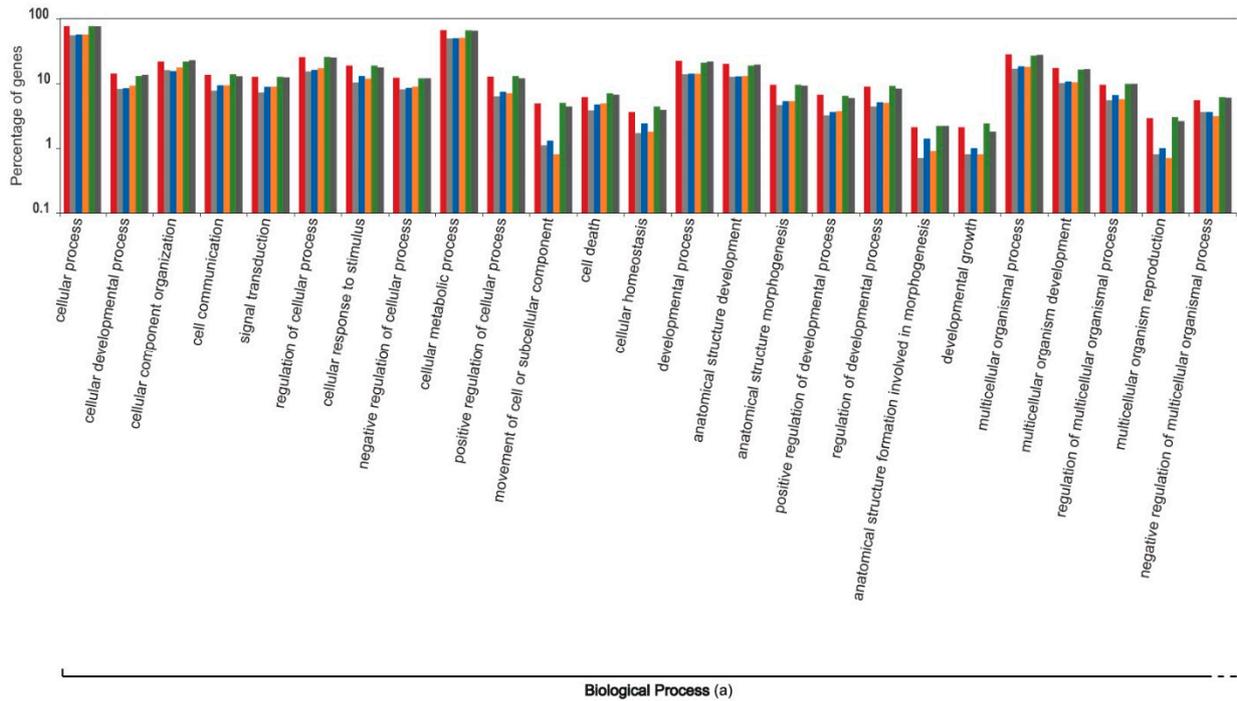


Figure 6. Intra-individual variation in the functional content of the top one percent of highly expressed transcripts within the liver transcriptome of South African sardines in terms of the Gene Ontology categorization “biological process”, shown for six individuals that participated in the 2019 sardine run. Each individual is represented by a unique colour. Values on the Y axis are log transformed.

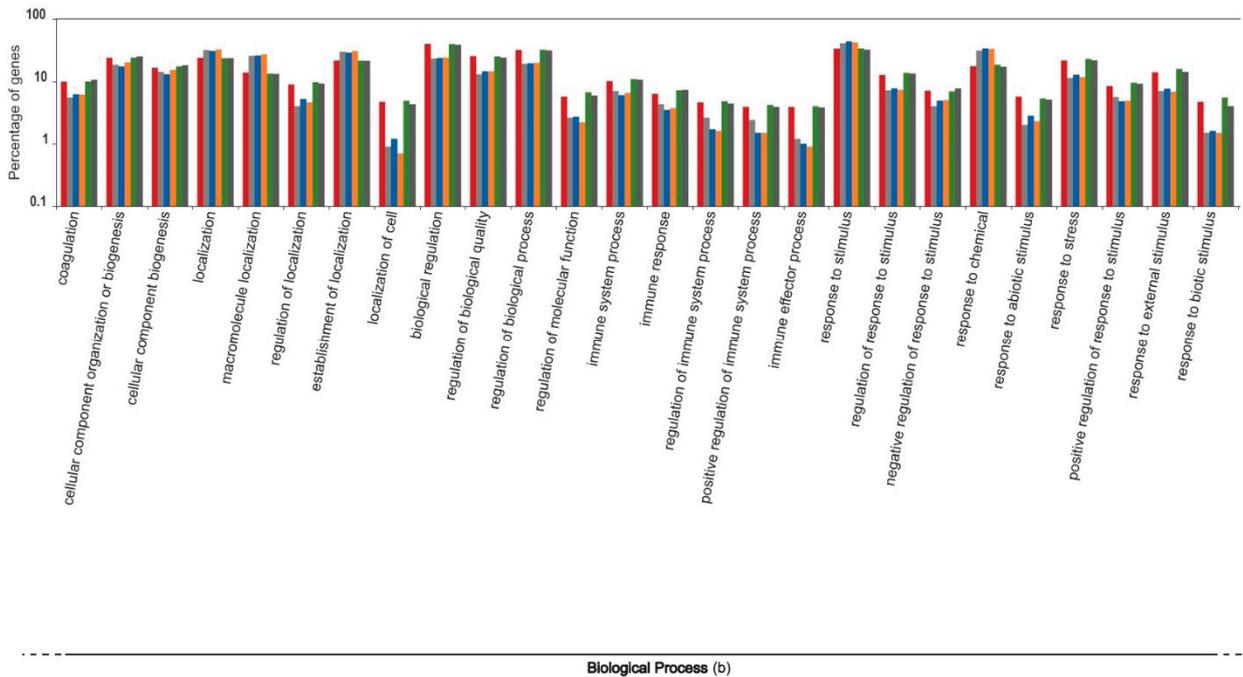


Figure 7. Intra-individual variation in the functional content of the top one percent of highly expressed transcripts within the liver transcriptome of South African sardines in terms of the Gene Ontology categorization “biological process”, shown for six individuals that participated in the 2019 sardine run. Each individual is represented by a unique colour. Values on the Y axis are log transformed (continued from Figure 6).

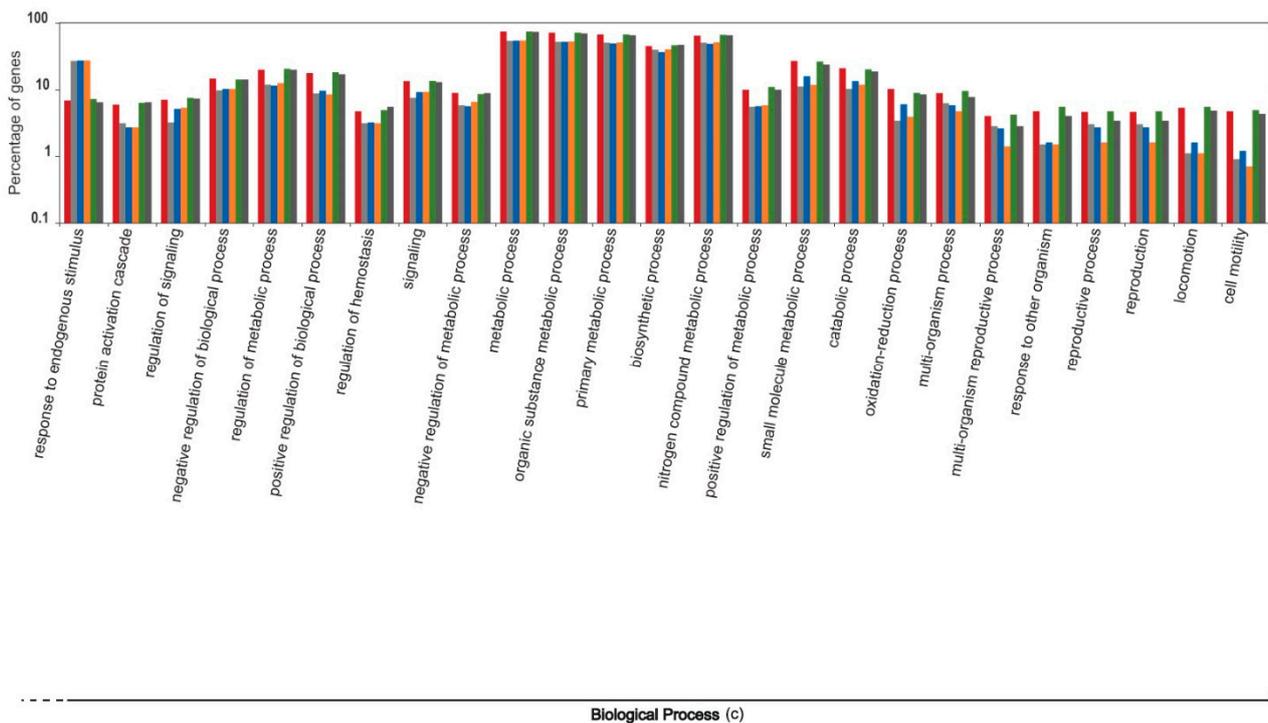


Figure 8. Intra-individual variation in the functional content of the top one percent of highly expressed transcripts within the liver transcriptome of South African sardines in terms of the Gene Ontology categorizations “biological process”, shown for six individuals that participated in the 2019 sardine run. Each individual is represented by a unique colour. Values on the Y axis are log transformed (continued from Figure 7).

Overexploitation of marine species, human-mediated climate change, and pollution threaten marine biodiversity globally. Marine species face an uncertain future, and their survival depends on their ability to adapt to this rapidly evolving environment. Many migratory species have already changed the route and timing of their migrations as a result of anthropogenic changes in their environment [57–59]. The long-term negative effects of human activities on the sardine run are yet to be fully understood. Mhlongo et al. [60] showed that an increase in sea surface temperature can negatively influence the spawning behaviour of sardines, which typically spawn at water temperatures between 16.0 °C and 22.0 °C. Other studies predict that the current rate of increase in sea surface temperature in EC and KZN could result in massive breeding failure, starvation, and, eventually, the complete disappearance of the sardine run [61,62].

Evaluating the functional content of an organism’s genome provides us with an in-depth understanding of the complex dynamics of local adaptations, which form an integral part of species’ adaptive responses to challenges that they will face in the near future. The present study provides a starting point for understanding the adaptive potential of South Africa’s sardine population to global change.

Supplementary Materials: The following are available online at <https://www.mdpi.com/2073-4425/12/3/368/s1>: Table S1: the number of quality-filtered sequences used for the liver assembly; Table S2: summary statistics of the liver transcriptome assembly; Table S3: gene symbols and full names of 1% highly expressed transcripts predicted from the liver transcriptome; Table S4: gene symbols and the full name of highly polymorphic transcripts (independent of their size) from the liver transcriptome.

Author Contributions: Conceptualization, P.R.T.; methodology, P.R.T., A.E.-K. and B.J.v.V.; validation, A.E.-K. and P.R.T.; formal analysis, A.E.-K., R.L.R., S.P.P., D.M.M., M.G.A. and C.M.S.; resources, P.R.T. and B.J.v.V.; writing—original draft preparation, A.E.-K., R.L.R., M.G.A., D.M.M., D.C.M., S.P.P., C.M.S., C.D.v.d.L., B.J.v.V. and P.R.T.; writing—review and editing, A.E.-K., R.L.R., M.G.A., D.M.M.,

D.C.M., S.P.P., C.M.S., C.D.v.d.L., B.J.v.V. and P.R.T.; funding acquisition, P.R.T., C.D.v.d.L. and B.J.v.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Research Foundation (NRF) CSUR Grant No. 87702 to P.R.T. and C.D.v.d.L., an FRC/URC grant from the University of Johannesburg to P.R.T., and DEFF funding to C.D.v.d.L.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Raw sequences for this study have been submitted to the NCBI Short Reads Archives (SRA) under Bio Project ID PRJNA701779. The assembled hepatic transcriptome is available in the Supplementary File sardine.fasta.

Acknowledgments: We would like to thank two anonymous reviewers for their constructive comments on the manuscript's initial version. A.E.K and D.M.M received bursaries from South Africa's National Antarctic Program (SANAP) grant No. 110728 to B.J.v.V. The computational resources for the current study were provided by the Center for High-Performance Computing (CHPC) in Cape Town, and the University of Johannesburg IT services.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Grantham, H.S.; Game, E.T.; Lombard, A.T.; Hobday, A.J.; Richardson, A.J.; Beckley, L.E.; Pressey, R.L.; Huggett, J.A.; Coetzee, J.C.; van der Lingen, C.D.; et al. Accommodating dynamic oceanographic processes and pelagic biodiversity in marine conservation planning. *PLoS ONE* **2011**, *6*. [[CrossRef](#)] [[PubMed](#)]
2. Hensor, E.; Couzin, I.D.; James, R.; Krause, J. Modelling density-dependent fish shoal distributions in the laboratory and field. *Oikos* **2005**, *110*, 344–352. [[CrossRef](#)]
3. Webster, M.M.; Goldsmith, J.; Ward, A.J.W.; Hart, P.J.B. Habitat-specific chemical cues influence association preferences and shoal cohesion in fish. *Behav. Ecol. Sociobiol.* **2007**, *62*, 273–280. [[CrossRef](#)]
4. Cattelan, S.; Griggio, M. Within-shoal phenotypic homogeneity overrides familiarity in a social fish. *Behav. Ecol. Sociobiol.* **2020**, *74*. [[CrossRef](#)]
5. van der Lingen, C.D.; Coetzee, J.C.; Hutchings, L. Overview of the KwaZulu-Natal sardine run. *Afr. J. Mar. Sci.* **2010**, *32*, 271–277. [[CrossRef](#)]
6. Hutchings, L.; Morris, T.; van der Lingen, C.D.; Lamberth, S.J.; Connell, A.D.; Taljaard, S.; van Niekerk, L. Ecosystem considerations of the kwaZulu-Natal sardine run. *Afr. J. Mar. Sci.* **2010**, *32*, 413–421. [[CrossRef](#)]
7. Myeza, J.; Mason, R.B.; Peddemors, V.M. Socio-economic implications of the kwaZulu-Natal sardine run for local indigenous communities. *Afr. J. Mar. Sci.* **2010**, *32*, 399–404. [[CrossRef](#)]
8. Andrew, S.C.; Primmer, C.R.; Debes, P.V.; Erkinaro, J.; Verta, J.P. The Atlantic salmon whole blood transcriptome and how it relates to major locus maturation genotypes and other tissues. *Mar. Genom.* **2020**, 100809. [[CrossRef](#)]
9. Ryu, T.; Kim, J.G.; Lee, J.; Yu, O.H.; Yum, S.; Kim, D.; Woo, S. First transcriptome assembly of a newly discovered vent mussel, *Gigantidas vrijenhoeki*, at Onnuri Vent Field on the northern Central Indian Ridge. *Mar. Genom.* **2020**, 100819. [[CrossRef](#)]
10. Pinosio, S.; Fratini, S.; Cannicci, S.; Oliva, M.; Pretti, C.; Baratti, M. De novo transcriptome assembly for *Pachygrapsus marmoratus*, an intertidal brachyuran crab. *Mar. Genom.* **2020**, 100792. [[CrossRef](#)]
11. Gnocchi, K.G.; Boldrini-França, J.; Gomes, L.C.; Chippari-Gomes, A.R. De novo assembly and annotation of the transcriptome of *Astyanax lacustris* liver unveil candidate genes to monitor response to environmental stress. *Mar. Genom.* **2020**, *54*, 100784. [[CrossRef](#)]
12. Emami-Khoyi, A.; Parbhu, S.P.; Ross, J.G.; Murphy, E.C.; Bothwell, J.; Monsanto, D.M.; van Vuuren, B.J.; Teske, P.R.; Paterson, A.M. De novo transcriptome assembly and annotation of liver and brain tissues of common brushtail possums (*Trichosurus vulpecula*) in New Zealand: Transcriptome diversity after decades of population control. *Genes* **2020**, *11*, 436. [[CrossRef](#)] [[PubMed](#)]
13. Cavallero, S.; Lombardo, F.; Salvemini, M.; Pizzarelli, A.; Cantacessi, C.; D'Amelio, S. Comparative transcriptomics reveals clues for differences in pathogenicity between *Hysterothylacium aduncum*, *Anisakis simplex sensu stricto* and *Anisakis pegreff*. *Genes* **2020**, *11*, 321. [[CrossRef](#)]
14. Tsehay, S.; Ortiz, R.; Johansson, E.; Bekele, E. New transcriptome-based SNP markers for Noug (*Guizotia abyssinica*) and their conversion to KASP markers for population genetics analyses. *Genes* **2020**, *11*, 1373. [[CrossRef](#)] [[PubMed](#)]
15. Benítez, R.; Trakooljul, N.; Núñez, Y.; Isabel, B.; Murani, E.; De Mercado, E.; Gómez-Izquierdo, E.; García-Casco, J.; López-Bote, C.; Wimmers, K.; et al. Breed, diet, and interaction effects on adipose tissue transcriptome in Iberian and Duroc pigs fed different energy sources. *Genes* **2019**, *10*, 589. [[CrossRef](#)]
16. van Dyk, J.C.; Pieterse, G.M.; van Vuren, J.H.J. Histological changes in the liver of *Oreochromis mossambicus* (Cichlidae) after exposure to cadmium and zinc. *Ecotoxicol. Environ. Saf.* **2007**, *66*, 432–440. [[CrossRef](#)]

17. Ren, M.; Li, J.; Xue, R.; Wang, Z.; Coll, S.L.; Meng, Q. Liver function and energy metabolism in hepatocellular carcinoma developed in patients with hepatitis B-related cirrhosis. *Medicine* **2019**, *98*, e15528. [[CrossRef](#)] [[PubMed](#)]
18. Sharma, A.; Singh, D.; Das, S.; Kumar, V. Hypothalamic and liver transcriptome from two crucial life-history stages in a migratory songbird. *Exp. Physiol.* **2018**, *103*, 559–569. [[CrossRef](#)]
19. Horton, W.J.; Jensen, M.; Sebastian, A.; Praul, C.A.; Albert, I.; Bartell, P.A. Transcriptome analyses of heart and liver reveal novel pathways for regulating songbird migration. *Sci. Rep.* **2019**, *9*, 1–13. [[CrossRef](#)]
20. Singh, D.; Swarup, V.; Le, H.; Kumar, V. Transcriptional Signatures in Liver Reveal Metabolic Adaptations to Seasons in Migratory Blackheaded Buntings. *Front. Physiol.* **2018**, *9*, 1–15. [[CrossRef](#)] [[PubMed](#)]
21. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)] [[PubMed](#)]
22. Haas, B.J.; Papanicolaou, A.; Yassour, M.; Grabherr, M.; Blood, P.D.; Bowden, J.; Couger, M.B.; Eccles, D.; Li, B.; Lieber, M.; et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **2013**, *8*, 1494–1512. [[CrossRef](#)]
23. Bryant, D.M.; Johnson, K.; DiTommaso, T.; Tickle, T.; Couger, M.B.; Payzin-Dogru, D.; Lee, T.J.; Leigh, N.D.; Kuo, T.H.; Davis, F.G.; et al. A tissue-mapped axolotl De Novo transcriptome enables identification of limb regeneration factors. *Cell Rep.* **2017**, *18*, 762–776. [[CrossRef](#)]
24. Finn, R.D.; Clements, J.; Eddy, S.R. HMMER web server: Interactive sequence similarity searching. *Nucleic Acids Res.* **2011**, *39*, 29–37. [[CrossRef](#)]
25. Krogh, A.; Larsson, B.; Von Heijne, G.; Sonnhammer, E.L.L. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J. Mol. Biol.* **2001**, *305*, 567–580. [[CrossRef](#)]
26. Lagesen, K.; Hallin, P.; Rødland, E.A.; Stærfeldt, H.H.; Rognes, T.; Ussery, D.W. RNAMmer: Consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* **2007**, *35*, 3100–3108. [[CrossRef](#)]
27. Seppely, M.; Manni, M.; Zdobnov, E.M. Busco: Assessing genome assembly and annotation completeness in gene prediction. In *Gene prediction*; Humana: New York, NY, USA, 2019; pp. 227–245.
28. Li, H. Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997.
29. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The Sequence Alignment/Map format and SAM tools. *Bioinformatics* **2009**, *25*, 2078–2079. [[CrossRef](#)]
30. Koboldt, D.C.; Zhang, Q.; Larson, D.E.; Shen, D.; McLellan, M.D.; Lin, L.; Miller, C.A.; Mardis, E.R.; Ding, L.; Wilson, R.K. VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequenc-ing. *Genome Res.* **2012**, *22*, 568–576. [[CrossRef](#)] [[PubMed](#)]
31. Li, B.; Dewey, C.N. RSEM: Accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinform.* **2011**, *12*, 323. [[CrossRef](#)] [[PubMed](#)]
32. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T. Gene Ontology: Tool for the Unification of Biology. *Nat. Genet* **2000**, *25*. [[CrossRef](#)] [[PubMed](#)]
33. Reimand, J.; Kull, M.; Peterson, H.; Hansen, J.; Vilo, J. G:Profiler—a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Res.* **2007**, *35*, 193–200. [[CrossRef](#)]
34. Ye, J.; Zhang, Y.; Cui, H.; Liu, J.; Wu, Y.; Cheng, Y.; Xu, H.; Huang, X.; Li, S.; Zhou, A. WEGO 2.0: A Web Tool for Analyzing and Plotting GO Annotations, 2018 Update. *Nucleic Acids Res.* **2018**, *46*, W71–W75. [[CrossRef](#)]
35. Supek, F.; Bošnjak, M.; Škunca, N.; Šmuc, T. Revigo summarizes and visualizes long lists of gene ontology terms. *PLoS ONE* **2011**, *6*. [[CrossRef](#)] [[PubMed](#)]
36. Kuznetsova, I.; Lugmayr, A.; Siira, S.J.; Rackham, O.; Filipovska, A. CirGO: An alternative circular way of visualizing gene ontology terms. *BMC Bioinform.* **2019**, *20*, 1–8. [[CrossRef](#)] [[PubMed](#)]
37. Strange, R.C.; Spiteri, M.A.; Ramachandran, S.; Fryer, A.A. Glutathione-S-transferase family of enzymes. *Mutat. Res. Mol. Mech. Mutagen.* **2001**, *482*, 21–26. [[CrossRef](#)]
38. Leaver, M.J.; Wright, J.; Hodgson, P.; Boukouvala, E.; George, S.G. Piscine UDP-glucuronosyltransferase 1B. *Aquat. Toxicol.* **2007**, *84*, 356–365. [[CrossRef](#)] [[PubMed](#)]
39. Bouzahouane, H.; Barour, C.; Sleimi, N.; Ouali, K. Multi-biomarkers approach to the assessment of the south-eastern Mediterranean Sea health status: Preliminary study on *Stramonita haemastoma* used as a bioindicator for metal contamination. *Chemosphere* **2018**, *207*, 725–741. [[CrossRef](#)]
40. Faheem, M.; Lone, K.P. Oxidative stress and histopathologic biomarkers of exposure to bisphenol-A in the freshwater fish, *Ctenopharyngodon idella*. *Braz. J. Pharm. Sci.* **2018**, *53*, 1–9. [[CrossRef](#)]
41. Samaraweera, A.V.; Sandamalika, W.M.G.; Liyanage, D.S.; Lee, S.; Priyathilaka, T.T.; Lee, J. Molecular characterization and functional analysis of glutathione S-transferase kappa 1 (*GSTκ1*) from the big belly seahorse (*Hippocampus abdominalis*): Elucidation of its involvement in innate immune responses. *Fish Shellfish Immunol.* **2019**, *92*, 356–366. [[CrossRef](#)]
42. Chen, B.; Peng, W.; Xu, J.; Feng, J.; Dong, C.; Xu, P.; Chen, B. Genomic Analysis of Glutathione S-transferases. *Pak. J. Zool.* **2017**, *49*, 1437–1448. [[CrossRef](#)]
43. Glisic, B.; Mihaljevic, I.; Popovic, M.; Zaja, R.; Loncar, J.; Fent, K.; Kovacevic, R.; Smital, T. Characterization of glutathione-S-transferases in zebrafish (*Danio rerio*). *Aquat. Toxicol.* **2015**, *158*, 50–62. [[CrossRef](#)]

44. Pérez-López, M.; Anglade, P.; Bec-Ferté, M.P.; Debrauwer, L.; Perdu, E.; Cravedi, J.P.; Rouimi, P. Characterization of hepatic and extrahepatic glutathione S-transferases in rainbow trout (*Oncorhynchus mykiss*) and their induction by 3,3',4,4'-tetrachlorobiphenyl. *Fish Physiol. Biochem.* **2000**, *22*, 21–32. [[CrossRef](#)]
45. Celius, T.; Walther, B.T. Oogenesis in Atlantic salmon (*Salmo salar*) occurs by zonagenesis preceding vitellogenesis in vivo and in vitro. *J. Endocrinol.* **1998**, *158*, 259–266. [[CrossRef](#)]
46. Conner, S.J.; Hughes, D.C. Analysis of fish ZP1/ZPB homologous genes—Evidence for both genome duplication and species-specific amplification models of evolution. *Reproduction* **2003**, *126*, 347–352. [[CrossRef](#)]
47. Li, J.; Chen, Z.; Gao, L.Y.; Colorni, A.; Ucko, M.; Fang, S.; Du, S.J. A transgenic zebrafish model for monitoring *xbp1* splicing and endoplasmic reticulum stress in vivo. *Mech. Dev.* **2015**, *137*, 33–44. [[CrossRef](#)]
48. Pacitti, D.; Lawan, M.M.; Feldmann, J.; Sweetman, J.; Wang, T.; Martin, S.A.M.; Secombes, C.J. Impact of selenium supplementation on fish antiviral responses: A whole transcriptomic analysis in rainbow trout (*Oncorhynchus mykiss*) fed supranutritional levels of Sel-Plex®. *BMC Genom.* **2016**, *17*, 1–27. [[CrossRef](#)] [[PubMed](#)]
49. Hung, I.C.; Hsiao, Y.C.; Sun, H.S.; Chen, T.M.; Lee, S.J. MicroRNAs regulate gene plasticity during cold shock in zebrafish larvae. *BMC Genom.* **2016**, *17*, 1–19. [[CrossRef](#)] [[PubMed](#)]
50. Horn, G.; Hofweber, R.; Kremer, W.; Kalbitzer, H.R. Structure and function of bacterial cold shock proteins. *Cell. Mol. Life Sci.* **2007**, *64*, 1457–1470. [[CrossRef](#)] [[PubMed](#)]
51. Chaikam, V.; Karlson, D.T. Comparison of structure, function and regulation of plant cold shock domain proteins to bacterial and animal cold shock domain proteins. *BMB Rep.* **2010**, *43*, 1–8. [[CrossRef](#)] [[PubMed](#)]
52. Roy, N.K.; Murphy, A.; Costa, M. Arsenic methyltransferase and methylation of inorganic arsenic. *Biomolecules* **2020**, *10*, 1351. [[CrossRef](#)] [[PubMed](#)]
53. Tierbach, A.; Groh, K.J.; Schönenberger, R.; Schirmer, K.; Suter, M.J.F. Glutathione S-Transferase protein expression in different life stages of zebrafish (*Danio rerio*). *Toxicol. Sci.* **2018**, *162*, 702–712. [[CrossRef](#)]
54. Micheletti, S.J.; Hess, J.E.; Zendt, J.S.; Narum, S.R. Selection at a genomic region of major effect is responsible for evolution of complex life histories in anadromous Steelhead. *BMC Evol. Biol.* **2018**, *18*, 1–11. [[CrossRef](#)]
55. Jindal, R.; Sinha, R.; Brar, P. Evaluating the protective efficacy of *Silybum marianum* against deltamethrin induced hepatotoxicity in piscine model. *Environ. Toxicol. Pharmacol.* **2019**, *66*, 62–68. [[CrossRef](#)]
56. Hygum, T.L.; Fobian, D.; Kamilari, M.; Jørgensen, A.; Schiøtt, M.; Grosell, M.; Møbjerg, N. Comparative investigation of copper tolerance and identification of putative tolerance related genes in tardigrades. *Front. Physiol.* **2017**, *8*. [[CrossRef](#)]
57. Le Corre, M.; Dussault, C.; Côté, S.D. Weather conditions and variation in timing of spring and fall migrations of migratory caribou. *J. Mammal.* **2016**, *98*, gyw177. [[CrossRef](#)]
58. Rakhimberdiev, E.; Duijns, S.; Karagicheva, J.; Camphuysen, C.J.; Dekinga, A.; Dekker, R.; Gavrilov, A.; ten Horn, J.; Jukema, J.; Saveliev, A.; et al. Fuelling conditions at staging sites can mitigate Arctic warming effects in a migratory bird. *Nat. Commun.* **2018**, *9*. [[CrossRef](#)]
59. Taylor, C.M.; Laughlin, A.J.; Hall, R.J. The response of migratory populations to phenological change: A Migratory Flow Network modelling approach. *J. Anim. Ecol.* **2016**, *85*, 648–659. [[CrossRef](#)]
60. Mhlongo, N.; Yemane, D.; Hendricks, M.; van der Lingen, C.D. Have the spawning habitat preferences of anchovy (*Engraulis encrasicolus*) and sardine (*Sardinops sagax*) in the southern Benguela changed in recent years. *Fish. Oceanogr.* **2015**, *24*, 1–14. [[CrossRef](#)]
61. Rouault, M.; Pohl, B.; Penven, P. Coastal oceanic climate change and variability from 1982 to 2009 around South Africa. *African J. Mar. Sci.* **2010**, *32*, 237–246. [[CrossRef](#)]
62. Mohanty, B.; Mohanty, S.; Sahoo, J. Climate Change: Impacts on Fisheries and Aquaculture. In *Climate Change and Variability*; Suzanne Simard, Ed.; Intech: Rijeka, Croatia, 2010; pp. 120–138.