Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A comprehensive RNA-seq dataset of embryonic and larval development in Senegalese sole (Solea senegalensis) across three thermal regimes

Lourdes Vega
UNIVERSIDAD DE CADIZ
Daniel Ramirez
UNIVERSIDAD DE CADIZ
Ismael Cross
UNIVERSIDAD DE CADIZ
Oscar Aramburu
Universidade de Santiago de Compostela
Dorinda Torres-Sabino
Universidade de Santiago de Compostela
Maialen Carballeda
Universidade de Santiago de Compostela
María Esther Rodríguez
UNIVERSIDAD DE CADIZ
Marco Mendizabal
UNIVERSIDAD DE CADIZ
Paulino Martínez
Universidade de Santiago de Compostela
Laureana Rebordinos
laureana.rebordinos@uca.es

UNIVERSIDAD DE CADIZ

data-descriptor

Keywords:

Posted Date: July 30th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10436543/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

This Data Descriptor presents a high-quality RNA-seq dataset profiling embryonic and larval development of Senegalese sole (*Solea senegalensis*) under three constant temperature regimes (14°C, 18°C and 21°C). A total of 48 samples were collected across eight developmental stages, from blastula to post-metamorphic larvae, and sequenced on the Illumina NovaSeq 6000 platform, generating approximately 30 million 150 bp paired-end reads per library. The dataset includes raw sequencing reads, gene-level count matrices and transcript abundance estimates (TPM), enabling comprehensive characterization of stage- and temperature-specific transcriptional dynamics. Technical validation confirms high data quality, with consistent sample clustering by developmental stage and biologically coherent differential expression patterns. This resource fills a critical gap in publicly available transcriptomic data for flatfish early development, providing a valuable foundation for investigating thermal plasticity, developmental regulation, and comparative flatfish biology, while supporting aquaculture research on temperature-driven phenotypic variation.

Background & Summary

Interest in the culture of *S. senegalensis* has grown considerably over the last decades, driven by its commercial value and adaptability to intensive and semi-intensive farming systems. Like other pleuronectiform fishes, its life cycle includes a complex developmental transition from a bilaterally symmetric pelagic larva to an asymmetric benthic juvenile, accompanied by extensive morphological and physiological reorganization [1–3]. These transformations involve extensive remodelling of sensory structures, body architecture, pigmentation patterns and feeding behaviour, ultimately establishing the benthic lifestyle characteristic of flatfishes.

Embryonic development progresses through blastula, gastrula and segmentation stages, during which cell proliferation, tissue differentiation and axis formation occur, followed by pre-hatching and larval phases characterized by increasing metabolic demands and organ maturation [2]. After hatching, larvae transition from endogenous to exogenous feeding while developing muscular, digestive and sensory functions. Subsequently, metamorphosis entails extensive anatomical remodelling, including eye migration, craniofacial restructuring, pigmentation asymmetry and acquisition of the asymmetric benthic morphology characteristic of flatfishes [1, 4].

Early developmental stages are among the most critical phases in flatfish aquaculture, frequently associated with elevated mortality, variable growth performance, skeletal abnormalities and metamorphic defects [3, 5]. Temperature has been identified as one of the major environmental factors influencing developmental rate, larval performance and gene expression patterns in Senegalese sole [6]. Previous transcriptomic studies have described molecular changes associated with thermal conditioning, skeletal development and metamorphosis, highlighting the involvement of developmental signalling pathways and metabolic regulation during early ontogeny [3, 5–6].

Over the last decade, genomic resources available for *S. senegalensis* have expanded considerably, including a chromosome-level genome assembly, studies on sex determination, and transcriptomic and epigenomic analyses of gonadal development. Nevertheless, publicly available transcriptomic datasets encompassing multiple developmental stages under controlled temperature conditions remain limited.

Here, we describe an RNA-seq dataset comprising 48 samples collected across eight developmental stages, spanning embryogenesis to post-metamorphic larvae, under three constant temperature conditions (14°C, 18°C and 21°C). Figure 1 provides an overview of the experimental design. By spanning the major embryonic and larval developmental stages under contrasting thermal conditions, this dataset provides a valuable resource for investigating gene expression dynamics during early ontogeny and supports future studies on developmental regulation, thermal plasticity, aquaculture and comparative flatfish biology.

Methods

1. Experimental design and sample collection

Sampling was conducted at the ECIMAT Marine Research Center (University of Vigo, Spain). Senegalese sole individuals were reared under controlled conditions and sampled across eight developmental stages, including six embryonic stages (blastula, gastrula, early segmentation, mid-segmentation, late segmentation and pre-hatching) and two larval stages (pre-metamorphic and post-metamorphic). Developmental stages were defined according to morphological criteria and standardized sampling time points expressed as hours post-fertilization (hpf) or days post-fertilization (dpf).

Individuals were reared under three constant temperature conditions (14°C, 18°C and 21°C). The experimental design followed a factorial structure including developmental stage and temperature as the main variables. Two biological replicates were collected for each developmental stage and temperature condition, resulting in a total of 48 RNA-seq libraries. Immediately after collection, samples were preserved in RNAlater™ (Thermo Fisher Scientific, USA), transported on ice to the University of Cádiz and stored at − 80°C until RNA extraction.

2. RNA extraction and sample processing

Samples were homogenized using a FastPrep-24™ 5G homogenizer (MP Biomedicals, USA), with three cycles of 20 s at 4.0 m s^{-1} with 1 min cooling on ice between cycles. Total RNA was extracted using the E.Z.N.A.® DNA/RNA Kit (Omega Bio-tek), following the manufacturer's instructions.

3. RNA quality control, library preparation and sequencing

RNA quantity and purity were assessed using a NanoDrop™ ND-1000 spectrophotometer (Thermo Fisher Scientific, USA). RNA integrity was evaluated using an Agilent 2100 Bioanalyzer with the RNA 6000 Nano

Kit (Agilent Technologies); only samples with an RNA Integrity Number (RIN) ≥ 7 were retained.

Library preparation and sequencing were performed by Novogene (Germany) following standard Illumina protocols on a NovaSeq 6000 platform (Illumina Inc., USA), generating 150 bp paired-end reads.

4. RNA-seq data processing and statistical analysis

Raw reads were processed using the nf-core/rnaseq pipeline (v3.10.1)[8]. Quality assessment was performed with FastQC [9], and adapter sequences together with low-quality bases were removed using Trimmomatic [10]. Clean reads were aligned to the *S. senegalensis* reference genome [7] using STAR [11]. Transcript abundance was estimated with RSEM [12], generating raw counts and TPM values.

All downstream analyses were conducted in R (v4.3.0)[13]. Gene-level count matrices were imported using tximport [14]. To define the active transcriptome, genes with TPM > 5 in at least two biological replicates were retained. Sample consistency was evaluated through Pearson and Spearman correlation analyses and principal component analyses (PCA). Differential expression analyses were performed using DESeq2 [15], with developmental stage and temperature as factors using the 18°C condition as the reference group. Genes with FDR < 0.05 were considered significantly differentially expressed.

Data visualization, including PCA plots, heatmaps, volcano plots and enrichment analyses, was performed using custom scripts in R, primarily based on ggplot2 [16] and associated packages.

5. Gene Ontology enrichment analysis

Gene Ontology (GO) enrichment analyses were performed across the Biological Process (BP), Molecular Function (MF), and Cellular Component (CC) ontologies using the R package clusterProfiler (v4.8.3)[17]. Differentially expressed genes were separated into up-regulated and down-regulated gene sets for each developmental stage and temperature comparison. Annotated *S. senegalensis* transcripts were mapped to *Danio rerio* gene identifiers using the org.Dr.eg.db annotation package (v3.17.0)[18]. Enrichment analysis was performed independently for each gene set using the enrichGO function. Statistical significance was assessed using the Benjamini–Hochberg procedure and GO terms with an adjusted P value (FDR) < 0.05 were considered significantly enriched. For visualization purposes, the five most significantly enriched GO terms identified for each comparison and direction of differential expression were retained, whereas complete enrichment results are provided in the supplementary material.

6. KEGG pathway enrichment analysis

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed using the R package clusterProfiler (v4.8.3)[17]. The same zebrafish orthologue mapping strategy described for GO enrichment was applied. KEGG pathway annotations were retrieved for *D. rerio* (organism code: dre)[19]. Differentially expressed genes were separated into up-regulated and down-regulated gene sets

for each developmental stage and temperature comparison. Annotated *S. senegalensis* transcripts were mapped to *D. rerio* gene identifiers using the org.Dr.eg.db annotation package (v3.17.0)[18]. Enrichment analysis was performed independently for each gene set using the enrichKEGG function with *D. rerio* as the reference organism. Multiple testing correction was performed using the Benjamini–Hochberg procedure, and pathways with an adjusted P value (FDR) < 0.05 were considered significantly enriched. For visualization purposes, the five most significantly enriched KEGG pathways identified for each comparison and direction of differential expression were retained, whereas complete enrichment results are provided in the supplementary material.

Data Records

The processed sequencing data supporting this Data Descriptor are available through the Figshare repository [https://figshare.com/s/cd311b80cebd16cc8927; DOI: 10.6084/m9.figshare.33040091]. The deposited files include quality-trimmed sequencing reads, gene-level expression quantification data, including raw counts and transcript abundance estimates expressed as transcripts per million (TPM), and differential gene expression results.

Additional supporting data are provided as Supplementary Information accompanying this Data Descriptor (Table 1). These include sequencing quality metrics, summaries of differential expression analyses, filtered and candidate gene sets, complete Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment results, and sample metadata describing developmental stage, temperature condition, and biological replicate.

Table 1
Summary of supplementary files accompanying the Data Descriptor.

Supplementary file	Description
Supplementary Table S1	RNA-seq quality metrics across developmental stages. Values are presented as mean ± standard deviation (SD) for libraries corresponding to each developmental stage.
Supplementary Table S2	Summary of differential expression analyses across developmental stages. Values correspond to the number of significant differentially expressed genes (DEGs) and the mean absolute $\log_2$ fold-change ($|\log_2 FC| \pm SD$) identified in the 14°C vs 18°C and 21°C vs 18°C comparisons. Differential expression analyses were performed independently for each developmental stage using DESeq2, with the 18°C treatment as the reference condition. Significant DEGs were defined using a false discovery rate (FDR) threshold of < 0.05.
Supplementary Table S3	Summary of filtered differentially expressed genes (DEGs) and candidate genes retained for downstream functional analyses across developmental stages and temperature comparisons. Candidate genes were selected according to annotation quality and expression magnitude ($|\log_2 FC| \geq 1.5$). For comparisons with more than 60 annotated DEGs, the 30 most up-regulated and 30 most down-regulated genes were retained.
Supplementary Table S4	Complete Gene Ontology (GO) enrichment results for temperature-responsive genes across developmental stages and temperature comparisons. Differentially expressed genes were analysed separately as up-regulated and down-regulated gene sets following mapping to *D. rerio* gene annotations and analysed using clusterProfiler. Enriched GO terms are classified according to the three Gene Ontology domains: Biological Process (BP), Molecular Function (MF), and Cellular Component (CC). For each comparison and direction of differential expression, the five most significantly enriched terms from each ontology are reported together with the GeneRatio, number of associated genes (Count), Benjamini–Hochberg adjusted P value (FDR), and the contributing genes.
Supplementary Table S5	Complete Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment results for temperature-responsive genes across developmental stages and temperature comparisons. Differentially expressed genes were analysed separately as up-regulated and down-regulated gene sets following mapping to *D. rerio* gene annotations and analysed using clusterProfiler. For each comparison and direction of differential expression, the five most significantly enriched KEGG pathways are reported together with the GeneRatio, number of associated genes (Count), Benjamini–Hochberg adjusted P value (FDR), and the genes contributing to each enriched pathway.

Technical Validation

1. Sequencing quality assessment

RNA sequencing generated between approximately 29 and 37 million paired-end reads per library across developmental stages. Following quality trimming, less than 1% of bases were removed from all libraries, indicating high initial sequencing quality. GC content remained highly consistent among developmental stages (47–49%), while duplication levels ranged between 56% and 64%. Average read length remained close to the expected 150 bp after trimming. Overall, these quality metrics demonstrate the consistency

of the sequencing dataset and support its suitability for downstream transcriptomic analyses. Detailed sequencing statistics are provided in Supplementary Table S1.

2. Sample clustering and principal component analysis

Principal component analysis (PCA) of normalized gene expression data revealed that samples clustered primarily according to developmental stage (Fig. 2). Biological replicates grouped closely together within each developmental stage, supporting the overall consistency and reproducibility of the dataset. A clear developmental trajectory was observed from embryonic to larval stages, with early embryonic stages forming distinct clusters that progressively transitioned towards larval stages, reflecting extensive transcriptomic remodelling throughout ontogeny.

Additional PCA analyses performed independently for each developmental stage revealed varying degrees of separation among temperature treatments. Distinct clustering patterns according to temperature were observed during the blastula, gastrula, early segmentation and pre-metamorphic stages, suggesting that temperature contributed substantially to transcriptomic variation during these developmental periods. Similar trends were observed during the mid- and late-segmentation stages, although with greater overlap among temperature conditions. In contrast, pre-hatching and post-metamorphic stages displayed increased dispersion among samples, resulting in less defined clustering patterns. Despite this variability, samples remained associated with their corresponding developmental stage and temperature treatment. No obvious outlier samples were detected, and all RNA-seq libraries were retained for downstream analyses.

Pairwise Spearman correlation analyses were performed to further evaluate sample similarity across the dataset (Fig. 3). Hierarchical clustering revealed a clear organization of samples according to developmental stage, with biological replicates consistently clustering together. These results further support the overall consistency and quality of the RNA-seq dataset.

3. Temperature-associated differential gene expression

Differential expression analyses were performed independently for each developmental stage using the 18°C treatment as the reference condition. Comparisons between 14°C vs 18°C and 21°C vs 18°C revealed stage-specific transcriptional responses. The number of differentially expressed genes (DEGs) varied markedly across developmental stages (Supplementary Table S2), indicating substantial differences in transcriptomic responses to temperature throughout ontogeny.

Blastula exhibited the largest number of DEGs in both temperature comparisons (6,983 and 6,923 DEGs for 14°C vs 18°C and 21°C vs 18°C, respectively). DEG numbers generally decreased during gastrulation and segmentation, reaching their lowest values at the pre-hatching stage (335 and 242 DEGs, respectively). Although developmental stages with larger numbers of DEGs did not necessarily display

the largest expression changes, late segmentation and pre-hatching exhibited the highest mean absolute $\log_2$ fold-change values among all stages (Supplementary Table S2).

Differences between temperature treatments were also evident. Gastrula, early segmentation and late segmentation exhibited more DEGs in the 14°C vs 18°C comparison, whereas pre-metamorphic larvae displayed a substantially larger number of DEGs in the 21°C vs 18°C comparison. Mid-segmentation showed comparable numbers of DEGs under both temperature conditions.

Volcano plots provide a graphical overview of differential expression patterns across developmental stages and temperature comparisons (Fig. 4). Blastula displayed extensive transcriptional responses under both temperature treatments, whereas progressively fewer significant DEGs were detected during subsequent embryonic stages. The plots also illustrate the contrasting magnitude of transcriptional responses between temperature treatments across developmental stages, consistent with the DEG counts summarized in Supplementary Table S2.

4. Functional annotation of temperature-responsive genes

To focus on biologically interpretable responses, differentially expressed genes (DEGs) were filtered according to statistical significance (FDR < 0.05), fold-change magnitude ($|\log_2 FC| \geq 1.5$), and annotation quality. Genes lacking functional annotation or corresponding to poorly characterised loci were excluded. For each comparison, when more than 60 annotated DEGs remained after filtering, the 30 most up-regulated and 30 most down-regulated genes (ranked by $\log_2 FC$) were retained; otherwise, all filtered annotated DEGs were included. This strategy generated manageable candidate gene sets while preserving the strongest transcriptional responses (Supplementary Table S3).

Recurrent candidate genes were subsequently identified across developmental stages and temperature comparisons (Table 2). The most frequently detected gene was *cyp8b1*, which was identified in five independent contrasts, followed by *hsc70*, *hsp90aa1.2*, *hspa8b*, and *rbp2a*, each identified in four contrasts. *Gna15*, *hsp71*, and *slc43a2b* were identified in three contrasts, suggesting the existence of a subset of temperature-responsive genes shared across multiple developmental contexts.

Table 2
Recurrent candidate genes identified across developmental stages and temperature comparisons. Occurrences indicate the number of independent temperature contrasts in which each gene was retained following filtering and candidate gene selection. Gene descriptions correspond to annotations from the *Solea senegalensis* reference transcriptome.

Gene symbol	Occurrences	Annotated description
cyp8b1	5	cytochrome P450 family 8 subfamily B member 1
hsc70	4	heat shock cognate 70
hsp90aa1.2	4	heat shock protein 90, alpha (cytosolic), class A member 1, tandem duplicate 2
hspa8b	4	heat shock protein family A (Hsp70) member 8b
rbp2a	4	retinol binding protein 2a, cellular
gna15	3	guanine nucleotide-binding protein subunit alpha-15
hsp71	3	heat shock 70 kDa protein 1
slc43a2b	3	solute carrier family 43 member 2b

Functional enrichment analyses were subsequently performed to characterize the biological functions associated with the temperature-responsive genes identified at each developmental stage. Gene Ontology (GO) enrichment analyses were conducted separately for up-regulated and down-regulated genes in each temperature comparison following mapping to *D. rerio* gene annotations. Enrichment was evaluated across the three GO domains: Biological Process (BP), Molecular Function (MF), and Cellular Component (CC). A total of 278 significantly enriched GO terms were identified across all comparisons, comprising 106 Biological Process, 80 Molecular Function, and 92 Cellular Component enrichments. The global distribution of the most recurrent enriched GO Biological Process terms across developmental stages and expression directions is presented in Fig. 5, whereas the GO enrichment results for all three ontologies are provided in Supplementary Table S4.

Pathway enrichment analyses were subsequently performed to characterize the biological pathways associated with temperature-responsive genes across developmental stages. Differentially expressed genes were analysed separately as up-regulated and down-regulated gene sets following mapping to *D. rerio* gene annotations. A total of 129 significantly enriched KEGG pathways were identified across all developmental stages and temperature comparisons, comprising 69 enrichments associated with up-regulated genes and 60 associated with down-regulated genes. Among the most frequently enriched pathways were Cadherin signaling, Adherens junction, Tight junction, IgSF CAM signaling, TGF-beta signaling pathway, and Oxidative phosphorylation. The global distribution of the most recurrent enriched KEGG pathways across developmental stages and expression directions is presented in Fig. 6, whereas the KEGG enrichment results are provided in Supplementary Table S5.

Usage Notes

This dataset is suitable for a wide range of analyses, including:

- Characterization of stage- and temperature-specific gene expression dynamics.
- Identification of thermal response genes and pathways.
- Comparative studies with other flatfish or teleost species.
- Integration with epigenomic, proteomic, or metabolomic datasets.
- Development of biomarkers for thermal stress or developmental quality in aquaculture.

Declarations

All animal procedures were approved by the Ethical Committee (body authorized according to R.D. 53/2013) and with the authorization of the Xunta de Galicia regional government.

Acknowledgements

This research was funded by the Junta de Andalucía to the group BIO-219 and [FEDER CCMM 00014]. The project was co-funded by the Regional Ministry of University, Research and Innovation of the Junta de Andalucía and by the European Union through the Next Generation EU funds under the Recovery, Transformation and Resilience Plan. It was also funded by the MCIN/FEDER (PID2022-137821OB-C31) and by (ThinkInAzul) supported by “Ministerio de Ciencia e Innovación” and “Xunta de Galicia” with funding from European Union Next Generation EU (PRTR-C17.I1) and European Maritime and Fisheries Fund (2022-CP076). We thank the ECIMAT Marine Research Center (University of Vigo) for providing facilities and technical support during sampling.

Competing interests

The authors declare no competing interests.

Funding

acquisition: L.R., P.M.

These authors contributed equally: Lourdes Vega, Daniel Ramírez.

Author Contribution

• Conceptualization: Laureana Rebordinos (L.R.), Paulino Martínez (P.M.)• Methodology: L.R., P.M., Ismael Cross (I.C.)• Investigation: Lourdes Vega (L.V.), Dorinda Torres Sabino (D.T.), Oscar Aramburu (O.A.), Maialen Carballeda (M.C.), Marco Mendizabal (M.M.), María Esther Rodríguez (MER), I.C.• Formal

analysis: I.C., O.A., Daniel Ramírez (D.R.)• Writing – original draft: D.R., L.R., L.V.• Writing – review & editing: D.R., L.R., I.C., P.M.• Supervision: L.R., P.M.• Funding acquisition: L.R., P.M.These authors contributed equally: Lourdes Vega, Daniel Ramírez.

Acknowledgement

We thank the ECIMAT Marine Research Center (University of Vigo) for providing facilities and technical support during sampling.

Data Availability

The processed sequencing data supporting this Data Descriptor are available through the Figshare repository [https://figshare.com/s/cd311b80cebd16cc8927; DOI: 10.6084/m9.figshare.33040091]. The deposited files include quality-trimmed sequencing reads, gene-level expression quantification data, including raw counts and transcript abundance estimates expressed as transcripts per million (TPM), and differential gene expression results.Additional supporting data are provided as Supplementary Information accompanying this Data Descriptor (Table 1). These include sequencing quality metrics, summaries of differential expression analyses, filtered and candidate gene sets, complete Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment results, and sample metadata describing developmental stage, temperature condition, and biological replicate.

Code Availability

Data processing and statistical analyses were performed using publicly available software and R/Bioconductor packages as described in the Methods. Analyses were implemented following the standard workflows and functions provided in the corresponding package documentation. Software and package versions used in the analyses are specified in the Methods. No custom software or original algorithms were developed for this study.

References

1. Padrós, F., Villalta, M., Gisbert, E. & Estévez, A. Morphological and histological study of larval development of the Senegal sole (*Solea senegalensis*): an integrative study. J. Fish Biol. 79, 3–32 (2011).
2. Liu, L. et al. Early ontogenesis from embryo to juvenile in Senegalese sole (*Solea senegalensis*) under laboratory conditions. J. Fish Biol. 104, 1800–1812 (2024).
3. Mendizábal-Castillero, M. et al. Differential expression pattern of hox genes between larval development and metamorphosis of Senegalese sole (*Solea senegalensis*, Kaup 1858). Aquaculture 609, 742802 (2025).

4. Campinho, M. A. et al. A thyroid hormone regulated asymmetric responsive centre is correlated with eye migration during flatfish metamorphosis. Sci. Rep. 8, 12267 (2018).
5. Bouza, C. et al. A comprehensive coding and microRNA transcriptome of vertebral bone in postlarvae and juveniles of Senegalese sole (*Solea senegalensis*). Genomics 116, 110802 (2024).
6. Campos, C. et al. Thermal plasticity of the miRNA transcriptome during Senegalese sole development. BMC Genomics 15, 525 (2014).
7. de la Herrán, R. et al. A chromosome-level genome assembly enables the identification of the follicle-stimulating hormone receptor as the master sex-determining gene in the flatfish *Solea senegalensis*. Mol. Ecol. Resour. 23, 886–904 (2023).
8. Ewels, P. A. et al. The nf-core framework for community-curated bioinformatics pipelines. Nat. Biotechnol. 38, 276–278 (2020).
9. Andrews, S. FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics (2010).
10. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120 (2014).
11. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21 (2013).
12. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12, 323 (2011).
13. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2023).
14. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Research 4, 1521 (2015).
15. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550 (2014).
16. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York (2016).
17. Wu, T. et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation 2, 100141 (2021).
18. Carlson, M. org.Dr.eg.db: Genome wide annotation for Zebrafish. R package version 3.17.0 (2023).
19. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 28, 27–30 (2000).

Figures

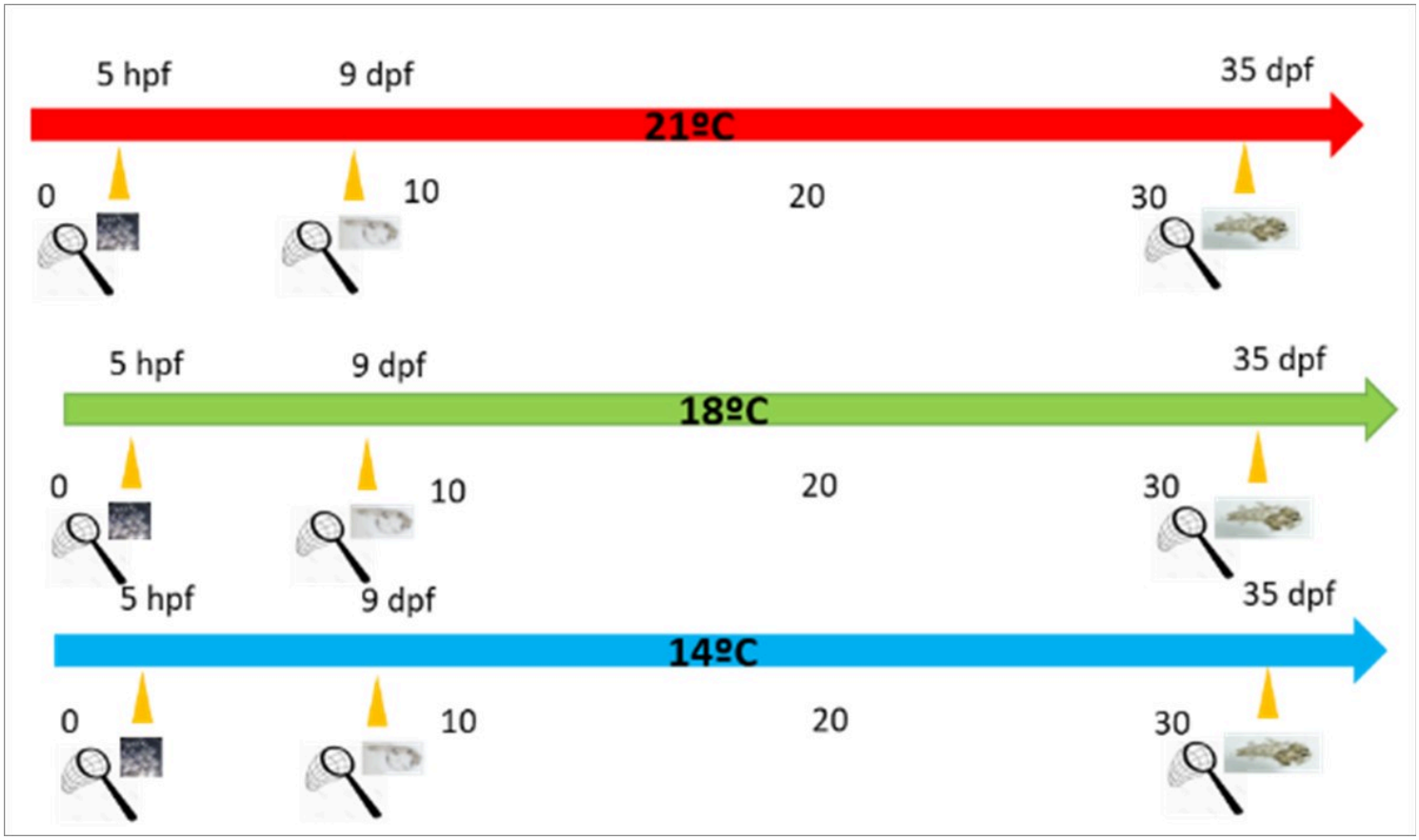


Figure 1

Overview of the experimental design. Senegalese sole (*Solea senegalensis*) individuals were sampled across eight developmental stages spanning embryogenesis, larval development and metamorphosis. Samples were obtained under three constant temperature conditions (14 °C, 18 °C and 21 °C). Developmental stages are expressed as hours post-fertilization (hpf) or days post-fertilization (dpf).

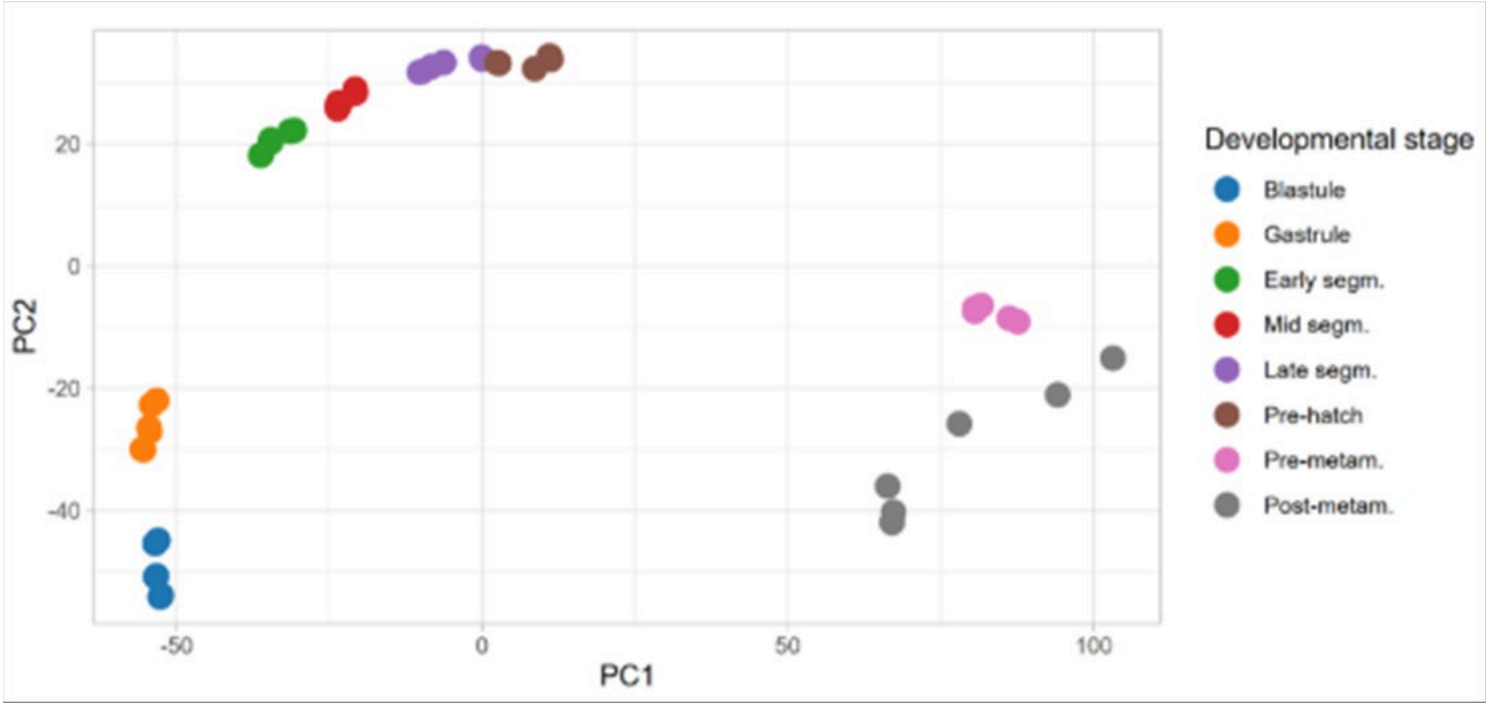


Figure 2

Principal component analysis (PCA) of normalized gene expression data from *Solea senegalensis*. PC1 and PC2 represent the first two principal components. Each point corresponds to an individual RNA-seq library coloured according to developmental stage. Samples cluster primarily according to developmental stage, illustrating the progressive transcriptomic changes associated with embryonic and larval development.

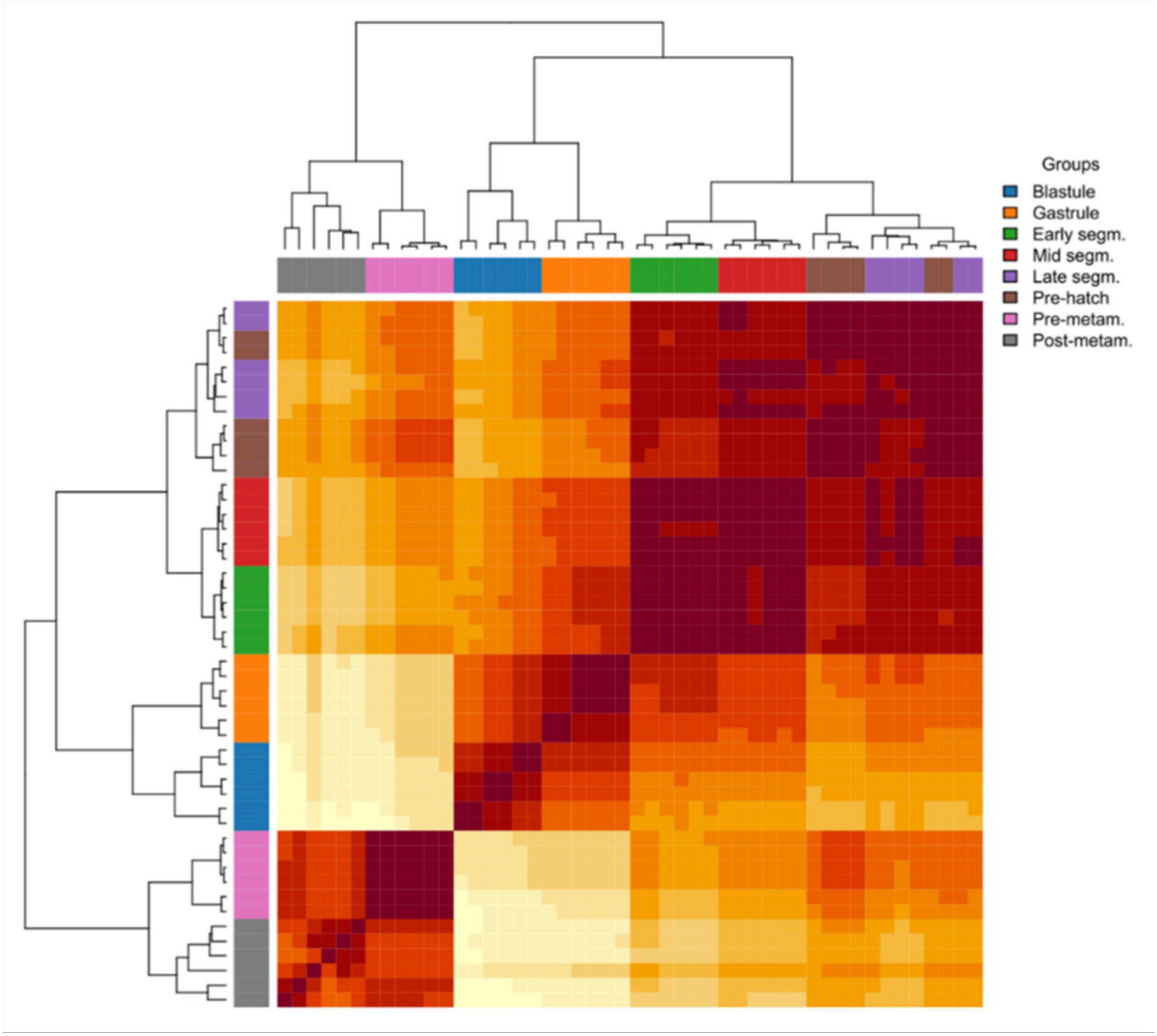


Figure 3

Spearman correlation heatmap of normalized gene expression profiles from *Solea senegalensis*. Rows and columns represent individual RNA-seq libraries. Colour intensity indicates pairwise Spearman correlation coefficients, while dendrograms represent hierarchical clustering based on transcriptomic similarity. Biological replicates cluster consistently within developmental stages, supporting the overall quality and reproducibility of the dataset.

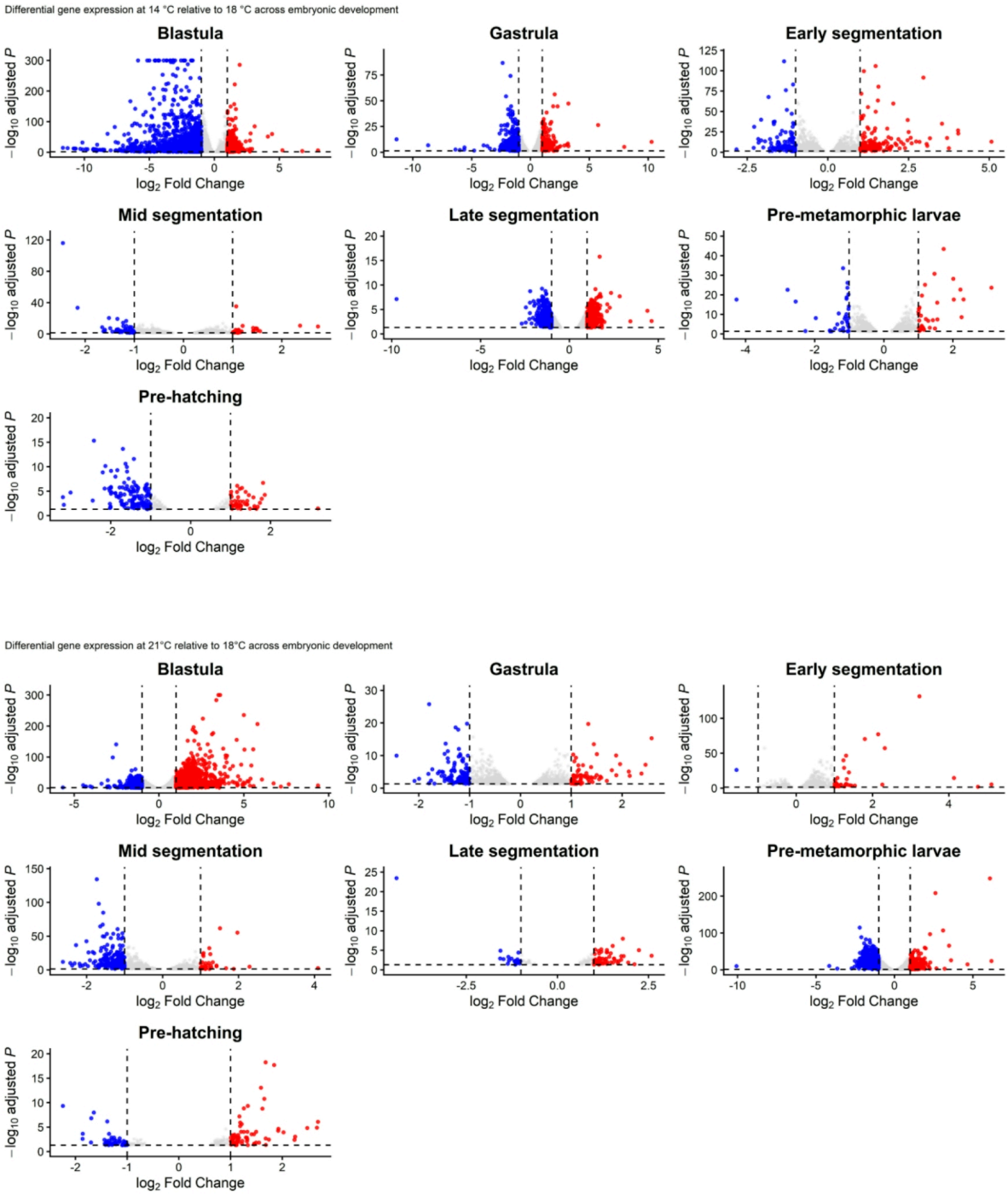


Figure 4

Volcano plots showing differential gene expression across developmental stages of *Solea senegalensis* under low- and high-temperature conditions relative to the 18 °C reference treatment. The upper panel corresponds to the 14 °C vs 18 °C comparison and the lower panel to the 21 °C vs 18 °C comparison. Each subplot represents an independent developmental stage. The x-axis shows $\log_2$ fold-change ($\log_2$FC), whereas the y-axis represents $-\log_{10}$ adjusted P values. Vertical dashed lines indicate the fold-

change threshold ($|\log_2 FC| \geq 1$), and the horizontal dashed line indicates the significance threshold (FDR < 0.05). Red and blue points represent significantly up-regulated and down-regulated genes, respectively, whereas grey points correspond to non-significant genes.

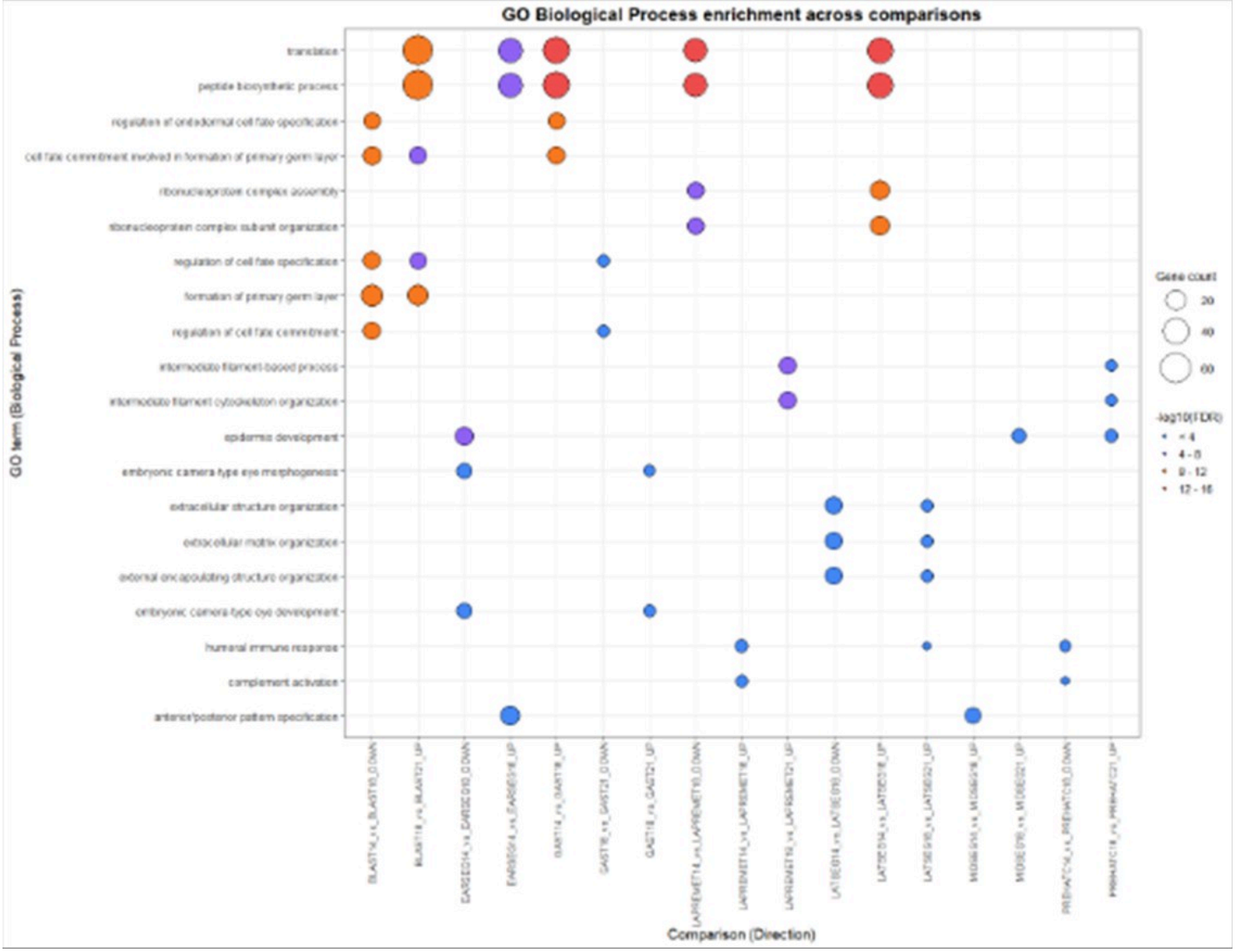


Figure 5

Global overview of Gene Ontology (GO) Biological Process enrichment across developmental stages and temperature comparisons. Differentially expressed genes were analysed separately as up-regulated and down-regulated gene sets following comparison against the 18 °C reference treatment. Each bubble represents one significantly enriched GO Biological Process term. Bubble size is proportional to the number of genes associated with each GO term (gene count), whereas colour indicates enrichment significance expressed as −log10 adjusted P value (FDR). Only the most recurrent significantly enriched Biological Process terms identified across comparisons are displayed.



Figure 6

Global overview of Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment across developmental stages and temperature comparisons. Differentially expressed genes were analysed separately as up-regulated and down-regulated gene sets following comparison against the 18 °C reference treatment. Each bubble represents one significantly enriched KEGG pathway. Bubble size is proportional to the number of genes associated with each pathway (gene count), whereas colour indicates enrichment significance expressed as −log10 adjusted *P* value (FDR). Only the most recurrent significantly enriched KEGG pathways identified across comparisons are displayed.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTableS1.xlsx

- SupplementaryTableS2.xlsx
- SupplementaryTableS3.xlsx
- SupplementaryTableS4.xlsx
- SupplementaryTableS5.xlsx