CHARACTERISING THE EARLY STAGES OF DIAPAUSE IN *DAPHNIA MAGNA:* A MULTI-OMICS APPROACH

by
ROSEMARY ELIZABETH BARNETT

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

School of Biosciences
University of Birmingham

December 2018

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

ABSTRACT

Diapause is a reversible developmental arrest in response to environmental challenges. The environmental genomics model organism *Daphnia* produces diapausing offspring. These dormant embryos produce healthy reproductive adults, even after 700 years of sustained developmental arrest. Despite this extraordinary adaptive feature of animal biology, prolonged diapause remained poorly understood. This thesis aims to develop *Daphnia magna* as a model system for embryonic diapause. I describe the reproducible induction of diapause in the laboratory and the development of embryos destined to diapause. My investigation reveals a delayed developmental trajectory compared to subitaneous embryos, until arrested at ca. 3500 cells. The collection and analysis of four omics datasets, including RNA sequencing paired with mass spectrometric data of the lipidome and metabolome, reveals distinct molecular differences during direct and diapause development. Furthermore, a significant developmental event is discovered by this molecular phenotyping of diapausing embryos at around 48 hours post-ovulation, when the molecular profiles are entirely different to pre-diapause, with changes manifesting later in the metabolome than the transcriptome. I identify additional molecular changes throughout the trajectory of diapause that are suggestive of molecular phenotypes describing the likely processes that produce an embryo that is capable of suspended animation for decades and even centuries.

This thesis is dedicated to Mr Peter Edis.

ACKNOWLEDGEMENTS

This research was carried out within the Environmental Genomics and Metabolomics laboratories in the School of Biosciences, University of Birmingham, UK and the Department of Animal Ecology, Evolution and Biodiversity, Ruhr-University Bochum, Bochum, Germany. Funding for this project was provided by Deutscher Akademischer Austasuschdienst Stipendien- und Betreuungsprogramm and a Midlands Integrative Biosciences Training Partnership studentship funded by the Biotechnology and Biological Sciences Research Council. Metabolomics research was supported by the UK Natural Environment Research Council Biomolecular Analysis Facility at the University of Birmingham. Transcriptomic research was supported by the Birmingham-Illinois Partnership for Discovery, Engagement and Education Project Seed Fund.

First and foremost I extend my sincerest gratitude to my remarkable supervisor **John Colbourne,** for having given me the opportunity to undertake a PhD in the first place, for the extensive amount of emotional and academic support throughout and for the infectious enthusiasm that meant I never stopped believing how awesome science really is. I am also grateful for my co-supervisor **Mark Viant**, who first captured my interest in omics technologies and has bestowed so much academic and professional guidance since then.

I would also like to thank **Linda Weiss**, who is a truly dedicated and inspirational researcher and a very good friend to me. Without her assistance and mentorship, a large portion of this research would be entirely guesswork. I would also like to thank **Luisa Orsini**, **Scott Hayward** and **Rocío Goméz** who in their individual capacities have contributed enormously to the direction and scope of this research.

I would like to acknowledge the significant contributions of my other colleagues. **Farhat** for always having a hug and some good advice. **Steve** for his endless technical assistance, music and movie suggestions. **Vignesh**, who saved the day in the eleventh hour. **Ian**, occasionally the only person I spoke to between long hours in the basement. And the wonderful **Luxi**, for spending many patient hours with me in dark rooms, and who along with everyone else at RUB made Germany feel like home.

More personally I would like to thank everyone, too numerous to name individually, that I have had the pleasure of sharing my life with over the last four years. Special mentions however are required for **Maria** who bought so many smiles and **Elliott** for patiently listening to my endless rants. **Martin** – you will always have one more friend in the world than you usually give yourself credit for (that's me by the way).

To my darling **Lauren**, thank you for saving my life, for being my best friend, for your unconditional love and respect. Stay fabulous.

Thank you to **my parents** who gave me a childhood filled with hydraulics and quadrats that ultimately meant I had to be a scientist, if only to keep up with the

dinner table conversation. Also, to the wonderful **Deb**, you will always be that one person I admire. Without the home away from home that you, **Oscar**, **Theo**, **Mark** and **Holly** have provided I probably would have given up a long time ago. And to **Carl**, **Saskia**, **Kate** and **Max** who also make my heart burst with family wishy-washiness.

For my big brother, I find that after four years I only have a (rather dull) book to show for myself and in that time, you have built a wonderful family, forged a successful career, fought and won the hardest of battles. I cannot begin to express how grateful I am that I have a man such as you walk ahead of me in life.

And lastly **Dave**: I love you.

TABLE OF CONTENTS

1. Introduction 1
1.1. Diapause: an alternative developmental trajectory 1
1.2. Diapause: current understanding and continuing ambiguity 6
1.2.1. Mechanisms of developmental arrest during diapause 9
1.2.2. Metabolic suppression during diapause 10
1.2.3. Mechanisms of stress tolerance during diapause 13
1.2.4. Molecular mechanisms of diapause 14
1.3. The reproductive biology of *Daphnia* 18
1.4. The developmental biology of *Daphnia* 22
1.5. *Daphnia magna* as a model system for embryonic diapause 24
1.6. Environmental omics approaches: overview and techniques 26
1.7. Thesis outline and its aims 33

2. Materials and methods 34
2.1. *Daphnia magna* isolate information 34
2.2. General culturing conditions 34
2.3. Culturing for the production of diapausing embryos 35
2.4. Culturing and preparation of *Chlorella vulgaris* algal feed 35
2.5. Selection of isolate and culturing parameters for the laboratory induction of diapause in *Daphnia magna* 37
2.6. Assessment of the longevity of laboratory-derived diapausing embryos 38
2.7. Assessment of the viability of laboratory-derived diapausing embryos 38
2.8. Temporal definition of the major morphological landmarks of embryonic diapause in *Daphnia magna* 39
2.9. Estimation of *Daphnia magna* sexual embryo dry mass 40
2.10. Collection of diapausing embryos for biomass optimisation of nESI-DIMS metabolomics 40
2.11. Identification of cellular diapause arrest and alternative developmental trajectories 41
2.12. Collection of embryos for nESI-DIMS and RNA-Seq analysis 43
2.13. Transcriptome sequencing 45
2.13.1. RNA extraction, library preparation and sequencing of *Daphnia magna* embryos 45
2.13.2. Illumina data processing 46
2.13.3. Differential gene expression analysis 47
2.13.4. *Daphnia magna* functional gene databases and gene ontology (GO) over-representation analysis 48
2.14. Direct infusion mass spectrometry (DIMS) metabolomics analysis 49

2.15. Statistical analysis of transcriptomic and metabolomic data 53
2.15.1. Univariate 53
2.15.2. Multivariate 53

3. Optimisation of protocols for the use of *Daphnia magna* as a model system for embryonic diapause 55
3.1. Introduction 55
3.2. Selection of isolate and culturing parameters for the laboratory induction of diapause in *Daphnia magna* 58
3.3. Assessment of the viability and longevity of laboratory-derived diapause offspring in *Daphnia magna* 66
3.4. Temporal definition of the major morphological landmarks of embryonic diapause in *Daphnia magna* 74
3.5. Biomass optimisation study for the application of nanoelectrospray ionisation-direct infusion mass spectrometry (nESI-DIMS) to diapausing embryos 82
3.6. Conclusions 87

4. Cellular and molecular investigations reveal alternative rates of development and points of departure in subitaneous and diapause-destined embryos 89
4.1. Introduction 89
4.2. Identification of the point of developmental arrest to demarcate the phases of diapause 92
4.3. Alternative developmental gene expression and metabolism 95
4.4. Identification of metabolic and transcriptomic differences between subitaneous and diapause-destined development 103
4.5. Discovering the function of diapause associated genes 107
4.6. Discovering the function of diapause associated metabolites 113
4.6.1. Discussion of the non-polar metabolite differences between diapause-destined and subitaneous *Daphnia magna* embryos 121
4.6.2. Discussion of the polar metabolite differences between diapause-destined and subitaneous *Daphnia magna* embryos 124
4.7. Conclusions 130

5. Molecular consequences of diapause initiation in *D. magna* 134
5.1. Introduction 134
5.2. Results and Discussion 136
5.2.1. Interrogating the transcriptome 138
5.2.2. Interrogating the metabolome 139
5.2.3. Molecular signatures of pre- and post-onset diapausing embryos 139
5.2.4. Putative significantly changed metabolites 143
5.2.5. Putative significantly expressed genes 157
5.3. Conclusions 168

6. Molecular profiling of embryonic diapause in *D. magna* 172
6.1. Introduction 172
6.2. Results and discussion 176
6.2.1. *Comparative analysis of the molecular modalities* 176
6.2.2. The molecular phenotype of diapause 204
6.3. Conclusions 217
6.3.1. The molecular phenotypes of diapause 217
7. Conclusions and future perspectives 222
8. References 230
9. Appendix A 255
9.1. Summary of DAPI-stained nuclei count data describing the development of sexually and asexually produced *Daphnia magna* embryos 255
9.2. Summary of sequencing metrics, data processing and quality assessment of results from the RNA-sequencing of sexually and asexually produced *Daphnia magna* embryos 257
9.2.1. Removal of low-expression genes and data normalisation 259
9.2.2. Assessment and removal of outlying samples 260
9.3. Summary of metabolite data processing and sample peak composition 262
9.3.1. Pairwise comparisons of subitaneous and diapause developmental *Daphnia magna* embryo samples reveals distinct peak compositions at pre-diapause and diapause 267
9.3.2. Alternative data processing reveals metabolic features unique to diapause and subitaneous development 275
9.4. Leverage/SPE plots of ANOVA-simultaneous component analysis (ASCA) variable sub-models for lipidomic and metabolomic datasets 279
9.5. Gene ontology annotation and overrepresentation analysis of genes differentially expressed in *Daphnia magna* sexual and asexual embryos 2–48 h post-ovulation 280
9.6. Investigation of genes related to glutathione 297
10. Appendix B 300
10.1. Metabolite fold changes at 24 h versus 10 h and 48 h versus 24 h 300
10.2. Investigation of metabolite peaks that are significantly changing around the onset of diapause developmental arrest that are unique only to diapause trajectory dataset 305
10.3. Investigation of the role of trehalose in the diapausing embryos of *Daphnia magna* 308

10.4. Gene ontology annotation and overrepresentation analysis of genes differentially expressed in relation to the initiation of diapause in sexually produced embryos of *Daphnia magna*.....312
10.5. Investigation of the mRNA expression of putatively annotated genes within the 'response to stress' gene ontology term.....319
10.6. Investigation of putatively annotated superoxide dismutase genes in the sexually produced embryos of *Daphnia magna.*.....320
10.7. Investigation of genes annotated with 'oxidoreductase activity' and/or 'lipid metabolic process' gene ontology.....322
10.8. Investigation of genes down-regulated upon initiation of diapause (48 h versus 10 h post-ovulation) annotated within the gene ontology terms 'DNA replication' and 'cell cycle'.....330
11. Appendix C.....338
11.1. Sparse principal component analysis of RNA-Seq data obtained from sexually produced *Daphnia magna* embryos collected between 2 h and 408 h post-ovulation.....338
11.2. Sparse principal component analysis of nESI-DIMS data from negative ion analysis of non-polar extracts of *Daphnia magna* embryos collected between 2 h and 17,592 h (2y) post-ovulation.....342
11.3. Sparse principal component analysis of nESI-DIMS data from negative ion analysis of polar extracts of *Daphnia magna* embryos collected between 2 h and 17,592 h (2 y) post-ovulation.....347
11.4. Sparse principal component analysis of nESI-DIMS data from positive ion analysis of polar extracts of *Daphnia magna* embryos collected between 2 h and 17,592 h (2 y) post-ovulation.....352

LIST OF FIGURES

Figure 1-1 'The eco-physiological phases of diapause'. The alternative developmental trajectory of diapause encompasses three progressive phases; pre-diapause, diapause and post-diapause of variable length and duration according to environmental conditions. Key events include the induction of diapause, the initiation of diapause (as onset of developmental arrest) and the termination of diapause (resumption of development). The term "diapause intensity" indicates, relative to a given time point, the duration of developmental arrest under set environmental conditions and thus the propensity of an organism to terminate diapause. Potential developmental trajectories are also highlighted including resumption of direct development upon termination of diapause, with (a) or without (c) a period of post-diapause quiescence and failure to reinitiate development (b) resulting in death. This figure is reproduced from Koštál (2006).....2

Figure 1-2 The reproductive cycle of *Daphnia magna.* Under favourable conditions mature females reproduce by parthenogenesis (I) giving birth to free-swimming clonal female offspring (II). In response to environmental stress, clonal male offspring (III) are produced and mature females switch to sexual reproduction (V) producing haploid eggs encased in an ephippial structure that require fertilisation by adult males (IV). The egg-containing ephippia is released during ecdysis (VI) and undergoes a period of diapause before hatching. 21

Figure 2-1 Sampling scheme for the collection of *Daphnia magna* embryos for transcriptomic and metabolomic analysis. Time (h) post-ovulation of sampling time points are indicated, shaded boxes (refer to legend bottom right) indicate collection for metabolomics (blue) and transcriptomics (green). 44

Figure 3-1 The reproductive output of L7 and B2 isolates of *Daphnia magna* over 21 d under "full" (0.08 mgC Daphniid^{-1}) and "reduced" (0.03 mgC Daphniid^{-1}) *Chlorella vulgaris* food ration at a density of 20 *Daphnia* L^{-1} (L7, n = 4; B2, n = 3). Mean cumulative number of neonates (A) and ephippia (B) produced by total cultures (20 *Daphnia*) over 21 d. Error bars denote standard deviation. Proportion of total offspring produced by cultures over 21 d that were male (C). (D) Maternal age (d) at first release of respective offspring type.....59

Figure 3-2 The effect of adult male presence on viable ephippia production over 21 d under "full" (0.08 mgCDaphniid^{-1}) and "reduced" (0.03 mgCDaphniid^{-1}) *Chlorella vulgaris* food ration at a density of 20/25 *Daphnia* L^{-1}. A) Total number of ephippia produced is corrected for the number of adult females present in each culture, B) Proportion of viable ephippia produced (ephippia that contained at least one egg upon release were considered viable). 63

Figure 3-3 Viable ephippia production over 21 d under "full" (0.08 mgCDaphniid^{-1}; n = 4) and "reduced" (0.03 mgCDaphniid^{-1};n = 3) *Chlorella vulgaris* food ration at

a density of 62 *Daphnia* L^{-1}. A) Total number of ephippia produced, B) Proportion of ephippia produced that contained at least one egg upon release. 64
Figure 3-4 Longevity of male (M; n = 3) and female (F; n = 5) L7 *Daphnia magna* under “full” (0.08 mgC $Daphniid^{-1}$) and “reduced” (0.03 mgC $Daphniid^{-1}$) *Chlorella vulgaris* food ration. A) Mean proportion of surviving individuals in cultures of 5 *D. magna* (error bars denote standard deviation). B) Age at total culture mortality after all 5 culture individuals had expired. 67
Figure 3-5 Hatching rates and survival to reproductive maturity of laboratory-produced diapausing embryos of *Daphnia magna* from ephippia of ages 14 d, 30 d, 60 d, and 728 d. A) Proportion of neonates hatching over 30 d post decapsulation, B) relationship between ephippial age and total proportion of hatching after 30 d, C) total proportion of hatched offspring to reach sexual maturity, D) relationship between proportion of hatching to reach sexual maturity and age of ephippia. 71
Figure 3-6 The maternal production of *Daphnia magna* asexual (A-B) and sexual embryos (C-D) imaged every 12 h. Scale bar = 500 µm. White bars and annotation indicate time (h) post-ovulation when imaging was performed. White triangles indicate ovulation which was defined as 0 h. Annotations in red indicate morphological points of interest including; the onset of oogenesis (□), the position of oocytes within the maternal ovary are outlined (---), compound eye formation (>), increasing pigmentation of the maternal carapace(*), release of sexual embryos into maternal brood pouch (♦).....77
Figure 3-7 The embryonic development of *Daphnia magna* upon removal from the maternal brood pouch (A) of asexual (B) and sexual (C) offspring. White triangles indicate embryos immediately (< 15 min) following ovulation which was defined as 0 h. White bars and annotation indicate time (h) post-ovulation when imaging was performed. Scale bar = 500 µm. Annotations in red indicate morphological points of interest including the formation of antennae (□) and compound eye formation (>). 81
Figure 3-8 Number of peaks occurring in >80% of extract samples across embryo pools of different biomass. A) negative ion mode of non-polar fractions, B) positive ion mode of polar fractions. 84
Figure 3-9 Relative standard deviation of all peaks occurring in >80% of samples across embryos pools of different biomass. A) negative ion mode of non-polar fraction, B) positive ion mode of polar fraction. 85
Figure 4-1 Cell count growth curves of subitaneous and diapause-destined embryos of *Daphnia magna*. A) Number of cells (DAPI-stained nuclei) in embryos sampled post-ovulation, points indicate raw counts from single biological replicates (n ≥ 3), lines represent respective logistic regression models for embryo types of cell count over time; B) Cell division rate of embryos over time estimated from first derivative with respect to time of logistic regression models of cell count over time. 93

Figure 4-2 Principal component scores plots showing the first two components from PCA of log-transformed gene counts (A) and metabolite intensities (B-D) of diapause-destined (blue) and subitaneous (red) embryos of *Daphnia magna* collected between 2 h and ~2 y (17,592 h) post ovulation; B) Negative ion non-polar extracts, C) Negative ion polar extracts, D) Positive ion polar extracts. 99
Figure 4-3 Principal component scores plots showing the first two components from PCA of log-transformed gene counts (A) and metabolite intensities (B-D) of diapause-destined (blue) and subitaneous (red) embryos of *Daphnia magna* collected between 2 h and 48 h post ovulation; B) Negative ion non-polar extracts, C) Negative ion polar extracts, D) Positive ion polar extracts. 101
Figure 4-4 Gene ontology (GO) over-representation analysis of genes differentially expressed between subitaneous and diapause-destined embryos of *Daphnia magna*. 'Up-regulated' genes refer to those with higher expression in diapause-destined embryos and vice versa for 'down-regulated' genes. 109
Figure 4-5 Heat map and hierarchical clustering of genes within the over-represented (p-adjusted <0.05) translation gene ontology annotation that displayed differential expression in samples of subitaneous and diapause-destined *Daphnia magna* embryos collected between 2 h and 48 h post-ovulation. Row labels indicate PANTHER-assigned gene family, coloured column top panels indicate embryo type and time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale. 111
Figure 4-6 Mean intensity of peaks from A) negative ion metabolomics analysis of non-polar extracts, B) negative ion, C) positive ion metabolomics analysis of polar extracts of *Daphnia magna* embryos at different time (h) post-ovulation with significant differences in intensity between sexually produced diapause-destined (blue) and asexually produced subitaneous (red) embryos in ANOVA simultaneous component analysis (ASCA) with a leverage threshold of 0.9 and squared prediction error alpha level of 0.05. Error bars denote standard deviation. Putative identities of peaks were assigned are annotated (TG = triacylglycerides). 123
Figure 4-7 Mean intensity of putatively annotated glutathione metabolites analysed by negative ion nESI-DIMS metabolomics analysis in *Daphnia magna* subitaneous and diapause-destined embryos between 2 h and 48 h post-ovulation. Error bars denote standard deviation. Panel annotations indicate putatively assigned compound identities with corresponding ion form detected. 125
Figure 4-8 Heat map and hierarchical clustering of genes relating to glutathione that displayed differential expression (p-adjusted <0.01) in samples of subitaneous and diapause-destined *Daphnia magna* embryos collected between 2 h and 48 h post-ovulation and a high positive or negative correlation ($-0.9 > r > 0.9$) with glutathione metabolite levels. Row labels indicate OrthoDB-assigned gene name and coloured column top panels indicate embryo type and time post-ovulation

(refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale. 127
Figure 5-1 Volcano plot of differentially expressed genes (A) and metabolites (B-D) with significant intensity fold changes between pre-onset and onset of diapause (48 h v 10 h) and between onset of diapause and embryos in diapause (74 h v 48 h). Metabolomics analysis implemented nESI-DIMS analysis of non-polar extracts (B; negative ion mode) and polar extracts analysed in negative ion mode (C) and positive ion mode (D). Filled points represent $p < 0.01$ (after Benjamini-Hochberg false discovery rate correction), red points indicate $\log_2$ fold change ratio > 2. 142
Figure 5-2 Intensity of putatively annotated trehalose (red, top panel) and uridine diphosphate (UDP)-glucose (blue, lower panel) peaks measured utilising nanoelectrospray ionisation-direct infusion mass spectrometry analysis of polar (positive and negative ion analysis respectively) extracts of *Daphnia magna* diapausing embryos. Points indicate individual biological replicates, lines fitted with locally weighted scatterplot smoothing (LOESS; span = 0.5), grey shading indicate model 95% confidence interval. 153
Figure 5-3 Biosynthesis of trehalose from Miao *et al*., (2016). Uridine diphosphate (UDP)-glucose and glucose-6-phosphate are condensed to form trehalose-6-phosphate catalysed by the enzyme trehalose-6- phosphate synthase (TPS). Trehalose is formed in a final enzymatic step catalysed by trehalose-6-phosphate phosphatase (TPP). 154
Figure 5-4 Heat map and hierarchical clustering of genes relating to trehalose in samples of sexually produced *Daphnia* embryos collected between 2h and 2 weeks (408h) post-ovulation. Row labels indicate gene name with corresponding annotation based on OrthoDB orthologous gene family group and coloured column top panels indicate time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale. 155
Figure 5-5 Heat map and hierarchical clustering of putatively annotated superoxide dismutase genes in samples of sexually produced *Daphnia* embryos collected between 2 h and 2 weeks (408 h) post-ovulation that had significant (false discovery rate corrected $p < 0.01$) large $\log_2$ fold changes between 10 h and 4 h or 48 h and 74 h (Dapma7bEVm004779 only) post-ovulation. Row labels indicate gene name with corresponding annotation based on OrthoDB orthologous gene family group and coloured column top panels indicate time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale. 161
Figure 5-6 Gene ontology (GO) over-representation analysis of genes that are differentially expressed between 10 h and 48 h post-ovulation in sexually produced embryos of *Daphnia magna*. ‘Up-regulated’ genes refer to those with higher expression in 48 h embryos and vice versa for ‘down-regulated’ genes.

Each gene set (i.e. up- or down-regulated) was investigated for overrepresented annotations within each GO domain covering biological processes, cellular components and molecular functions. 163
Figure 5-7 Heatmap and hierarchical clustering of genes annotated with the gene ontology terms ‘cell cycle’ (A; 95 genes), ‘DNA replication’ (B; 21 genes), and shared between the two terms (C; 22 genes) that were significantly (false discovery rate corrected $p < 0.1$), differentially expressed between 10 h and 48 h in sexually produced embryos of *Daphnia magna.* Coloured column top panels indicate time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale. 166
Figure 6-1 Scores plots from the sparse principal component analysis (sPCA) of log-transformed gene counts (A) and metabolite intensities (B-D) of diapause-destined and diapausing embryos of *Daphnia magna* collected between 2 h and 2 y (~17592 h) post ovulation; Metabolomics analysis of B) Negative ion non-polar extracts, C) Negative ion polar extracts, D) Positive ion polar extracts. ..177
Figure 6-2 Sparse principal component analysis (sPCA) scores plots (1) and mean (log-normalised) expression of genes with non-zero loadings for associated sparse component (2) from sPCA of RNA-Seq data from sexually-produced *Daphnia magna* embryos collected between 2 h and 408 h (2 weeks) post-ovulation. Sparse components 1–4 (A–D) are shown (refer to Figure 6-1 legend for point characters). For expression plots, coloured lines indicate individual genes (mean expression at each time point calculated across 2-3 biological replicates), vertical dashed lines indicate approximate onset of developmental arrest (48 h). 182
Figure 6-3 Heat map and hierarchical clustering of expression of genes with non-zero loadings in the third sparse principal component from the sparse principal component analysis of log-normalised gene counts in *Daphnia magna* sexually produced embryos collected between 2 h and 408 h post-ovulation. Log-normalised counts in subitaneous embryos are shown for comparison only. Row labels indicate gene name with corresponding annotation based on OrthoDB orthologous gene family group and coloured column top panels indicate embryo type and time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale. 183
Figure 6-4 Scores plots (1) and mean log-normalised intensity of peaks with non-zero loadings for associated component (2) from sparse principal component analysis of log-normalised peak intensity from negative ion mode metabolomics analysis of non-polar extracts of sexual *Daphnia magna* embryos collected between 2 h and 2 y (17592 h) post-ovulation. Sparse principal components 1 (A) through to 4 (D) are shown. Vertical dashed line indicates approximate onset of developmental arrest (48 h). Coloured lines indicate mean intensity of individual peaks with time post ovulation. 187

Figure 6-5 Intensity of peaks with a non-zero loading in the third sparse principal component identified through sparse principal component analysis of log-normalised peak intensity matrix obtaining follow the negative ion metabolomics analysis of non-polar *Daphnia magna* embryo extracts collected between 2 h and 2 y (17592 h) post-ovulation that displayed a general increase (A) and decrease (B) in intensity with time post-ovulation. Individual points represent biological replicates. Lines are fitted with locally weighted scatterplot smoothing (LOESS; span = 0.3). Grey shading indicates model 95% confidence interval. 188
Figure 6-6 Intensity of peaks with a non-zero loading in the fourth sparse principal component identified through sparse principal component analysis of log-normalised peak intensity matrix obtaining follow the negative ion metabolomics analysis of non-polar *Daphnia magna* embryo extracts collected between 2 h and 2 y (17592 h) post-ovulation that displayed a significant increase (A) and decrease (B) in intensity between 408 h and 17592 h post-ovulation. Peaks with putatively assigned compound identities are shown in legend boxes. Individual points represent biological replicates. Lines are fitted with locally weighted scatterplot smoothing (LOESS; span = 0.3). Grey shading indicates model 95% confidence interval. 190
Figure 6-7 Scores plots (1) and mean log-normalised intensity of peaks with non-zero loadings for associated component (2) from sparse principal component analysis of log-normalised peak intensity from negative ion mode analysis of polar extracts of sexual *Daphnia magna* embryos collected between 2 h and 2 y (17592 h) post-ovulation. Sparse principal components 1 (A) through to 4 (D) are shown. Vertical dashed line indicates approximate onset of developmental arrest (48 h). Coloured lines indicate mean intensity of individual peaks. 194
Figure 6-8 Scores plots (1) and mean log-normalised intensity of peaks with non-zero loadings for associated component (2) from sparse principal component analysis of log-normalised peak intensity from positive ion mode analysis of polar extracts of sexual *Daphnia magna* embryos collected between 2 h and 2 y (17592 h) post-ovulation. Sparse principal components 1 (A) through to 4 (D) are shown. Vertical dashed line indicates approximate onset of developmental arrest (48 h). Coloured lines indicate mean intensity of individual peaks.....195
Figure 6-9 Intensity of peaks with a non-zero loadings in the first sparse principal component identified through sparse principal component analysis of log-normalised peak intensity matrix obtaining follow the negative ion analysis of polar *Daphnia magna* embryo extracts collected between 2 h and 2 y (17592 h) (post-ovulation that displayed a general decrease in intensity. Individual points represent biological replicates. Lines are fitted with locally weighted scatterplot smoothing (LOESS; span = 0.3). Grey shading indicates the modelled 95% confidence intervals. 197

Figure 6-10 Intensity of peaks with a non-zero loadings in the second sparse principal component that demonstrated a prominent spike in intensity at 74 h post-ovulation. These are identified through sparse principal component analysis of log-normalised peak intensity matrix obtaining from the negative ion (A) and positive ion (B) metabolomics analysis of polar *Daphnia magna* embryo extracts collected between 2 h and 2 y (17592 h) post-ovulation. Additional titles indicate putative peak annotations. Individual points represent biological replicates. Lines are fitted with locally weighted scatterplot smoothing (LOESS; span = 0.3). Grey shading indicates the modelled 95% confidence intervals. 200
Figure 6-11 Intensity of putatively annotated dipeptide peaks with a non-zero loadings in the third sparse principal component that demonstrated a prominent spike in intensity at 48 h post-ovulation. These are identified through sparse principal component analysis of log-normalised peak intensity matrix obtaining from the positive ion analysis of polar *Daphnia magna* embryo extracts collected between 2 h and 17592 h (2 y) post-ovulation. Individual points represent biological replicates. Lines are fitted with locally weighted scatterplot smoothing (LOESS; span = 0.3). Grey shading indicates the modelled 95% confidence intervals.....202
Figure 6-12 The molecular phenotype of diapause anchored by the rate of cell division through to developmental arrest. Main panel shows the stages of diapause in *Daphnia magna* as defined within this thesis based on the arrest of active development determined from the counting of DAPI-stained nuclei of sexually produced embryos at various time points post-ovulation. This encompasses pre-diapause (blue shading) wherein direct development proceeds, the initiation of diapause at the onset of developmental arrest (ca. 48 h) followed by a period of diapause (red shading). The duration of diapause investigated was further divided into “early” diapause covering the immediate days following initiation and “late” covering a longer period of 2 y, indicated by horizontal black arrows. Sampling time points are indicated in the grey annotation panel for each chart and by vertical dashed lines on the main growth curve panel. Bar charts indicate the percentage of the total number of detected features that show a significant change in magnitude between the time points indicated. Significantly changing features were determined by $p < 0.01$ (after Benjamini-Hochberg false discovery rate correction). Actual counts are given in Table 6.5. 207
Figure 6-13 Plot of shared and unique differentially expressed genes identified in pairwise comparisons of successive sampling time points taken during early diapause (48 h–408 h). ‘Total’ genes refers to the total number within each time point that displayed a significant change in expression determined by $p < 0.01$ (after Benjamini-Hochberg false discovery rate correction). Empty intersections were no genes were shared between respective time points are not shown....212

Figure 6-14 Number differentially expressed genes (A) and metabolites (B-D) with significant intensity fold changes between the indicated time points. Blue fractions of bars indicate the number of features that demonstrate a significant change only within the given developmental time interval (i.e. unique) or within at least one additional time interval (red fraction). Significantly changing features were determined by $p < 0.01$ (after Benjamini-Hochberg false discovery rate correction). 213
Figure 6-15 Volcano plots of differentially expressed genes (A) and metabolites (B-D) with significant fold changes between *Daphnia magna* embryos sampled at successive time points during diapause development. Metabolite analysis implemented nESI-DIMS metabolomics analysis of non-polar extracts (B; negative ion mode) and polar extracts analysed in negative ion mode (C) and positive ion mode (D). Filled points represent $p < 0.01$ (after Benjamini-Hochberg false discovery rate correction), red points indicate $\log_2$ fold change ratio > 2. 215
Figure 6-16 Conceptual figure describing the molecular phenotypes of embryonic diapause as described within this thesis. Coloured lines represent distinct phenotype trajectories characterised by element amount (i.e. gene expression or metabolite intensity) with time in diapause development. Shaded plot areas indicated pre-diapause (blue) and diapause (red) phases, where vertical hashed black line indicates the initiation of diapause....219
Figure 9-1 Boxplots of mean gene expression (n = 3) for pre-normalisation (A) and normalised (B) data for each time pint post-ovulation collected from asexually and sexually produced *Daphnia magna* embryos. 259
Figure 9-2 Distribution of total gene counts ($\log_{10}$) across all biological samples (n = 30) A) before and B) after the removal of genes with <10 counts across all samples....260
Figure 9-3 Principal component scores plots showing the first two components from PCA of log-transformed gene counts, three samples circled in red (arbitrarily drawn) were determined to be outliers based on sample scores being three or more standard deviations away from group mean score....261
Figure 9-4 Heatmap showing the clustering of sample types and biological replicates based on Euclidean distance calculated on log-transformed gene counts, outlying samples that failed to cluster with other biological replicates of the same sample type are highlighted in red. 261
Figure 9-5 Upset plot of shared and unique genes and metabolites between diapause-destined and subitaneous embryo samples across all time points. A) Number of genes with >10 counts that occur in all biological replicates of diapause-destined and/or subitaneous samples; B-D) number of peaks that occur in at least 80% of all biological replicates of diapause-destined and/or subitaneous samples; B) Negative ion analysis of non-polar extracts; C)

Negative ion analysis of polar extracts; D) Positive ion analysis of polar extracts. 263
Figure 9-6 Peak characteristics of non-polar extracts of *Daphnia magna* embryos collected at various times post ovulation following nESI-DIMS analysis in negative ion mode and data processing to yield intensity matrices consisting of peaks present in at least 80% of biological samples (n = 8) within each biological sample class. A) Total number of peaks per class including number of unique (yellow) features. B) Percentage of shared features in pairwise comparisons between all sample types, values in cells indicate the percentage of total peaks present in sample row that are also present in given sample column. 268
Figure 9-7 Peak characteristics of polar extracts of *Daphnia magna* embryos collected at various times post ovulation following nESI-DIMS analysis in negative ion mode and data processing to yield intensity matrices consisting of peaks present in at least 80% of biological samples (n = 8) within each biological sample class. A) Total number of peaks per class including number of unique (yellow) features. B) Percentage of shared features in pairwise comparisons between all sample types, values in cells indicate the percentage of total peaks present in sample row that are also present in given sample column. 269
Figure 9-8 Peak characteristics of polar extracts of *Daphnia magna* embryos collected at various times post ovulation following nESI-DIMS analysis in positive ion mode and data processing to yield intensity matrices consisting of peaks present in at least 80% of biological samples (n = 8) within each biological sample class. A) Total number of peaks per class including number of unique (yellow) features. B) Percentage of shared features in pairwise comparisons between all sample types, values in cells indicate the percentage of total peaks present in sample row that are also present in given sample column. 270
Figure 9-9 Euler diagrams indicating the number of intersecting metabolite peaks between datasets processed including both subitaneous and diapause embryo samples (All) and diapause embryo samples only (Sexual) from the nano-electrospray ionisation direct infusion mass spectrometry analysis of non-polar extracts (A; negative ion mode) and polar extracts analysed in negative ion mode (B) and positive ion mode (C).....277
Figure 9-10 Leverage/SPE scatter plots of ANOVA-simultaneous component analysis (ASCA) variable submodels for (A) phenotype (embryo type), (B) time (h post-ovulation) and the (C) interactions thereof for analysis conducted in negative ion mode on polar extracts (.1) positive ion mode on polar extracts (.2) and negative ion analysis on non-polar extracts (.3). Permutation testing of sub-models for time and interactions were insignificant ($p > 0.05$) for analysis of non-polar extracts.....279
Figure 9-11 Heat map and hierarchical clustering of genes relating to glutathione that displayed differential expression (p-adjusted < 0.01) in samples of subitaneous and diapause-destined *Daphnia magna* embryos collected between 2 h and 48 h

post-ovulation. Row labels indicate gene name with corresponding annotation based on OrthoDB orthologous gene family group and coloured column top panels indicate time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale. 299
Figure 10-1 Volcano plots of metabolites with significant intensity fold changes between 24 h v10 h and 48 h v 24 h post-ovulation in relation to the onset of diapause (ca. 48 h) in embryos of *Daphnia magna.* Time points compared are indicated by surrounding boxes. Metabolite analysis implemented nESI-DIMS analysis of non-polar extracts (A; negative ion mode) and polar extracts analysed in negative ion mode (B) and positive ion mode (C). Filled points represent $p < 0.01$ (after Benjamini-Hochberg false discovery rate correction), red points indicate a $\log_2$ fold change ratio > 2.....301
Figure 10-2 Heat map and hierarchical clustering of genes relating to trehalose in samples of subitaneous and diapause-destined *Daphnia* embryos collected between 2 h and 2 weeks (408 h) post-ovulation. Row labels indicate gene name with corresponding annotation based on OrthoDB orthologous gene family group and coloured column top panels indicate time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale. 310
Figure 10-3 Intensity of putatively annotated trehalose peak identified utilising nESI-DIMS analysis of polar (positive ion analysis) extracts of *Daphnia magna* subitaneous (Asexual) and diapause-destined (Sexual) embryos. Points indicate individual biological replicates, lines fitted with locally weighted scatterplot smoothing (LOESS; span = 0.5), grey shading indicate model 95% confidence interval.....311
Figure 10-4 Heat map and hierarchical clustering of putatively annotated genes within the ‘response to stress’ gene ontology term in samples of subitaneous and diapause-destined *Daphnia magna* embryos collected between 2 h and 2 weeks (408 h) post-ovulation. Row labels indicate gene name with corresponding annotation based on OrthoDB orthologous gene family group and coloured column top panels indicate time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale. 319
Figure 10-5 Heat map and hierarchical clustering of expression of putatively annotated superoxide dismutase genes in samples of subitaneous and diapause-destined *Daphnia magna* embryos collected between 2 h and 2 weeks (408 h) post-ovulation. Row labels indicate gene name with corresponding annotation based on OrthoDB orthologous gene family group and coloured column top panels indicate time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale. 321
Figure 10-6 Heat map and hierarchical clustering of expression of genes assigned to the ‘lipid metabolic process’ gene ontology term that were significantly up-regulated in expression in samples of sexual *Daphnia magna* embryos collected

at 48 h versus 10 h post-ovulation. Column labels individual biological replicates (see key for embryo type and collection time). Colours indicate low expression (blue) and high expression (red) on log-normalised scale.....324
Figure 10-7 Heat map and hierarchical clustering of expression of genes assigned to the ‘oxidoreductase activity’ gene ontology term that were significantly up-regulated in expression in samples of sexual *Daphnia magna* embryos collected at 48 h versus 10 h post-ovulation. Column labels individual biological replicates (see key for embryo type and collection time). Colours indicate low expression (blue) and high expression (red) on log-normalised scale.....329
Figure 10-8 Heat map and hierarchical clustering of expression of genes assigned to the ‘cell cycle’ gene ontology term that were significantly down-regulated in expression in samples of sexual *Daphnia magna* embryos collected at 48 h versus 10 h post-ovulation. Row labels indicate gene name with corresponding annotation based on OrthoDB orthologous gene family group (also listed in Table 10-11) and coloured column top panels indicate time post-ovulation (refer to legend).. Colours indicate low expression (blue) and high expression (red) on log-normalised scale. 330
Figure 10-9 Heat map and hierarchical clustering of expression of genes assigned to the ‘DNA replication’ gene ontology term that were significantly down-regulated in expression in samples of sexual *Daphnia magna* embryos collected at 48 h versus 10 h post-ovulation. Row labels indicate gene name with corresponding annotation based on OrthoDB orthologous gene family group (also listed in Table 10-12) and coloured column top panels indicate time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale. 334
Figure 10-10 Heat map and hierarchical clustering of expression of genes assigned to both ‘cell cycle’ and ‘DNA replication’ gene ontology terms that were that were significantly down-regulated in expression in samples of sexual *Daphnia magna* embryos collected at 48 h versus 10 h post-ovulation. Row labels indicate gene name with corresponding annotation based on OrthoDB orthologous gene family group (also listed in Table 10-11 and Table 10-12) and coloured column top panels indicate time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale. 335
Figure 11-1 Heat map and hierarchical clustering of expression of genes with non-zero loadings in the second sparse principal component in sparse principal component analysis of RNA-Seq data from diapause-destined *Daphnia magna* embryos collected between 2 h and 2 weeks (408 h) post-ovulation, subitaneous embryos are shown for comparison only. Row labels indicate gene name and coloured column top panels indicate time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale. 341

Figure 11-2 Intensity of peak (*m*/z = 1145.591560) with a non-zero loading in the second sparse principal component with high mean intensity at 120 h post-ovulation identified through sparse principal component analysis of log-normalised peak intensity matrix obtaining following the negative ion analysis of non-polar *Daphnia magna* embryo extracts collected between 2 h and 17,592 h (2 y) post-ovulation. Individual points represent biological replicates, blue trend line fitted with locally weighted scatterplot smoothing (LOESS; span = 0.5), grey shading indicates model 95% confidence interval. 346
Figure 11-3 Intensity of peak (*m*/z = 370.1933764) with a none-zero loading in the first sparse principal component with high mean intensity at 24 h post-ovulation identified through sparse principal component analysis of log-normalised peak intensity matrix obtaining follow the positive ion analysis of polar *Daphnia magna* embryo extracts collected between 2 h and 17592 h (2 y) post-ovulation. Individual points represent biological replicates, lines fitted with locally weighted scatterplot smoothing (LOESS; span = 0.5), grey shading indicates model 95% confidence interval. 357
Figure 11-4 Heat map and hierarchical clustering of genes relating to protein synthesis in samples of diapause-destined *Daphnia* embryos collected between 2 h and 2 weeks (408 h) post-ovulation. Row labels indicate corresponding annotation based on OrthoDB orthologous gene family group and coloured column top panels indicate time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale. 358

LIST OF TABLES

Table 1-1 Summary of the life stage at which model species in studies referenced within this thesis initiate diapause. (*) Indicates that three distinct embryonic forms of *Austrofundulus limnaeus* have been described. 8
Table 2-1 Summary of the major ionic constituents of water drawn from the Sherwood Sandstone aquifer via bore holes on the University of Birmingham campus. Source: Hydrogeological Services International Ltd report (2001) “The University of Birmingham Groundwater Supply – Test Pumping of Environment Agency Borehole”. 35
Table 2-2 Stock composition of Bold’s basal medium 36
Table 2-3 Concentrations of *Chlorella vulgaris* algal food used in the selection of culturing parameters for the induction of sexual reproduction in *Daphnia magna* 38
Table 2-4 Electrospray ionisation settings for analysis of *Daphnia magna* embryos with nESI-DIMS 51
Table 3-1 Analysis of variance table investigating the effect of gender and *Chlorella vulgaris* food ration on age (d) of *Daphnia magna* at 100% culture mortality 68
Table 3-2 Estimated dry mass of pooled Daphnia magna sexual embryos. SD = standard deviation. 83
Table 3-3 Estimated dry mass of samples analysed for the biomass optimisation of *Daphnia magna* sexual embryos utilising nanoelectrospray ionisation-direct infusion mass spectrometry. 83
Table 3-4 Table of median peak relative standard deviation (RSD) and peak count obtained by nano-electrospray ionisation-direct infusion mass spectrometry analysis of polar and non-polar extracts from different numbers of *Daphnia magna* sexual embryos. QC = quality control samples. 86
Table 4-1 Summary of the first and second principal components and significant embryo type and time separation by mean sample group scores determined by ANCOVA (type III SS) from PCA of log-transformed gene counts and metabolite intensities of diapause-destined and subitaneous embryos of *Daphnia magna* collected between 2 h and ~2 y (17,592 h) post ovulation; Italic values = $p < 0.05$, red values = $p < 0.01$ 100
Table 4-2 Summary of the first and second principal components and significant embryo type and time separation by mean sample group scores determined by ANCOVA (type III SS) from PCA of log-transformed gene counts and metabolite intensities of diapause-destined and subitaneous embryos of *Daphnia magna* collected between 2 h and 48 h post ovulation; Italic values = $p < 0.05$, red values = $p < 0.01$ 103
Table 4-3 Summary of the number of significantly changing metabolites or genes from ANOVA-simultaneous component analysis and differential gene expression

analysis respectively, associated with time post-ovulation, embryo type (subitaneous or diapause-destined) and the interaction thereof. Significantly changing metabolites were determined based on a leverage threshold of 0.9 and squared prediction error alpha level of 0.05. Significantly changing genes were selected based on a false-discovery rate corrected *p*-value threshold of 0.01. 104
Table 4-4 Summary of the GO-annotation of differentially expressed genes after false discovery rate correction ($p < 0.1$) between time post-ovulation (h), embryo type (subitaneous versus diapause-destined) and the interaction thereof. Numbers represent the unique functional categories assigned under each domain. Separate "background' column indicates the total number of genes within the *Daphnia magna* transcriptome, used as a reference set for over-representation analysis. 107
Table 4-5 Table showing putative annotation of spectral peaks from negative ion analysis of non-polar *Daphnia magna* embryo extracts that showed significant changes in intensity between subitaneous and diapause-destined embryos based on ANOVA simultaneous component analysis (ASCA) models with a leverage threshold of 0.9 and squared prediction error (SPE) alpha level of 0.05. Observed peak *m/z* is shown in addition to ASCA model statistics (Leverage and SPE). Formula indicates putative empirical formulae assigned to each peak using BEAMS, with corresponding annotation including number of possible species attributable to the aforementioned formula, respective putative compound identity with theoretical mass. Mass error (ppm) is the difference between theoretical and experimental *m/z*, normalised to experimental *m/z.* Other putative formula without compound ID assignment are indicated, for peaks with a high number of possible formulae only number of formulae are indicated for brevity. 115
Table 4-6 Table showing putative annotation of spectral peaks from negative ion analysis of polar *Daphnia magna* embryo extracts that showed significant changes in intensity between subitaneous and diapause-destined embryos (embryo type), time (h) post-ovulation or the interaction thereof based on ANOVA simultaneous component analysis (ASCA) models with a leverage threshold of 0.9 and squared prediction error alpha level of 0.05. Observed peak *m/z* is shown in addition to ASCA model statistics (Leverage and SPE). Formula indicates putative empirical formulae assigned to each peak using BEAMS, with corresponding annotation including number of possible species attributable to the aforementioned formula, respective putative compound identity with theoretical mass. Mass error (ppm) is the difference between theoretical and experimental *m/z*, normalised to experimental *m/z.* Other putative formula without compound ID assignment are indicated, for peaks with a high number of possible formulae only number of formulae are indicated for brevity. 116
Table 4-7 Table showing putative annotation of spectral peaks from positive ion analysis of polar *Daphnia magna* embryo extracts that showed significant

changes in intensity between subitaneous and diapause-destined embryos (embryo type), time (h) post-ovulation or the interaction thereof based on ANOVA simultaneous component analysis (ASCA) models with a leverage threshold of 0.9 and squared prediction error alpha level of 0.05. Observed peak *m/z* is shown in addition to ASCA model statistics (Leverage and SPE). Formula indicates putative empirical formulae assigned to each peak using BEAMS, with corresponding annotation including number of possible species attributable to the aforementioned formula, respective putative compound identity with theoretical mass. Mass error (ppm) is the difference between theoretical and experimental *m/z*, normalised to experimental *m/z.* Other putative formula without compound ID assignment are indicated, for peaks with a high number of possible formulae only number of formulae are indicated for brevity. 119
Table 5-1 Table of number of genes and metabolites that displayed a significant (false discovery rate (FDR) adjusted *p*-value) changes relative to the initiation of diapause in *Daphnia magna* embryos*.* FC = fold change. Total indicates the number of features detected within the two time points compared. 'large' log_2 fold change indicates the number of features with a $-2 > log_2$ FC > 2. 138
Table 5-2 Table showing putatively annotated spectral peaks from the negative and positive ion metabolomics analysis of sexually produced *Daphnia magna* embryo polar extracts with signifcant ($p < 0.01$; after false discovery rate correction) large peak intensity log_2 fold change ratios (> 2) between 48 h and 10 h post-ovulation. Other putative formulae without compound identity assignment are indicated, for peaks with a high number of possible formulae only numbers of formulae are indicated for brevity. 145
Table 5-3 Table showing putative annotation of spectral peaks from negative ion metabolomics analysis of sexually produced *Daphnia magna* embryo polar extracts with signifcant ($p < 0.01$; after false discovery rate correction) large peak intensity log_2 fold change ratios (> 2) between 74 h and 48 h post-ovulation. Other putative formulae without compound identity assignment are indicated, for peaks with a high number of possible formulae only numbers of formulae are indicated for brevity. 146
Table 5-4 Table showing putative annotation of spectral peaks from positive ion metabolomics analysis of sexually produced *Daphnia magna* embryo polar extracts with signifcant ($p < 0.01$; after false discovery rate correction) large peak intensity log_2 fold change ratios (> 2) between 74 h and 48 h post-ovulation. Other putative formulae without compound identity assignment are indicated, for peaks with a high number of possible formulae only numbers of formulae are indicated for brevity. 148
Table 5-5 Table of number of genes that displayed a significant (false discovery rate (FDR) adjusted *p*-value < 0.1) change in expression relative to the initiation of diapause at 48 h v 10 h or 74 h v 48 h post-ovulation in *Daphnia magna* embryos. Up- or Down-regulated genes are based on expression at the later

time point versus the earlier. Of the differentially expressed genes the number with putative PANTHER gene family annotations used for further GO-annotation overrepresentation analysis are indicated. 158
Table 6-1 Summary of published studies investigating diapause development including relevant organism and diapause life history and experimental design. Where reported, total duration indicates maximum time in diapause until onset of mortality. Data type collected and statistical design used in analysis are also indicated (PCA = principal component analysis, GLM = generalized linear models). Experimental parameters include induction of diapause method by changing photoperiod (LD = Long Day; SD = Short Day; values reported as light:dark) and sampling time points (days) are indicated by grey cells relative to organism diapause phase. 174
Table 6-2 Summary of sparse principal component analysis conducted on the log-normalised gene count and metabolite peak intensity matrices from the analysis of *Daphnia magna* sexually produced embryos collected between 2 h and 2 y (17592 h) post-ovulation, including the proportion of variance explained and the number of elements (peaks or genes) assigned non-zero loadings for the first four sparse principal components. 179
Table 6-3 Table of metabolite peak intensities detected during the negative ion metabolomics analysis of non-polar *Daphnia magna* sexual embryos collected at 408 h and 17592 h (2 y) post-ovulation. FC = fold change, calculated relative to 2 y. Peak *m/z* values indicated in bold denote peaks specific to sexually produced embryos. Peaks with significant (false discovery rate (FDR) adjusted *p*-value < 0.05) changes in intensity between the two time points are also indicated in bold and were determined based on a Student's t-test. Green shading denotes peaks assigned at least one putative annotation. 189
Table 6-4 Table showing the number of peaks detected in polar metabolomic analysis with non-zero loadings for the first and second sparse principal components (sPC) that demonstrated significant differences in intensity between time points. Significance was determined by a Student's t-test and false discovery rate-adjusted *p*-value of less than 0.01. 196
Table 6-5 Summary of the number of features (i.e. genes or metabolites) with significant changes in expression (genes) or intensity (metabolites) from successive pairwise comparisons between *Daphnia magna* embryos sampled during diapause development. Significantly changing features were determined by $p < 0.01$ (after Benjamini-Hochberg false discovery rate correction) and are also plotted in Figure 6-12. Elements that were significant only within one time interval (plotted in Figure 6-14) and with large (> 2-fold) $\log_2$ fold changes (FC; plotted in Figure 6-15) are also reported. 208
Table 9-1 Mean number of cells (DAPI-stained nuclei) in asexually produced subitaneous embryos of *Daphnia magna* in the hours post-ovulation 255

Table 9-2 Mean number of cells (DAPI-stained nuclei) in sexually produced diapause-destined embryos of *Daphnia magna* in the hours post-ovulation 256
Table 9-3 Table of *Daphnia magna* embryo sample sequencing metrics, read quality and successful alignment to transcriptome 257
Table 9-4 Number of mass spectral peaks retained in intensity matrices throughout successive filtering stages during the processing of data obtained from the nESI-DIMS mass spectrometry metabolomics analysis of negative and positive ion analysis of *Daphnia magna* sexual and asexual embryo polar and non-polar extracts. RSD = relative standard deviation 265
Table 9-5 Table of the number of biological replicates retained in final data matrices (bold) following nESI-DIMS metabolomics analysis of polar and non-polar extracts of sexually and asexually produced *Daphnia magna* embryos collected at various time points post-ovulation. All time points contained an initial 8 biological replicates and 'processed' columns refer to the number of samples for which data was successfully acquired. 266
Table 9-6 The proportion of metabolite peaks detected during the nESI-DIMS analysis of subitaneous and diapause embryo samples that were unique to each respective time point sampled. 273
Table 9-7 Table of mass spectral peaks retained in intensity matrices throughout successive filtering stages during the processing of data obtained from the nESI-DIMS mass spectrometry negative and positive ion analysis of *Daphnia magna* diapause embryo polar and non-polar extracts. RSD = relative standard deviation. 276
Table 9-8 Table of the number of biological replicates retained in final data matrices (bold) following nESI-DIMS metabolomics analysis of polar and non-polar extracts of sexual *Daphnia magna* embryos collected at various time points post-ovulation. All time points contained an initial 8 biological replicates and 'processed' columns refer to the number of samples for which data was successfully acquired. 278
Table 9-9 Gene ontology overrepresentation analysis of genes differentially expressed with time (h) post-ovulation after false discovery rate correction ($p < 0.1$). Background contained 19,244 genes, and original input contained 4989 genes 280
Table 9-10 Gene ontology overrepresentation analysis of genes differentially expressed based on an interaction between embryo type and time (h) post-ovulation after false discovery rate correction ($p < 0.1$). Background contained 19,244 genes, and original input contained 3013 genes 289
Table 9-11 Table of all genes with OrthoDB gene families associated with glutathione that were differentially expressed with time (h) post ovulation of the interaction of embryo type and time (h) post-ovulation. 297
Table 10-1 Table showing putative annotation of spectral peaks from postive and negative ion analysis of polar sexual *Daphnia magna* embryo extracts with

signifcant (p < 0.01; after false discovery rate correction) large peak intensity fold change ratios (> 2) between 24 h and 48 h post-ovulation. Other putative formulae without compound identity assignment are indicated, for peaks with a high number of possible formulae only numbers are indicated for brevity. 303
Table 10-2 Table showing spectral peaks from the negative ion analysis of polar extracts of sexually produced *Daphnia magna* embryos with signifcant (p < 0.01; after false discovery rate correction), large peak intensity log_2 fold change ratios (> 2) in relation to the onset of diapause (ca. 48 h post-ovulation). Peaks indicated in bold were unique only to diapause developmental samples when subitaneous embryos were excluded from analysis. 306
Table 10-3 Table showing spectral peaks from the positive ion analysis of polar extracts of sexually produced *Daphnia magna* embryos with signifcant (p < 0.01; after false discovery rate correction), large peak intensity log_2 fold change ratios (> 2) in relation to the onset of diapause (ca. 48 h post-ovulation). Peaks indicated in bold were unique only to diapause developmental samples when subitaneous embryos were excluded from analysis. 307
Table 10-4 Mean intensity of putatively annotated trehalose peak (*m/z* 381.0799831) in *Daphnia magna* sexual embryos collected between 2 h and 2 y (17592 h) post-ovulation. 308
Table 10-5 Table showing log_2 fold changes between 10 h, 48 h and 74 h post-ovulation, significance (p-values adjusted for false discovery rate), and relationship between mean log-normalised gene expression and trehalose metabolite intensity assessed with Spearman's rank correlation coefficient (ρ) of genes related to trehalose metabolism in the sexual embryos of *Daphnia magna.* NA represents genes that with outlying counts (in sample replicates) determined by Cook's distance (cut off = 0.99 quantile), (–) represents genes not detected in the dataset. 309
Table 10-6 Table of gene ontology terms significantly overrepresented in genes differentially expressed between sexually produced *Daphnia magna* embryos collected 48 h and 10 h post-ovulation after false discovery rate correction (p < 0.1). Background contained 19,244 genes, and original input contained 5,127 genes (2681 up-regulated, 2,446 down-regulated; direction refers to gene expression at 48 h relative to 10 h). 312
Table 10-7 Table of gene ontology terms overrepresented in genes differentially expressed between sexually produced *Daphnia magna* embryos collected 74 h and 48 h post-ovulation after false discovery rate correction (p < 0.1). Background contained 19,244 genes, and original input contained 229 genes (135 up-regulated, 94 down-regulated; direction refers to gene expression at 74 h relative to 48 h). 314
Table 10-8 Table showing log_2 fold changes between 10 h, 48 h and 74 h post-ovulation, significance (p-values adjusted for false discovery rate) of putatively annotated superoxide dismutase genes in the sexually produced embryos of

Daphnia magna. NA represents genes with outlying counts (in sample replicates) determined by Cook's distance. 320
Table 10-9 Table of *Daphnia magna* genes with significant (p-adjust < 0.1) increased expression at 48 h versus 10 h post-ovulation that were assigned putative identities and membership to 'lipid metabolic process' gene ontology during Protein ANalysis THrough Evolutionary Relationships (PANTHER) overrepresentation analysis. 322
Table 10-10 Table of *Daphnia magna* genes with significant (p-adjusted < 0.1) increased expression at 48 h versus 10 h post-ovulation that were assigned putative identities and membership to 'oxidoreductase activity' gene ontology during Protein ANalysis THrough Evolutionary Relationships (PANTHER) overrepresentation analysis. 325
Table 10-11 Table of *Daphnia magna* genes assigned to the 'cell cycle' gene ontology term, their corresponding OrthodDB gene family and putative protein identity that were significantly up-regulated in expression in samples of sexual embryos collected at 74 h versus 48 h post-ovulation (see Figure 10-8 for expression). 331
Table 10-12 Table of *Daphnia magna* genes assigned to the 'DNA replication' gene ontology term, their corresponding OrthodDB gene family and putative protein identity that were significantly up-regulated in expression in samples of sexual embryos collected at 74 h versus 48 h post-ovulation. 336
Table 11-1 Table of genes with non-zero loadings for sparse principal components (sPC) 1–4 from the sparse principal component analysis of RNA-Seq data from sexually-produced *Daphnia magna* embryos collected between 2 h and 408 h (2 weeks) post-ovulation. *Daphnia magna* gene ID is indicated with corresponding annotation with OrthoDB gene family designation and identity. 338
Table 11-2 Table of non-polar metabolite spectral peaks with non-zero loadings for sparse principal components (sPC) 1–4 from the sparse principal component analysis of nESI-DIMS data from negative ion analysis of non-polar extracts of *Daphnia magna* embryos collected between 2h and 17,592 h (2 y) post-ovulation. Observed peak *m/z* is shown in addition to sPC loadings, bold values indicate peaks not detected in subitaneous embryo sample dataset. Formula indicates putative empirical formulae assigned to each peak using BEAMS, with corresponding annotation including number of possible species attributable to the aforementioned formula, respective putative compound identity with theoretical mass. Mass error (ppm) is the difference between theoretical and experimental *m/z*, normalised to experimental *m/z.* Other putative formula without compound ID assignment are indicated, for peaks with a high number of possible formulae only number are indicated for brevity. 342
Table 11-3 Table of polar metabolite spectral peaks with non-zero loadings for sparse principal components (sPC) 1–4 from the sparse principal component analysis of nESI-DIMS data from negative ion analysis of polar extracts of

Daphnia magna embryos collected between 2 h and 17,592 h (2 y) post-ovulation. Observed peak *m/z* is shown in addition to sPC loadings, bold values indicate peaks not detected in subitaneous embryo sample dataset. Formula indicates putative empirical formulae assigned to each peak using BEAMS, with corresponding annotation including number of possible species attributable to the aforementioned formula, respective putative compound identity with theoretical mass. Mass error (ppm) is the difference between theoretical and experimental *m/z*, normalised to experimental *m/z.* Other putative formula without compound ID assignment are indicated, for peaks with a high number of possible formulae only number are indicated for brevity. 347

Table 11-4 Table of polar metabolite spectral peaks with non-zero loadings for sparse principal components (sPC) 1–4 from the sparse principal component analysis of nESI-DIMS data from positive ion analysis of polar extracts of *Daphnia magna* embryos collected between 2 h and 17,592 h (2 y) post-ovulation. Observed peak *m/z* is shown in addition to sPC loadings, bold values indicate peaks not detected in subitaneous embryo sample dataset. Formula indicates putative empirical formulae assigned to each peak using BEAMS, with corresponding annotation including number of possible species attributable to the aforementioned formula, respective putative compound identity with theoretical mass. Mass error (ppm) is the difference between theoretical and experimental *m/z*, normalised to experimental *m/z.* Other putative formula without compound ID assignment are indicated, for peaks with a high number of possible formulae only number are indicated for brevity. 352

LIST OF ABBREVIATIONS

ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
ASCA	ANOVA-simultaneous component analysis
BEAMS	Birmingham Metabolite Annotation for Mass Spectrometry
DAPI	4',6-diamidino-2-phenylindole
DE	Differential expression
DIMS	Direct-infusion mass spectrometry
FDR	False discovery rate
FOXO	Forkhead transcription regulatory factor
GO	Gene ontology
HMDB	Human Metabolite DataBase
HSP	Heat shock protein
ID	Identity
JH	Juvenile hormone
KNN	K-nearest-neighbour
LipidMAPS	Lipid Metabolites and Pathways Strategy
LOESS	Locally estimated scatterplot smoothing
m/z	Mass:charge ratio
MS	Mass spectrometry
nESI	Nano electrospary ionization
NMR	Nuclear magnetic resonance
PANTHER	Protein ANalysis THrough Evolutionary Relationships
PBS	Phosphate buffered saline
PC	Principal component
PCA	Principal component analysis
PCNA	Proliferating cell nuclear antigen
ppm	Parts per million
PQN	Probabilistic quotient normalisation
PQN	Probabilistic quotient normalisation
PTTH	Prothoracicotropic hormone
QC	Quality control
SD	Standard deviation
SIM	Single ion monitoring
sPC	Sparse principal component
sPCA	Sparse principal component analysis
SPE	Squared prediction error
TPP	Trehalose-6-phosphate phosphatase
TPS	Trehalose-6-phosphate synthase
UDP	Uridine-diphosphate
UV	Ultra-violet
RSD	Relative standard deviation

1. Introduction

1.1. Diapause: an alternative developmental trajectory

Diapause is a form of animal dormancy characterised by a reversible arrest of development and metabolic suppression, that is typically utilised as a means to survive through inimical environmental conditions (Koštál, 2006). Diapause is unique from other dormancies (e.g. quiescence) as the stage, entry, and length of arrest is endogenously programmed in advance of onset. Moreover, direct development does not necessarily resume once an inducing period of environmental hardship is over (Koštál, 2006). Diapause is an important stage in the life cycle of species across the animal kingdom; it has been described across taxa ranging from nematodes (*Caenorhabditis elegans* “dauer” larvae), mammals (in excess of 130 species, reviewed by Fenelon, Banerjee and Murphy (2014)) to arthropods including crustaceans and branchipods. But it is perhaps most widely studied in insects (Denlinger, 2002; Fielenbach and Antebi, 2008; Fritsch and Richter, 2015; Fenelon *et al.*, 2017).

Diapause may be either facultative or obligatory. For example a wide range of mammals including, the roe deer *Capreolus capreolus* –the first scientifically documented example of diapause (Ziegler, 1843)– exhibit obligate, pre-implantation embryonic diapause that occurs every gestation (Fenelon *et al.*, 2017). Insects may also undergo obligate adult diapause including the solitary bee *Osmia lignaria*, which commences diapause in autumn and terminates in spring (Sgolastra *et al.*, 2010). However, the majority of described diapauses are facultative and closely linked to

environmental conditions including seasonal patterns (Denlinger, 2002). For the remainder of this thesis, the term diapause is taken to refer to facultative diapause. Despite manifesting as a developmental arrest, diapause is by no means a static phenomenon. Instead, diapause is a dynamic, alternative developmental program consisting of sequential phases. This program is marked by multiple physiological changes and differential gene expression, occurring throughout the trajectory (Ragland, Denlinger and Hahn, 2010). The phases of diapause are defined most clearly by Koštál (2006) as three “eco-physiological phases”: pre-diapause, diapause and post-diapause (Figure 1-1). Although these phases are largely based on generalisations of insect diapause, these are applied throughout this thesis.

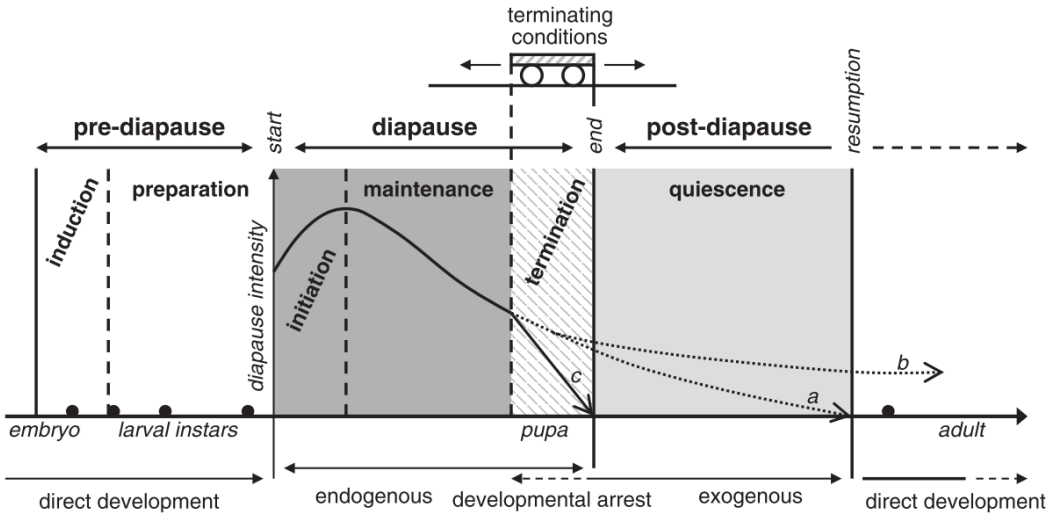


Figure 1-1 ‘The eco-physiological phases of diapause’. The alternative developmental trajectory of diapause encompasses three progressive phases; pre-diapause, diapause and post-diapause of variable length and duration according to environmental conditions. Key events include the induction of diapause, the initiation of diapause (as onset of developmental arrest) and the termination of diapause (resumption of development). The term “diapause intensity” indicates, relative to a given time point, the duration of developmental arrest under set environmental conditions and thus the propensity of an organism to terminate diapause. Potential developmental trajectories are also highlighted including resumption of direct development upon termination of diapause, with (a) or without (c) a period of post-diapause quiescence and failure to reinitiate development (b) resulting in death. This figure is reproduced from Koštál (2006).

There is frequently a temporal separation between the induction of diapause and its onset; this phase is known as ‘pre-diapause’ (Koštál, 2006). Here direct development

continues but the organism is committed to later enter diapause. Often, a series of diapause-specific molecular, behavioural and physiological changes may occur including prolonged development, that are thought to prepare the organism for diapause (Denlinger, 2002). The initiation of diapause is defined as the point at which direct morphogenesis stops (Koštál, 2006). A concomitant decrease in metabolism at initiation is also considered a hallmark of diapause (Ragland, Denlinger and Hahn, 2010). Respiration rate is often used as a proxy measurement of metabolism and has been shown to be considerably lower during diapause in comparison to non-diapausing individuals: the diapausing embryos of *Artemia franciscana* show a rapid decrease in respiration rate over the first five days of diapause with respiration after 20 days reducing to just 2% of original levels (Podrabsky and Hand, 2015).

During the maintenance of diapause, developmental arrest and metabolic suppression continue. Organisms enter a refractory period where arrest is maintained even if environmental conditions become permissive to direct development (Gyllström and Hansson 2003). Diapause maintenance is relatively poorly understood in contrast to the other phases of diapause. The molecular mechanisms that maintain diapause and changes that occur that move the organism towards the end of refractory period are largely unclear in many organisms. Diapause duration in the copepod *Paracartia latisetosa* was found to directly correlate with the length of maternal photoperiod (Belmonte and Pati, 2007) suggesting maternal control; but it is unknown how this information is transferred to the offspring. Furthermore, the conditions that are experienced and the length of diapause can also have a profound effect on later life stages of the organism. With increasing time spent in diapause, an increase in mortality, decrease in fat reserves and subsequent

reduction in the size and fitness of emerging adults have been observed in the parasitioid *Asobara tabida* (Ellers and Van Alphen, 2002). Diapausing larvae of the mosquito *Wyeomyia smithii* exhibit over 60% loss in fitness when exposed to simulated winter versus summer conditions (Bradshaw, Armbruster and Holzapfel, 2009).

Termination of diapause is marked by a return of the potential for development, often in response to the perception of specific termination stimuli (Koštál, 2006). For example, light, desiccation and reduced temperature promotes breaking of *A. franciscana* cyst diapause (Robbins *et al.*, 2010). Organisms that have terminated diapause may also enter post-diapause quiescence where developmental arrest persists but is mediated by external stimuli (Padilla and Ladage, 2012). *S. crassipalpis* pupae will terminate diapause after around 40 days at 20°C but will remain developmentally arrested for at least 110 days if held at 10°C (Hayward, *et al.,* 2005).

The perception of environmental cues and the regulatory mechanisms controlling the entry and exit to diapause are likely subject to strong selective pressure (Gyllström and Hansson 2003). As an overwintering survival mechanism the optimal timing of diapause events may be highly decisive of organism success by determining the extent of impact of adverse seasonal conditions (Dong *et al.*, 2014). Chiefly and by its nature, diapause results in the interruption of the continuous life cycle. In short-lived species such as monogonont rotifers or cladoceran branchiopods like *Daphnia sp.,* there is no guarantee that the diapause inducing stimuli experienced are truly representative of the onset of winter, thus the significant investment of resources into diapause represents a certain ‘bet-hedging’ (Gilbert, 2003).

Furthermore, intra-species differences in response to diapause-inducing stimuli (e.g., photoperiod), have previously been reported (Ferrari and Hebert, 1982; Fitzsimmons and Innes, 2006; Roulin *et al.*, 2015). A notable example is a latitudinal cline in diapause induction in response to photoperiod in *Daphnia magna*. Canadian populations produced diapausing offspring in response to reduced photoperiod. This treatment had no observable effect on populations originating from England (Ferrari and Hebert, 1982). Intra-specific variation is also observed in *Nasonia vitripennis,* whereby northern populations demonstrate a longer critical photoperiod for diapause induction (Paolucci, Van de Zande and Beukeboom, 2013). This variation was previously attributed to local environmental adaption (Deng, 1996; Paolucci, Van de Zande and Beukeboom, 2013) thereby demonstrating the complex genetic mechanisms underlying diapause induction, and highlighting a significant limitation in that, there may exist differences in response to diapause environmental stimuli at a species level.

There are also notable disadvantages to diapause that should not be overlooked. Immobile diapausing individuals are also more susceptible to predation, increasing diapause-associated mortality (Sujii *et al.*, 2002; Murphy and Lill, 2010). The post-diapause offspring of *D. magna* are smaller in size than subitaneous offspring and significantly less resistant to starvation (Arbačiauskas and Lampert, 2003). The suppression of biological process such as protein turnover was also observed during diapause. However, this may limit the ability of organisms to mount an effective translational response to environmental challenges during diapause (Podrabsky and Hand, 2000). It was suggested that critical proteins (e.g. heat shock proteins) are produced during pre-diapause preparatory period to provide later protection. But these

preparations may not be entirely sufficient to ensure organism survival in an ever-changing environment (Podrabsky and Hand, 2015).

1.2. Diapause: current understanding and continuing ambiguity

Diapause represents a complex alternative developmental trajectory progressing through multiple stages (Koštál, 2006). However, a clear understanding of the molecular mechanisms behind the manifestation of diapause has largely eluded researchers (Poelchau et al., 2013). This knowledge gap is in part, due to an apparent lack of phylogenetic conservation in the ontogenetic stage at which diapause occurs. The model diapause species referenced within this thesis are listed in Table 1-1including the developmental stage at which diapause is observed, and ranges from embryonic through to adult diapauses. This remarkable diversity of manifestations persists within taxa, for example: not all insects exhibit diapause, the dipteran *Sarcophaga crassipalpis* enters diapause in a pupal form, whereas *Aedes albopictus* enters diapause at the egg stage (Denlinger, 1986). However, diapause typically occurs at only one distinct ontogenic stage within each species with the notable exception of the annual killifish *Austrofundulus limnaeus,* where three distinct ontogenic stages of diapause are possible during embryogenesis (Denlinger, 2002; Podrabsky and Hand, 2015). It has been suggested that diapause may have evolved on multiple separate occasions, even within insect taxa (Denlinger, 1986; Ragland and Keep, 2017). This uncertainty raised the question as to whether observations made in one species regarding diapause are applicable to the phenomena as a whole, including whether observations made at one developmental stage (e.g. adult)

are truly applicable to other stages such as embryonic diapause (Ragland, Denlinger and Hahn, 2010; Podrabsky and Hand, 2015). Nevertheless, concerted research efforts (particularly in insects) have given some insight into the physiology of diapause and several recurring phenotypes include, developmental arrest, alterations to metabolism and gene expression, in addition to increased stress tolerance are observed.

Table 1-1 Summary of the life stage at which model species in studies referenced within this thesis initiate diapause. (*) Indicates that three distinct embryonic forms of *Austrofundulus limnaeus* have been described.

Species	Class	Order	Life stage	Reference
Artemia franciscana	Branchipoda	Anostraca	Embryo	Podrabsky and Hand, 2015
Artemia parthenogenetica	Branchipoda	Anostraca	Embryo	Meller *et al.*, 2012a
Capreolus capreolus	Mammalia	Artiodactyla	Embryo	Fenelon *et al.*, 2017
Diaptomnus sanguineas	Maxillopoda	Calanoida	Embryo	Hairston *et al.*, 1995
Paracartia latisetosa	Maxillopoda	Calanoida	Embryo	Belmonte and Pati, 2007
Daphnia magna	Branchipoda	Cladocera	Embryo	Arbačiauskas and Lampert, 2003
Daphnia pulicaria	Branchipoda	Cladocera	Embryo	Frisch, *et al*., 2014
Daphnia pulex	Branchipoda	Cladocera	Embryo	Hiruta and Tochinai, 2014
Austrofundulus limnaeus	Actinopterygii	Cyprinodontiformes	Embryo*	Podrabsky and Hand, 2015
Kryptolebias marmortus	Actinopterygii	Cyprinodontiformes	Embryo	Mesak, Tatarenkov and Avise, 2015
Aedes albopictus	Insecta	Diptera	Embryo	Poelchau *et al.*, 2013b
Chymomyza costata	Insecta	Diptera	Larvae	Koštál *et al.*, 2009
Culex pipiens	Insecta	Diptera	Adult	Sim and Denlinger, 2011
Delia antiqua	Insecta	Diptera	Pupae	Hao *et al.*, 2016
Drosophila melanogaster	Insecta	Diptera	Adult	Kubrak *et al.*, 2014
Drosophila montana	Insecta	Diptera	Adult	Kankare *et al.*, 2016
Sarcophaga crassipalpis	Insecta	Diptera	Pupae	Ragland, Denlinger and Hahn, 2010
Wyeomyia smithii	Insecta	Diptera	Larvae	Bradshaw, Armbruster and Holzapfel, 1998
Dolycoris baccarum	Insecta	Hemiptera	Adult	Bashan and Cakmak, 2005
Piezodorus lituratus	Insecta	Hemiptera	Adult	Bashan and Cakmak, 2005
Asobara tabida	Insecta	Hymenoptera	Prepupa	Ellers and Van Alphen, 2002
Chouioia cunea	Insecta	Hymenoptera	Larvae	Zhao *et al.*, 2018
Nasonia vitripennis	Insecta	Hymenoptera	Larvae	Paolucci, Van de Zande and Beukeboom, 2013
Osmia bicornis	Insecta	Hymenoptera	Imagines	Dmochowska-Ślęzak *et al.*, 2015

Osmia lignaria	Insecta	Hymenoptera	Adult	Sgolastra *et al.*, 2010
Antheraea mylitta	Insecta	Lepidoptera	Pupae	Jena *et al.*, 2013
Bombyx mori	Insecta	Lepidoptera	Embryo	Nakagaki *et al.*, 1991
Danaus plexippus	Insecta	Lepidoptera	Adult	Paul, 2002
Diatraea grandiosella	Insecta	Lepidoptera	Larvae	Kuthiala and Chippendale, 1989
Helicoverpa armigera	Insecta	Lepidoptera	Pupae	Bao and Xu, 2011
Helicoverpa zea	Insecta	Lepidoptera	Pupae	Zhang and Denlinger, 2012
Ostrinia nubilalis	Insecta	Lepidoptera	Larvae	Jovanović-Galović *et al.*, 2004
Pieris napi	Insecta	Lepidoptera	Pupae	Lehmann *et al.*, 2018
Sesamia nonagrioides	Insecta	Lepidoptera	Larvae	Perez-Hedo, Sanchez-Lopez and Eizaguirre, 2012
Allonemobius socius	Insecta	Orthoptera	Pupae	Reynolds and Hand, 2009
Brachionus calyciflorus	Monogononta	Ploimida	Embryo	Gilbert, 2004
Brachionus plicatilis	Monogononta	Ploimida	Embryo	Ziv *et al.*, 2017
Caenorhabditis elegans	Chromadorea	Rhabditida	Larvae	Houthoofd *et al.*, 2002

1.2.1. Mechanisms of developmental arrest during diapause

Alterations to the cell cycle have been reported as a means for the manifestation of diapause-associated developmental arrest. Over 80% of the cells in the central nervous system of diapausing *N. vitripennis* (Shimizu, Mukai and Goto, 2016), *C. costata* (Koštál *et al.*, 2009), and *S. crassipalpis* (Tammariello and Denlinger, 1998a), as well as over 90% of the cells in the diapause embryos of *A. limnaeus* and *Artemia parthenogenetica* are halted in G_0/G_1 phase (Meller *et al.*, 2012a; Podrabsky and Culpepper, 2012; Chen *et al.*, 2016); while cells in diapausing *Bombyx mori* are commonly arrested in G_2 phase (Nakagaki *et al.*, 1991). Progression of the eukaryotic cell cycle is modulated by cyclin and cyclin-dependent kinase activities, which are subject to alterations in gene transcription, but can also be altered by regulatory proteins (e.g. inhibitors and post-translational modifications). For example,

increased phosphorylated (inactive) cyclin-dependent kinase 1 protein levels are observed in the brains of diapausing *S. crassipalpis* (Pavlides, Pavlides and Tammariello, 2011). A p27 homolog and cyclin-dependent kinase inhibitor (*dacapo*) also displays increased mRNA expression in diapausing *C. costata* larvae (Koštál *et al.*, 2009). Proliferating cell nuclear antigen (PCNA), is a protein involved in diverse cellular processes including DNA replication, repair and control of the cell cycle (Maga, 2003). *Pcna* has been implicated in the molecular processes of diapause in several species. A decrease in *pcna* expression levels coincide with the onset of G_0/G_1 cell cycle arrest in *S. crassipalpis* (Tammariello and Denlinger, 1998b; Hayward *et al.*, 2005). *Pcna* is also down-regulated during diapause in *Delia antiqua* (Hao *et al.*, 2016), *Chymomyza costata* (Kostal et al., 2009), and PCNA protein is less abundant in the diapausing eggs of the rotifer *Brachionus plicatilis* (Ziv *et al.*, 2017).

1.2.2. Metabolic suppression during diapause

Respiration rate is often used as a proxy measurement of metabolism and was shown to be considerably lower during diapause in comparison to non-diapausing individuals. A 90% reduction in respiration rate was observed in the diapausing pupae of *S. crassipalpis* versus non-diapausing pupae, with an increase in oxygen consumption upon termination of diapause (Ragland, Denlinger and Hahn, 2010). The diapausing embryos of *A. franciscana* also show a rapid decrease in respiration rate over the first five days of diapause with respiration after 20 days reducing to just 2% of original levels (Podrabsky and Hand, 2015). The respiration rate of *A. limnaeus* embryos is reduced by 70% upon entry to diapause and drops further by

up to 90% (Podrabsky and Hand, 2015). *Allonemobius socius* pupae in diapause have a similar respiration rate to directly developing pupae until around 5 days, where respiration of directly developing pupae continues to increase whilst diapause respiration rate is maintained at the same level for at least 15 days (Reynolds and Hand, 2009).

Cryptobiosis is a state when an organism's "metabolic activity becomes hardly measurable, or comes reversibly to a standstill" (Keilin, 1959), and typically presents in response to anoxia, reduced temperatures and/or desiccation, allowing cryptobiotic organisms to survive extreme environmental changes (Clegg, 2001). Thus, based on this definition regarding metabolism, the diapausing organisms above can be considered cryptobiotic. But these examples are of immobile embryos or pupae, however adult diapauses are not necessarily immobile, and furthermore, do not show the same extreme metabolic depression. The monarch butterfly *Danaus plexippus* and *Drosophila melanogaster* both exhibit an adult diapause characterised by a delay in development of the reproductive system (Paul, 2002; Kubrak *et al.*, 2014). During diapause, *D. plexippus* maintain flight ability, often migrating over long distances and do not appear to show any reduction in metabolism (Paul, 2002). *D. melanogaster* also remain mobile, continuing to feed, albeit at a reduced rate. However, there are increased levels of circulating carbohydrates and proteins in diapausing individuals compared to controls, in addition to an increase in the expression of a glucagon-like adipokinetic hormone that is involved in the regulation of carbohydrate and lipid metabolism (Kubrak *et al.*, 2014). Thus, some organisms in diapause may also be cryptobiotic, however cryptobiosis can occur in the absence of diapause and *vice versa*. For example, in tardigrades diapause and cryptobiosis are independent states

that may or not co-present: several species of tardigrades are capable of cryptobiosis at any point within their life cycle, however diapause is only observed in tardigrade eggs (Guidetti, Altiero and Rebecchi, 2011). The key distinction between the two states is that cryptobiosis arises as a direct consequence of exogenous environmental parameters and will end once a permissive environment is established, whilst the diapause state (under endogenous control) will persist without terminating stimuli (Koštál, 2006).

Non-feeding diapausing organisms must have sufficient reserves to maintain molecular processes during diapause, and to provide sufficient energy to complete development upon diapause termination (Hahn and Denlinger, 2007). There is evidence to suggest that the composition of the three major macronutrient energy reserves (lipids, carbohydrates and amino acids) may be different in diapausing organisms. *S. crassipalpis* pupae show an enrichment of glucose, pyruvate and alanine in diapausing pupae, suggesting a reliance on glycolysis and gluconeogenesis. Furthermore, key enzymes in both these pathways and the tricarboxylic acid cycle displayed altered mRNA levels in relation to diapause (Michaud and Denlinger, 2007; Ragland, Denlinger and Hahn, 2010). An up-regulation of two gluconeogenesis enzymes (phosphoenolpyruvate carboxykinase and pyruvate carboxylase) was found in *S. crassipalpis, D. melanogaster* and *C. elegans* by at least nine-fold and two-fold respectively (Ragland, Denlinger and Hahn, 2010). A decline in all free fatty acids measured in *N. vitripennis*, was observed over the course of diapause, indicating their utilisation as an energy source, and emphasising the dynamic nature of diapause (Li *et al.*, 2015). An

increase in anaerobic metabolism during diapause is also described in *D. melanogaster*, *C. elegans,* and*, W. smithii* (Ragland, Denlinger and Hahn, 2010).

1.2.3. Mechanisms of stress tolerance during diapause

There are numerous examples where diapausing organisms have been shown to tolerate extreme environmental perturbations that are often lethal to the organism during direct development. Some of the physical challenges that organisms are likely to encounter during diapause include temperature, desiccation and anoxia. Organisms that over-winter as a diapausing life-stage are typically exposed to sub-zero temperatures that are below the normal organism range and, in many cases species exhibit diapause-specific increases in cold tolerance (Bale, 1996; Wilches *et al.*, 2016; Hemmati, Moharramipour and Talebi, 2017). For example, diapausing pupae of the wild silkworm, *Antheraea pernyi* exhibit 100% survival after 1 hour exposure to -15°C whereas only 3.6% of non-diapausing first instar larvae of this species survive the same exposure (Liu *et al.*, 2016). There are numerous factors that are thought to contribute to this increased thermal tolerance including accumulation of cyroprotectants and other biochemical adaptions, membrane remodelling and utilisation of cellular protective mechanisms such as heat shock proteins (HSPs). The ability of diapausing life stages to tolerate low temperatures also facilitates the colonisation of seasonally hostile environments increasing organism range, and is considered to contribute to the success of insect taxa (Denlinger, Yocum and Rinehart, 2012).

Desiccation is another potential stress that diapausing organisms often encounter. This is particularly important when considering non-feeding diapausing stages that

will be unable to take on additional water to maintain tissue osmolality. Some aquatic species, particularly those inhabiting temporary water bodies, utilise diapause to mitigate drying and/or freezing and often display a high tolerance for desiccation. Diapausing cysts of *A. franciscana*, express a diapause-specific ferritin homolog called artemin, which is believed to play a role in desiccation resistance as RNAi knockdown of *artemin* lead to a 30% reduction in survival of cysts when exposed to desiccation and freezing stress (King, Toxopeus and Macrae, 2014). Although the exact function of artemin protein is unknown, *in vitro* studies have shown both mRNA binding and protein chaperone capacity (King, Toxopeus and Macrae, 2014).
A. limnaeus diapausing embryos display the highest anoxia tolerance identified in an invertebrate, with a maximum lethal time to 50% mortality under anoxic exposure of 65 days at 25°C. The mechanism is believed to involve a depression of oxidative metabolism (Duerr and Podrabsky, 2010; Meller *et al.*, 2012b; Podrabsky and Culpepper, 2012). Furthermore, these embryos also show a high tolerance to UV-C radiation exposure, possibly through light-mediated repair of DNA lesions (Wagner and Podrabsky, 2015; Hand *et al.*, 2016).

1.2.4. Molecular mechanisms of diapause

Many of the physiological and ecological components of diapause including cues for entry and exit are reasonably well documented (Podrabsky and Hand, 2015; Hand *et al.*, 2016; Wilches *et al.*, 2016; Baumgartner and Tarrant, 2017). However, the molecular mechanisms of diapause remain unclear which is astonishing considering that for many species diapause represents a major life history trait (Mesak, Tatarenkov and Avise, 2015; Qi *et al.*, 2015). As a quantitative trait, diapause is the

summation of complex interactions involving the transcriptional changes of multiple genes (Kang *et al.*, 2016). Large-scale shifts in transcriptional activity have been documented in multiple diapausing species and studies are often contradictory to each other with some studies reporting an increase in global transcript abundance during diapause (Dong *et al.*, 2014; Qi *et al.*, 2015), whilst others report a decrease (Durant *et al.*, 2016). There are also examples where such large transcriptional shifts are not present. For example, the late embryonic diapause of *Kryptolebias marmortus,* differentially expresses a minute fraction (0.1%) of transcripts compared to pre- and post-diapause embryos. The majority (67%) of these differentially expressed genes have reduced expression during diapause (Mesak, Tatarenkov and Avise, 2015). Genes implicated in diapause have been identified in many species and cover broad biological functions including; stress response (Wadsworth and Dopman, 2015), hormone signalling (Poelchau *et al.*, 2013a; Qi *et al.*, 2015; Wadsworth and Dopman, 2015; Kang *et al.*, 2016), cell cycling (Poelchau *et al.*, 2013a; Wadsworth and Dopman, 2015; Thompson and Ortí, 2016), development (Mesak, Tatarenkov and Avise, 2015; Wadsworth and Dopman, 2015), metabolism (Poelchau *et al.*, 2013b; Qi *et al.*, 2015; Wadsworth and Dopman, 2015; Kang *et al.*, 2016), and the immune system (Mesak, Tatarenkov and Avise, 2015; Qi *et al.*, 2015; Kang *et al.*, 2016).

The organism-wide coordination of the physiological and molecular events described above likely involves complex regulation and the involvement of hormonal control is an obvious mechanism to achieve this. The insulin-signalling pathway plays a central role in carbohydrate and lipid metabolism and a potential role in diapause has been investigated in variety of species. RNAi knockdown of the insulin receptor induced

the initiation of a diapause-like phenotype in non-diapausing *C. pipiens* females (Sim and Denlinger, 2008). A downstream target of the insulin pathway is a gene that encodes phosphatidylinositol 3-kinase (PI3K), which is down-regulated in diapausing *D. melanogaster* and *C. elegans* (Williams *et al.*, 2006; Kučerová, Kubrak, *et al.*, 2016). Juvenile hormones (JHs) are a large family of hormones primarily found in insects, which have numerous essential functions in reproduction and development and are also regulated by the insulin pathway (Herman, 1973, 1975; Lessman, Rollins and Herman, 1982; Lessman and Herman, 1983; Tatar and Yin, 2001; Flatt, Tu and Tatar, 2005; Tu, Yin and Tatar, 2005; Denlinger and Armbruster, 2014). In direct development, JH levels are high in larvae, dropping to permit metamorphosis (Denlinger, 2002). However alternative JH titre patterns have been observed during the diapause developmental trajectory. For example, *Sesamia nonagrioides* larvae that enter diapause exhibit continued JH expression preventing pupation and leading to the retention of larval phenotype (Perez-Hedo, Sanchez-Lopez and Eizaguirre, 2012). By contrast, diapause in *C. pipiens* results in a reduction of JH titre, associated with an up-regulation of forkhead transcription regulatory factor (FOXO) expression and concurrent fat accumulation (Denlinger and Armbruster, 2014), and topical application of JH terminates diapause, resulting in ovarian development (Tu, Yin and Tatar, 2005; Sim and Denlinger, 2009).

The larval and pupal diapauses of insects have been directly linked to a decrease in prothoracicotropic hormone (PTTH) titre (Zhang and Denlinger, 2012). PTTH is a neuropeptide that stimulates the prothoracic gland to synthesise the steroid hormone ecdysone and is absent in the diapausing pupae of *Helicoverpa zea* (Zhang and Denlinger, 2012). The role of PTTH and ecdysone in diapause is further

demonstrated by numerous experiments that indicate diapause can be terminated and a resumption of development achieved by exogenous ecdysteroid (20-hydroxyecdysone) application (Zhang and Denlinger, 2012; Dong *et al.*, 2014). Upon the termination of larval diapause in *Ostrinia nubilalis*, Wadsworth and Dopman (2015) discovered the up-regulation of *Neverland (Nvd), Spook (Spo)* and *Phantom (Phm)* genes involved in the production of 20-hydroxyecdysone. Loss of *Nvd* function in *Drosophila* is associated with developmental arrest and reduced growth through a reduction in 20-hyrdroxyecdysone production (Wadsworth and Dopman, 2015). Denlinger (2002) further reviewed the involvement of ecdysteroids and associated receptors in several diapausing insects.

Arguably, one of the best studied embryonic diapauses is that of *B. mori,* where the initiation of diapause is governed by diapause hormone, a neuropeptide hormone that triggers trehalase activity in ovaries, leading to subsequent alterations in carbohydrate metabolism, and action of the polyols sorbitol and trehalose result in diapause (Yamashita, 1996; Denlinger, 2002). Similar diapause hormones have subsequently been identified in excess of 30 insect species. However, it is unclear whether this is conserved outside of insects (Denlinger, Yocum and Rinehart, 2012; Chang and Ramasamy, 2014).

A further understanding of the diapause phenomena including elucidation of the complex regulatory mechanisms, molecular and physiological changes involved will likely shed a powerful light on several fundamental biological questions and have practical uses in understanding the mystery of biological longevity and developmental biology. Considering that development is a complex and canonical process that is fundamental to all animal life, how is it then that in a few organisms this is drawn to

an extensive or complete halt without documented ill effect? I argue that the utility of diapause as a model for ageing is woefully under exploited. In some instances, organisms that enter diapause persist for a much greater length of time than their subitaneous contemporaries. Indeed, extreme examples exist if we consider the crustacean *Daphnia* which has a typical lifespan in the order of weeks to months. However, viable specimens have successfully been resuscitated after centuries (~700 years) in diapause (Frisch *et al.*, 2014). Indeed, planktonic invertebrates appear to exhibit particularly long successful periods of dormancy (reviewed by Radzikowski; Table II; 2013) that regularly span decades. Other extreme examples include the copepod *Diaptomus sanguineas,* which was reported to have successfully hatched after 330 years in diapause (Hairston *et al.*, 1995). The differential expression of genes associated with longevity and ageing has been observed in diapausing *D. melanogaster* (Kučerová, Kubrak, *et al.*, 2016). But there are few other studies investigating the effect of diapause on senescence and the molecular processes behind such prolonged lifespans (Tatar and Yin, 2001). The mechanisms of extreme stress tolerance underlying the unique characteristics that facilitate organism survival in the face of the possible challenging environmental conditions experienced during the diapause period is worthy of study.

1.3. The reproductive biology of *Daphnia*

The model organism *Daphnia sp.* inhabits water bodies likely to dry, freeze or experience other large environmental fluctuations (Gyllström & Hansson 2004). *Daphnia* is the nominal genus of the family Daphniidae, capable of both

parthenogenetic and sexual reproduction (Figure 1-2), which is a feature of all 12 families of the order Cladocera (the water fleas, including the Daphniidae) and a single known species of Cyclestheria (the clam shrimp) among the branchiopods (Fritsch and Richter, 2015). Under permissive environmental conditions, *Daphnia sp.* populations are almost exclusively composed of females that reproduce by cyclical parthenogenesis, thus giving rise to clonal, subitaneous, broods that can number up to 50 offspring in larger species (Hobaek and Larsson, 1990; Burns, 2000; Shaw *et al.*, 2008). *Daphnia sp.* maintain a similar level of fecundity throughout their lifespan; given that a brood is produced every few days, and adult female may produce hundreds of offspring during her lifetime, facilitating rapid population expansion (Li *et al.*, 2017; Booksmythe *et al.*, 2018).

Sexual reproduction in *Daphnia sp.* is induced in response to a variety of isolated or combined stimuli including photoperiod, population density, reduction in food quantity and quality, and fish predation (Stross and Hill, 1965; Brunner and Halcrow, 1977; Alekseev and Lampert, 2001; Ślusarczyk, Dawidowicz and Rygielska, 2005; Koch, von Elert and Straile, 2009). This deterioration of environmental conditions leads to the induction of clonal males and the formation of a single haploid egg within each ovary that requires fertilization (Hiruta and Tochinai, 2014). These sexual offspring are referred to as ‘resting eggs’ and enter diapause, facilitating genotype survival under challenging conditions (Hobaek and Larsson, 1990; Cáceres and Tessier, 2004; Kawasaki, Sakata and Namiki, 2004). The formation of sexually derived embryos has been shown to be more energetically demanding than subitaneous embryos (Lynch, 1983). During their formation, modifications to the maternal carapace occur over two moult cycles and produce a protective ephippum that

encases the embryos in two cuticle layers with a crystalline calcium phosphate layer between (Kawasaki, Sakata and Namiki, 2004; Hiruta and Tochinai, 2014). The ephippial structure is highly melanised and sclerotized which protects the eggs within from desiccation and predatory digestion but is permeable to some chemicals (Davison 1969; Gyllström & Hansson 2004; Navis et al. 2013). The ephippia are shed with the second moult and released into the surrounding environment, possibly floating, attaching to flora and fauna or sinking (Brendonck and Meester, 2003). Large numbers of *Daphnia* embryos fail to emerge from diapause each season and therefore accumulate in the sediment and form an ‘egg bank’. These can potentially hatch after abnormal amounts of time. Most recently *D. pulicaria* have been successfully hatched from ephippia dated to ca. 700 years old (Frisch, *et al.,* 2014).

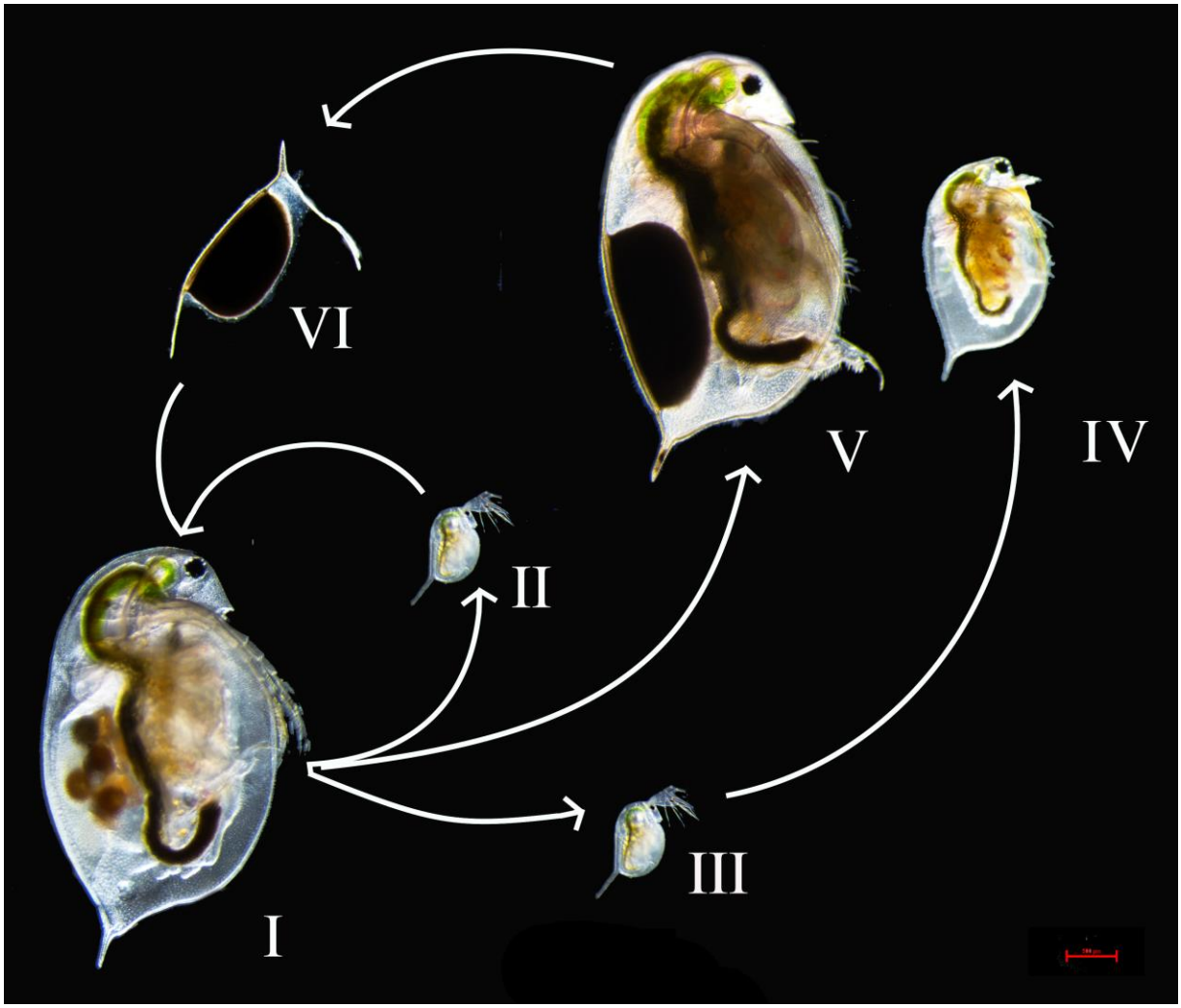


Figure 1-2 The reproductive cycle of *Daphnia magna.* Under favourable conditions mature females reproduce by parthenogenesis (I) giving birth to free-swimming clonal female offspring (II). In response to environmental stress, clonal male offspring (III) are produced and mature females switch to sexual reproduction (V) producing haploid eggs encased in an ephippial structure that require fertilisation by adult males (IV). The egg-containing ephippia is released during ecdysis (VI) and undergoes a period of diapause before hatching.

Despite a central importance in their life cycle, the process of diapause in *Daphnia sp.* remains poorly understood. Many stimuli are shared between the induction of males and resting egg formation but it has been suggested that the induction of male offspring is independent from ephippia production, evidenced by the presence of males and absence of ephippium in response to inductive stimuli (Hobaek and Larsson, 1990). Sex determination in *Daphnia magna* is environmentally determined during the later stages of oocyte development

immediately (under 12 hours) prior to oocyte release into the brood pouch (Ignace, Dodson and Kashian, 2011). Although the precise molecular mechanism is unknown, strong evidence suggests that it is hormonal in origin, mediated by the juvenile hormone methylfarnesoate (Olmstead and Leblanc, 2002). However, ephippia and male offspring production should be closely coupled by natural selection to ensure that males are present to fertilise the haploid sexual embryos when they are produced (Roulin *et al.*, 2013; Booksmythe *et al.*, 2018). The diversity of the cues involved, the prior need to change reproductive strategy, known genetic variation in response to these cues and cross-generational transfer of photoperiodic information suggest that the regulatory mechanisms behind diapause induction in *Daphnia sp.* are likely to be complex, but necessary, to couple entry to diapause to the multiple types of environmental deterioration these animals may experience (Deng, 1996; Alekseev and Lampert, 2001; Koch, von Elert and Straile, 2009). The mechanisms of diapause induction pertaining to the transduction of environmental stimuli are not addressed in this thesis.

1.4. The developmental biology of *Daphnia*

An excellent detailed description of the embryogenesis of asexually produced *D. magna* is already available (Gulbrandsen and Johnsen, 1990; Kotov and Boikova, 2001; Mittmann *et al.*, 2014) and has more recently been expanded to include male offspring (Toyota, Hiruta, *et al.*, 2016). However, there is limited information regarding the development of amphigonic, diapausing embryos (von Baldass, 1941; Zaffagnini, 1987; Hiruta & Tochinai 2014). From the outset of development, the two

embryos appear morphologically distinct. Histological sections of *D. pulex* reveal differences in the shape and colour of ovaries between the two forms of reproduction (Hiruta and Tochinai, 2014). Following progression through the ovary, oocytes are deposited into the dorsal brood pouch; diapause-destined oocytes are larger, more oval and lack the visible oil droplets present in subitaneous oocytes (Hiruta and Tochinai, 2014; Mittmann, *et al.*, 2014). The first two stages of embryonic development as defined by Mitmman, *et al.,* (2014) in parthenogenetic *D. magna* embryos encompass the early cell cleavages, resulting in a blastomere of 512 cells after 9 cycles of synchronous division. The authors describe gastrulation as the third stage of embryonic development and describe a polar gastrulation and epibolic movement of cells, although the number of cells in the embryo through these stages is not reported. Embryos progress through a further 5 stages defined by the formation of rudimentary appendages corresponding to the first and second antenna, mandible, and thoracic segments. Stages 9-11 involve overall changes to the shape of the carapace and elongation of the antennae. Hatching from the chorion and vitelline membrane occurs at stage 12 demarcating the end of embryonic development. Each daphniid is released from the brood pouch as a partially mobile neonate, retaining a thin cuticle layer that prevents full body extension. The cuticle is rapidly shed, and a now juvenile daphniid, fully resembling the adult form, progresses through 4-5 juvenile instars until reaching sexual maturity (Figure 1(II); Ebert, 1991; Mittmann *et al.*, 2014). Development rate is modulated by temperature, progressing faster at higher temperatures (Gulbrandsen and Johnsen, 1990).

There is limited evidence that the sexually derived embryos of *D. pulex* enter diapause at ~1,000 cells in early gastrula stage, but corresponding evidence from

other *Daphnia* species is absent (Baldass 1941). Gastrulation is a fundamental stage in embryonic development where cellular migration determines internal and external organization via the formation of distinct germ layers (Solnica-Krezel and Sepich, 2012). Embryonic diapause is conserved across cladocerans, branchiopods and copepods of the super family Centropagoidea as well as being relatively frequent in insect taxa (Orthoptera, Hemiptera, Lepidoptera, and Diptera) (Brendonck and Meester, 2003; Denlinger, Yocum and Rinehart, 2012; Baumgartner and Tarrant, 2017). Often, the precise cellular developmental stage of embryonic diapause is unknown. But the best described embryonic diapauses include those of *Artemia* and *B. mori* which both enter diapause in the late gastrula stage (MacRae, 2010; Moribe *et al.*, 2010). It is possible therefore that developmental arrest occurs in this phase of embryogenesis as key processes such as cell-lineage allocation must be completed to ensure successful resumption of development once diapause is determined (Arendt and Nübler-Jung, 1997). By contrast however, mammalian embryos enter diapause at the blastula stage, this typically represents the end of autonomous embryonic development in mammals beyond which uterine interaction is necessary (Renfree and Fenelon, 2017).

1.5. *Daphnia magna* as a model system for embryonic diapause

Whilst classical model species have been instrumental in furthering biological understanding they can fall short in some areas of specific scientific interest. For example, *Drosophila* is the powerhouse of genetic research, yet exhibits a comparatively weak diapause phenotype and thus may not be the most appropriate

model system to investigate the genetic basis of diapause (Poynton *et al.*, 2007; Denlinger and Armbruster, 2014; Russell *et al.*, 2017). *Daphnia sp.* has a long history in ecological research and is an established test species in toxicology (Shaw *et al.*, 2008). As a primary consumer, *Daphnia sp.* occupy an important trophic level in aquatic food webs, are highly abundant in freshwater ecosystems and are widely distributed throughout temperate climates (Roulin *et al.*, 2013). Their small size, short life cycle and relative ease of culture facilitates their use in laboratories and their parthenogenetic life cycle increases experimental reproducibility as defined genetic backgrounds can be maintained (Shaw *et al.*, 2008). The characterisation of the *D. pulex* genome (Colbourne *et al.*, 2011) established *Daphnia* as a model species for ecological and evolutionary genomics research, yet with its profound plasticity (Simon *et al.*, 2011), the use of this organism to investigate a wide variety of biological questions is growing. In this thesis, I utilise *D. magna*, which is a much larger animal for omics investigations and also benefits from maturing genomic resources including a published transcriptome (Orsini *et al.*, 2016) and anticipated genome sequence. Furthermore, functional genomic technologies are also available including RNAi (Kato *et al.*, 2011), and targeted gene editing with the CRISPR/Cas system (Nakanishi *et al.*, 2014).

Despite these resources, *D. magna* has been overlooked as a model system for the study of diapause. However, I suggest the plasticity of *D. magna* diapause makes it an ideal model system to study diapause. Through the manipulation of environmental conditions in the laboratory, both diapausing and subitaneous offspring can be produced under the same conditions allowing for direct comparisons between the two developmental trajectories, thereby disentangling developmental, genetic,

environmental and diapause specific changes. The ability of diapausing organisms to tolerate environmental conditions (e.g., reduced temperature and anoxia) that would be fatal to their non-dormant counterparts is well documented (Bale, 1996; Meller *et al.*, 2012a; Podrabsky and Hand, 2015; Wilches *et al.*, 2016; Hemmati, Moharramipour and Talebi, 2017). However, diapause also represents an interruption to the continuous lifecycle and for comparatively short-lived species can result in a significant expansion in the lifespan of an organism. The successful hatching of *Daphnia sp.* diapausing embryos is feasible after extended periods of dormancy amounting to ~60 years and ephippial offspring have been utilised in palegenomic experiments to investigate local adaption to predation, urban land use and parasite stress (Orsini, Spanier and De Meester, 2012). Furthermore the successful resurrection of *D. pulicaria* isolates have been reported after even greater periods in dormancy of up to 700 years (Frisch *et al.*, 2014). There is limited information regarding lifespan of *D. magna* but it is typically observed to be in the order of a few months (Pietrzak, Bednarska and Grzesiuk, 2010; Latta IV, Frederick and Pfrender, 2011; Li *et al.*, 2017). Therefore, diapause in *D. magna* represents a potential to expand organism lifespan by several orders of magnitude and is thus an appropriate model system to unravel the mechanisms behind extended diapause.

1.6. Environmental omics approaches: overview and techniques

Diapause is a highly complex phenotype that manifests at the organism level altering physiology, life history, metabolism and gene expression across an entire developmental trajectory that is often condition dependant (Denlinger, 2002; Denlinger and Armbruster, 2014; Kang *et al.*, 2016). Therefore, it is not easily

associated with DNA variation alone but rather is likely the product of the complex interactions among transcribed elements, functional proteins and metabolism. For each of these layers of cellular events, a corresponding ‘omics’ technology has emerged that strives to comprehensively characterise and quantify all of the perturbations associated with a given phenotype. Transcriptomics aims to capture information on the identity and quantity of the full range of functional RNAs, or genes expressed by an organism under a given condition (Wang, 2001). Proteomics aims to describe the entire protein complement of a sample, including primary peptide sequence as well as post-translational modifications and may provide direct functional evidence of gene expression that is otherwise not achievable by considering the transcriptome alone (Aebersold and Mann, 2003). Further downstream is metabolomics, which is the quantitative determination of the low molecular weight metabolites present within a given biological sample, ranging from tissue or biofluid to whole organism homogenates (Fiehn, 2002). The ability to capture information about the current molecular status of an organism and how this corresponds to observable phenotypes (e.g. development), is a powerful tool in the elucidation of molecular networks and gene function (Macel, van dam and Keurentjes, 2010).

Metabolomics and transcriptomics are the two orthogonal data types collected and investigated in this thesis in the context of embryonic diapause in *D. magna.* They were selected as both disciplines aim to comprehensively (without bias) define all the molecular changes associated with a given experimental condition or phenotype. Typical shotgun proteomic experiments require extensive preparation including enzymatic digestion, and generally include further separative techniques to retain low

abundance proteins, significantly increasing time and cost of analyses (Chandramouli and Qian, 2009; Giansanti *et al.*, 2016). Furthermore, current methodologies are still limited by starting sample input (Feist and Hummon, 2015) and thus due to these limitations, particularly given the challenges regarding collection of sufficient biomaterial from *Daphnia* embryos, proteomics was not included in this thesis.

The analytical approach for the study of the transcriptome applied during this thesis is the application of next-generation sequencing technologies to directly quantify expressed transcripts through cDNA sequencing (RNA-Seq). Illumina platforms are currently the most utilised in RNA-Seq experiments and were thus applied during my investigations owing to depth of transcriptome coverage, low error rate and low cost (Schadt, Turner and Kasarskis, 2010; Nguyen *et al.*, 2018).

The metabolome is diverse in chemical and physical properties consisting of thousands of molecules with varying atomic composition, polarity, volatility, concentration and structure (Macel, van dam and Keurentjes, 2010). Thus, the task of metabolomics is a formidable one and highly sophisticated analytical techniques and a variety of bioinformatic and statistical approaches are utilised (Dunn, Bailey and Johnson, 2005; Han *et al.*, 2008; Lei, Huhman and Sumner, 2011). Despite significant advancements in the field, currently no one analytical platform is capable of capturing the entire metabolome (Gertsman and Barshop, 2018).

Direct infusion mass spectrometry (DIMS) analysis is performed on biological samples without prior separation, therefore benefits from reduced acquisition times (Han and Gross, 2005). This thesis implemented spectral-stitching nano electrospray ionization (nESI) DIMS, whereby the nESI source is directly coupled to a high-resolution Orbitrap mass spectrometer. DIMS-based metabolomics has been

previously used to detect metabolic perturbations in *D. magna* in response to xenobiotic exposure (Poynton *et al.*, 2011; Taylor, White and Viant, 2017) demonstrating the suitability for use with low biomass samples of less than a milligram of tissue. Furthermore a recently published complete workflow, including data analysis implemented in Galaxy has demonstrated that this approach is both reproducible and accessible, opening it to wider research communities (Southam *et al.*, 2017). However, this approach is not without its limitations. Full-scan DIMS data lack structural information and thus unequivocal compound identification is a particular challenge (Gertsman and Barshop, 2018). Furthermore, without prior knowledge of peaks of interest and the introduction of standards, quantitation is based on relative abundance of peaks between samples, therefore care must be taken to ensure samples are as comparable as possible (Fiehn, 2002).

No single 'omics' technology in isolation is likely to comprehensively elucidate the underlying molecular events behind biological phenotypes (Nie *et al.*, 2007; Zhang, Li and Nie, 2010). It is therefore becoming increasingly apparent that investigations involving several technologies (i.e. multi-omics) which strive to integrate observations at multiple levels will likely prove useful. In this thesis I present both metabolomic and transcriptomic datasets. However, given the remit to capture as many elements (i.e. genes or metabolites) as possible, omics data are highly multivariate and advanced computational resources and statistics are required for combined analyses, with approaches varying from deriving basic correlations (Hasin, Seldin and Lusis, 2017), through to advanced network approaches (Bersanelli *et al.*, 2016). The statistical analysis presented within this thesis is limited to the tandem analysis of both data

types, which is a necessary starting point to comprehensively understand the nature of the data prior to further analysis including integration into a multi-omics dataset. PCA is an unsupervised multivariate technique that is particularly useful in the interpretation of high-dimensional datasets whereby a drastically reduced set of new ‘principal components’ are computed from linear combinations of original variables (Abdi and Williams, 2010). Furthermore, this transformation is achieved without the loss of information such that the original variance within the dataset is preserved (Bonner and Beyene, 2017). The first component describes the maximum amount of variation within the dataset with subsequent orthogonal components describing smaller portions (Bro and Smilde, 2014). In this way, the respective expression and intensity data of thousands of genes and metabolite peaks is represented in a handful of principal components (PCs). For each PC, original variables are assigned a ‘loading’ based on how much variation they contribute and individual samples are assigned a coefficient or ‘score’, which can then be plotted (Jolliffe, 2002). In scores plots, samples that are similar in terms of the values of original variables will also have comparable scores and inherently cluster together. Thus, patterns of gene expression and metabolite intensities can be assessed across all samples in an unbiased manner and significant separation of sample groups along component axes can also be determined by the application of univariate statistics.

The transition from observations of sample clustering in PC scores plots to a biological interpretation can be made through assessing variable loadings for the relevant PC. Peaks or metabolites with large loadings are those contributing most to the aforementioned sample separation. However, given that PCs are a linear combination of all original variables, each with a loading, deciding which variable

loadings to consider becomes more problematic in large datasets (Bonner and Beyene, 2017). Sparse principal component analysis (sPCA) is an alternative approach to PCA where PCs are estimated from a subset of original features rather than the total (Zou, Hastie and Tibshirani, 2006). SPCA models are optimised so that the variance explained is maximised and the number of variables with non-zero loadings is drastically reduced, however these variables are predominantly those responsible for sample separation along PCs and thus likely to be biologically meaningful (Zou, Hastie and Tibshirani, 2006). This reduced set of original variables means the loadings results from sPCA are more interpretable than PCA (Bonner and Beyene, 2017). However, a notable limitation is that additional parameters that determine loadings sparseness need to be arbitrarily chosen prior to modelling (Erichson *et al.*, 2018). Drawing on the strengths of both methods, a combination of PCA to visually represent variation in both metabolomic and transcriptomic data and sPCA for selection of genes and metabolites of interest for further study were implemented in this thesis. Additional multivariate statistics implemented included the application of ANOVA-simultaneous component analysis (ASCA) (Smilde *et al.*, 2005) to the analysis of metabolomics data. This was utilised, with accompanying permutation testing (Vis *et al.*, 2007), to assess peak intensities that changed significantly with embryo type, time or the interaction thereof. The majority of the statistical analysis of transcriptomic data was performed through well-established pipelines, particularly *DEseq2* (Love, Huber and Anders, 2014). In addition, a variety of univariate statics i.e. Analysis of Variance (ANOVA) and Student's t-tests were also implemented. Where applicable, appropriate correction was applied to account

for multiple hypothesis testing (e.g. false discovery rate correction; Benjamini and Hochberg, 1995).

1.7. Thesis outline and its aims

This thesis aims to present *Daphnia magna* as a model system to understand the molecular events of embryonic diapause. I therefore identified several key objectives that are addressed in subsequent chapters to achieve this aim:

- Determine the appropriate experimental conditions for the reliable induction of sexual reproduction and production of viable diapause-destined offspring in *D. magna* (Section 3).
- Identify cellular differences between diapause-destined and subitaneous development in *D. magna*, including the stage of developmental arrest, thus allowing me to clearly define the temporal phases of pre-diapause and diapause (Section 4).
- Assess the molecular differences between directly-developing and diapause-destined embryos in terms of metabolite composition and expressed genes, in addition to the relative changes over pre-diapause development (Section 4).
- Characterise changes to metabolism and gene expression during the initiation of diapause, comparing sexually derived embryos only, sampled before and after the onset of developmental arrest (Section 5).
- Assess temporal dynamics of the transcriptome and metabolome throughout the phases of pre-diapause and diapause, including metabolic changes occurring between samples collected in ‘early’ and ‘late’ diapause (Section 6).

2. Materials and methods

Unless otherwise stated all data analysis was performed with R (v3.5.1) using R studio (v1.1.383), and all plots were generated with *ggplot2* (v3.1.0) package.

2.1. *Daphnia magna* isolate information

Clonal isolate 'L7' was obtained from a sediment core collected from Lake Ring, Denmark (55° 57' 51.83'' N, 9° 35' 46.87'' E). L7 was hatched from a core section 7.5 cm deep, dated roughly to 1990-1985 within the eutrophic stage in the lake's history (Cuenca Cambronero and Orsini, 2018). Clone 'B2' originated (ca. 2006) from a laboratory culture obtained from the University of Reading and is the routine *D. magna* clone utilised for studies by the University of Birmingham *Daphnia* group. Both cultures were maintained in the laboratory for numerous clonal generations prior to any experiments performed.

2.2. General culturing conditions

The *D. magna* were maintained in glass beakers (1 L) at a constant temperature (20 ± 1°C) with a light:dark photoperiod of 16:8 h. Cultures were maintained in ground water collected from a bore hole (60 m; pH 7.3-7.8) on the campus grounds of the University of Birmingham that draws water from the underlying Sherwood Sandstone aquifer (Table 2-1). Water was collected twice-monthly, stored (10°C) and equilibrated (24 h) to culture temperature before use. Culture medium was replaced once-weekly. Cultures were fed daily with a suspension of *Chlorella vulgaris* (Section 2.4) and offspring were removed via filtration. Stock cultures of

20 *Daphnia* L^{-1} were initiated every 2 weeks with neonates (<24 h) of the third brood or later from previous culture.

Table 2-1 Summary of the major ionic constituents of water drawn from the Sherwood Sandstone aquifer via bore holes on the University of Birmingham campus. Source: Hydrogeological Services International Ltd report (2001) "The University of Birmingham Groundwater Supply – Test Pumping of Environment Agency Borehole".

Determinant	Concentration (mg L^{-1})
Ca	72.5
Na	12.7
Mg	28
K	>2
Alkalinity - $CaCO_3$	22.2
Harness- $CaCO_3$	28.15
SO_4	46.8
Cl	27.5
NO_3	16.8
NO_2	<0.01
NH_3	<0.05

2.3. Culturing for the production of diapausing embryos

The *D. magna* were raised at a density of 62 *Daphnia* L^{-1} consisting of 50 female and 12 male individuals. All cultures were fed daily with *C. vulgaris* at 0.03 mgC *Daphnia*$^{-1}$ (corresponding to 1.6 mC L^{-1}). Offspring and ephippia produced each day were removed via filtration, all other culturing parameters were as described in Section 2.2.

2.4. Culturing and preparation of *Chlorella vulgaris* algal feed

The unicellular green algae *C. vulgaris* was cultured for use as feed for *D. magna*. Algal cultures were maintained in suspension in Bold's basal medium (Bold, 1949; Table 2-2) in conical flasks (2.5 L) under constant aeration and full spectrum light.

Cultures were renewed twice weekly through seeding with ca. 20 mL of previous culture in fresh autoclaved media.

Table 2-2 Stock composition of Bold's basal medium

Number	Compound	Empirical formula	Concentration (gL^{-1})	Volume prepared (L)	Volume (mL) added to prepare 5 L
1	di-Potassium hydrogen orthophosphate	$K_2 HPO_4$	7.5	1	50
2	Potassium di-hydrogen orthophosphate	KH_2PO_4	4.93	1	50
3	Magnesium sulphate	$MgSO_4 \cdot 7H_2O$	7.5	1	50
4	Sodium nitrate	$NaNO_3$	25	1	50
5	Calcium chloride	$CaCl_2 \cdot 2H_2O$	2.5	1	50
6	Sodium chloride	NaCl	2.5	1	50
7	EDTA tetrasodium salt	EDTA - Na_4	50	0.1	5
	Potassium hydroxide	KOH	31		
8	Ferrous Sulphate	$FeSO_4 \cdot 7H_2O$	4.98	0.1	5
	conc. Sulphuric acid	H_2SO_4	10 ml L^{-1}		
9	Boric acid	H_3BO_3	11.42	0.1	5
10	Zinc sulphate	$ZnSO_4 \cdot 7H_2O$	14.12	0.025	0.5
11	Manganese chloride	$MnCl_2 \cdot 4H_2O$	2.32	0.025	0.5
12	Copper sulphate	$CuSO_4 \cdot 5H_2O$	2.52	0.025	0.5
13	Cobaltous nitrate	$(Co(NO_3)_2 \cdot 6H_2O$	0.8	0.025	0.5
14	Sodium molybdate	$Na_2MoO_4 \cdot 2H_2O$	1.92	0.025	0.5

Algal feed was prepared by the removal of a known volume of culture suspension, the optical density (A_{440nm}) of a 1:10 dilution (dH_2O) was obtained. A final feed suspension of optical density 0.8 (A_{440nm}) corresponding to 800 $mgCL^{-1}$ was produced following centrifugation (3000 rpm) and resuspension of the algal pellet in a known volume (dH_2O). Feed was prepared twice weekly and stored (4°C) between use.

2.5. Selection of isolate and culturing parameters for the laboratory induction of diapause in *Daphnia magna*

The *D. magna* were raised from <24 h neonates in 1 L vessels at a density of 20 *Daphnia* L^{-1} (L7 and B2), 25 *Daphnia* L^{-1} and 62 *Daphnia* L^{-1} (L7 only). The 25 *Daphnia* L^{-1} cultures were initiated with either 25 female individuals or 20 female and 5 male individuals (n = 3). All 62 *Daphnia* L^{-1} cultures contained 50 female and 12 male individuals. All cultures were fed daily with *C. vulgaris* (Section 2.4) at different algal rations corresponding to a "full" (0.08 mgC *Daphnia*$^{-1}$) and "reduced" (0.03 mgC *Daphnia*$^{-1}$) diet ration (Table 2-3). Offspring and ephippia produced each day were removed via filtration, counted and assessed using a dissection microscope. Neonates were sexed and ephippia were opened with the use of fine forceps to assess viability. Ephippia containing at least one embryo were considered viable. Significant differences in total neonate and ephippia output between isolates and food ration were determined using two-way analysis of variance (ANOVA). Differences in the age of first release of both offspring types between isolates and food ration were determined using a factorial ANOVA. A generalised linear mixed-effects model (GLM) with binomial errors and logit link function was used to analyse the effect of food ration and isolate on offspring sex ratio. Comparing increased culture density, a two-sample Student's t-test was used to assess differences in neonate output and ephippia produced.

Table 2-3 Concentrations of *Chlorella vulgaris* algal food used in the selection of culturing parameters for the induction of sexual reproduction in *Daphnia magna*

Number of individuals	Amount of *C. vulgaris* *Daphnia*$^{-1}$ (mgC)	Volume of *C. vulgaris* (µL)	Final *C. vulgaris* concentration in beaker (mgC L^{-1})
20	0.08	2000	1.60
25		2500	2.00
62		6000	4.80
20	0.03	640	0.51
25		800	0.64
62		2000	1.60

2.6. Assessment of the longevity of laboratory-derived diapausing embryos

Male (n = 3) and female (n = 5) neonates (< 24 h) were collected from stock cultures and placed in glass jars containing 5 individuals of the same sex in bore water (200 mL). Cultures were fed daily with *C. vulgaris* at a full (0.08 mgC *Daphnia*$^{-1}$) or reduced (0.03 mgC *Daphnia*$^{-1}$) diet ration and any offspring were removed daily via filtration. Cultures were observed daily for mortality until all animals of each replicate beakers had died. The relationship between gender and food ration and the age of culture mortality was determined via two-way ANOVA.

2.7. Assessment of the viability of laboratory-derived diapausing embryos

Ephippia were collected daily from cultures of sexually reproducing *D. magna* (Section 2.3) over a period of 2 y and transferred to 15 ml amber polypropylene conical tubes (Greiner Bio-One) in approximately 7.5 ml of culture water leaving sufficient air gap, and stored at 4°C in darkness until use. Ephippia of various ages (14 d, 30 d, 60 d and 728 d) were then opened on the same day via manual

decapsulation. Embryos of the same age were transferred to a 96-well plate (Corning®); each well contained the embryo from a single ephippia in bore water (ca. 100 uL), where two embryos were obtained they were transferred in pairs and considered as a single replicate (therefore, n = 96). Plates were incubated (20 ± 1°C) under constant illumination with full spectrum light and observed every 24 h–48 h for hatched neonates. Upon hatching, neonates were transferred to glass vials in bore water (100 mL) and fed daily with *C. vulgaris* (0.08 mgC *Daphnia*$^{-1}$). Hatchlings were observed daily for mortality or their first brood of neonates. Separate linear regression models determined the relationship between the age of ephippia the proportion of successful hatching and survival to reproductive maturity.

2.8. Temporal definition of the major morphological landmarks of embryonic diapause in *Daphnia magna*

Prior to the onset of reproduction (ca. 7 d) *D. magna* were collected from sexual cultures (Section 2.3) and transferred to a 24-well plate (Corning®) in mating pairs containing bore water (1 mL). Plates were maintained under environmental conditions described previously (Section 2.2) with daily replacement of water and feed (0.03 mgC *D*aphnia^{-1}). Daphnia were observed and imaged every 12 h over 4 d. For the observation of embryo development outside the maternal brood pouch, embryos were removed from females raised in the same manner within 2 h of ovulation, transferred to new wells maintained in the same environmental conditions with the absence of algae and imaged every 12 h for 3 d. All imaging was conducted using a Nikon SMZ800 microscope and NIS-Elements F (v4.00.00; Nikon Instruments UK Ltd). Raw images were compiled into figures maintaining original scale using GNU Image Manipulation Program (GIMP; v2.8.22).

2.9. Estimation of *Daphnia magna* sexual embryo dry mass

The dry mass of 1.5 ml Eppendorf tubes was obtained in advance after incubation at 70°C overnight with open lids using an analytical balance (±0.1 mg; Sartorius). *D. magna* sexually produced embryos (produced as described in Section 2.3) were collected, manually decapsulated, combined into a 200 embryo pool (n = 5) and transferred to dry tubes in a minimum volume (dH_2O). The approximate wet mass was obtained immediately, and dry mass was obtained after overnight incubation (70°C). Embryo mass was estimated by subtraction of the mass of empty tubes from the final dry mass of tubes including embryos. Single embryo dry mass was further estimated by correction for the number of eggs (200) in the embryo pool.

2.10. Collection of diapausing embryos for biomass optimisation of nESI-DIMS metabolomics

D. magna sexually produced diapausing embryos were obtained from ephippia shed by running sexual cultures (Section 2.3) over several months. Ephippia were collected daily and transferred to 15 mL amber polypropylene conical tubes (Greiner Bio-One) in approximately 7.5 mL of culture medium leaving sufficient air gap. Tubes were stored at 4°C in darkness until use. When sufficient numbers of ephippia were obtained, all tubes were opened at the same time, embryos were removed from ephippia by manual decapsulation and combined into a large pool to distribute biological variation due to duration of storage. Embryos were then individually transferred to droplets (100 µL) of LC/MS Optima™ grade H_2O (Fisher Scientific) and divided into samples of 50, 25, 10, 5 and 1 embryo (n = 8). Additional quality control (QC) samples were prepared in the same manner and consisted of 50 embryos per

sample (n = 12). Each embryo sample was then washed several times via sequential transfer in a maximum volume of 6 µL through 4 droplets (100 µL) of LC/MS Optima™ grade H_2O in order to minimise contaminants originating from culture media (e.g., salts, ephippial remnants and, algae). Samples were then transferred in LC/MS Optima™ grade H_2O (60 µL) to soft tissue homogenising tubes (2 mL; Precellys), flash frozen in liquid nitrogen and stored at -80°C.

2.11. Identification of cellular diapause arrest and alternative developmental trajectories ¬

All experiments were performed at Ruhr Universität Bochum, Germany in collaboration with Luxi Chen and Dr Linda C. Weiss. All data collections and analysis were performed with Luxi Chen. Dr Linda C. Weiss oversaw all work and conceived of this study.

All animals were raised as described previously with the following alteration: a modified version of ADaM medium was used for culture water (2.3 mg $CaCl_2$•2 H_20; 2.2 mg $NaHCO_3$; 0.1 mg SeO_2; 12.5 mg sea salt, filled up to 1 L with ultrapure water; Klüttgen *et al.*, 1994). Mature, reproducing females raised in 62 *Daphnia* L^{-1} cultures (Section 2.3) were transferred to 24-well plates and observed every 15 min for evidence of exuviae and release of embryos into the maternal brood pouch. Sampling time points (n ≥ 3) were every 2 h following ovulation for 24 h with additional collections at 30 h, 36 h, 48 h post-ovulation and 2 h, 36 h and 2 weeks following maternal release of the ephippium. Upon release from mother, ephippia

¬ Work presented in this section was performed in collaboration with Dr Linda Weiss and Luxi Chen at the Ruhr Universität Bochum, Germany. And has been accepted for publication in BMC Cell Biology: "Mitotic activity patterns and cytoskeletal changes throughout the progression of diapause developmental program in *Daphnia*" Luxi Chen; Rosemary E. Barnett; Martin Horstmann; Verena Bamberger; Lea Heberle; Nina Krebs; John K. Colbourne; Rocío Gómez; Linda C. Weiss

were maintained under the same environmental conditions until collection. Embryos were removed from the mother or ephippial casing using fine forceps and fixed in 4% PFA-TX (formaldehyde 37%; Merck, Germany; diluted in phosphate buffered saline (PBS) 0.1 M, pH 7.4; with 0.05% Triton X; Serva, Germany) for 15 min. Following fixation eggs were washed (3 x 5 min) in PBS and individually transferred to a Polysine™ microscope slide (Thermo Scientific). After removal of excess PBS, eggs were squashed in VECTASHIELD® Mounting Medium with 4',6-diamidino-2-phenylindole (DAPI; Vector Laboratories, Inc., U.S.A.) using the cover slip and sealed with clear nail varnish. Samples were stored in darkness (4°C) until imaging. DAPI-stained sample slides were imaged using Zeiss Axiophot fluorescent microscope using 358/461 nm excitation/emission filter, 20x magnification. Images were collected with an Olympus XC10 monochrome digital camera using CellSens imaging software (Olympus, Germany). Single images of all squashed cells from each embryo were acquired from several images using the stitching function and cell numbers were determined using the counting function in CellSense. Nuclei were assumed to correspond to single cells (absence of polyploidy).

Logistic curves of cell number against time were fitted to data according to Equation 2-1.

Equation 2-1 y= embryo cell number, a_1 = the first parameter (first an approximation of the maximal cell number, later value is replaced by the modeled parameter), a_2=the second parameter of the logistic function, a_3= the third parameter of the logistic function, t=hours post ovulation)

$$y = \frac{a_1}{1 + e^{-(a_2 + a_3 \times t)}}$$

An initial model was fitted to count data using the logit function (Equation 2-2) where a_1 is first an approximation of maximal cell number (e.g. 4000 cells in sexually produced and 8000 cells in asexually produced embryos). Intercept and slope

parameters were extracted from initial model, and then used as approximated values for the parameters a_2 and a_3.

Equation 2-2

$$\boldsymbol{lm}\left(\boldsymbol{logit}\left(\frac{\boldsymbol{y}}{\boldsymbol{a_1}}\right) \sim \boldsymbol{t}\right)$$

Final coefficients were established from starting approximate parameters using the *nls*() function. Original data points and the modelled growth curves were plotted in a single plot. The cell division rate was calculated as the derivative with respect to time (Equation 2-3).

Equation 2-3

$$y' = \frac{\partial y}{\partial t}$$

2.12. Collection of embryos for nESI-DIMS and RNA-Seq analysis

Prior to the onset of reproduction (ca. 7 d) *D. magna* were collected from sexual cultures (Section 2.3) and transferred to a 24-well plate (Corning®) in mating pairs containing bore water (1 mL) and *C. vulgaris* (0.08 mgC *Daphnia*$^{-1}$). Plates were maintained under environmental conditions described previously (Section 2.2). Plates were observed every 15 min for a period of 24 h – females that failed to ovulate within 24 h were not sampled from, thus females spent no longer than 24 h in well plates. Sampling time points are shown in detail in Figure 2-1. Asexually produced embryos were collected at 2 h, 10 h, 24 h (metabolomics only), and 48 h post-ovulation. Sexually produced embryos were collected at; 2 h, 10 h, 24 h (metabolomics only), 48 h, 72 h, 82 h, 98 h (metabolomics only), 120 h, 408 h, ~17592 h (2 y; metabolomics only). 2–48 h samples were collected from the maternal brood pouch in the following manner: 2 h samples were obtained from females

directly from well plates. >2 h samples were obtained from females that were transferred upon ovulation to glass vials in bore water (100 mL) and fed daily (where appropriate) with *C. vulgaris* (0.08 mgC *Daphnia*$^{-1}$). 74–408 h samples were obtained after the ephippia had been shed by the female, upon release ephippia were transferred to glass vials in bore water (100 mL) and maintained under culture conditions until collection. ~17592 h samples represent embryos 2 y post-ephippia release, these embryos were produced in the same manner but upon collection were stored in 15 mL amber polypropylene conical tubes (Greiner Bio-One) in approximately 7.5 mL of culture medium leaving sufficient air gap, tubes were stored at 4°C in darkness until use and extracted 2 y (728 d) post-release but the number of hours are not exact and reported for consistency.

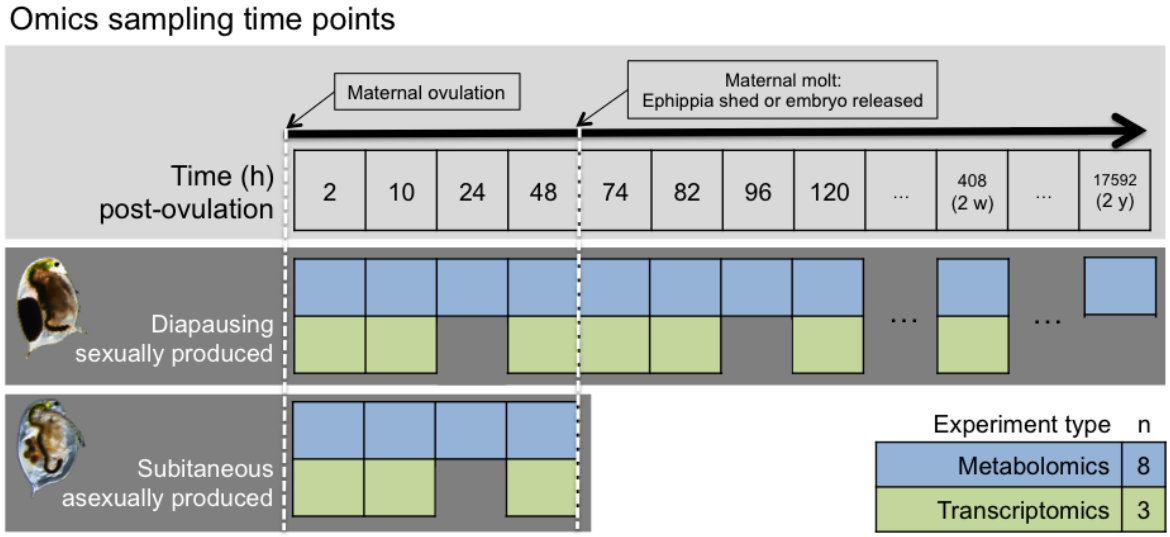


Figure 2-1 Sampling scheme for the collection of *Daphnia magna* embryos for transcriptomic and metabolomic analysis. Time (h) post-ovulation of sampling time points are indicated, shaded boxes (refer to legend bottom right) indicate collection for metabolomics (blue) and transcriptomics (green).

Embryos were obtained from the maternal brood pouch via dissection and from ephippia via manual decapsulation under a light microscope using fine forceps. Embryos were then individually transferred to droplets (100 µL) for metabolomics

LC/MS Optima™ grade H_2O and for transcriptomics molecular grade DNA/RNAse free H_2O was used. Each embryo sample was then washed several times via sequential transfer through 4 droplets (100 µL) of H_2O (respective grades). Samples were then transferred in H_2O (6 µL) to soft tissue homogenising tubes (2 mL; Precellys), flash frozen in liquid nitrogen and stored at -80°C. Single embryos (n = 3) were collected for transcriptomics, for metabolomics (n = 8) 10 embryos per sample were collected. Between additions tubes were stored (-80°C) and kept in liquid nitrogen whilst sampling took place. Collection occurred over a period of 5 months (960 embryos total), however embryos were added to tubes in a stepwise fashion such that all biological replicates contained embryos from early and later sampling events. Collection of all timed samples was randomised between diurnal and nocturnal events; as such I anticipate that zeitgeber effects if present are randomly distributed throughout.

2.13. Transcriptome sequencing

2.13.1. RNA extraction, library preparation and sequencing of *Daphnia magna* embryos

Embryo samples were collected as described in Section 2.12. mRNA (poly A+) was extracted, reverse transcribed and amplified using the REPLI-g® Single Cell RNA Library kit (Qiagen; Cat No. 150073). Reactions were performed in the original collection tube: embryos were manually burst with extra-long filtered pipette tips (5 µL) in lysis buffer provided (7 µL). All other steps were performed according to the manufacturers' protocol. Amplification nucleotides and primers were removed using Agencourt AMPure XP Beads (100 µL) after incubation (10 min; RT) tubes were

placed in a magnetic rack and supernatant was discarded. Two wash steps in 80% EtOH (200 µL) followed by drying (3 min; RT) were performed. cDNA was eluted from beads in EB buffer (60 µL; Qiagen), 10, 50 and 100-fold dilutions were quantified in duplicate using the Infinite® 200 PRO plate reader (TECAN) against a Promega™ lambda DNA concentration curve (n = 3; 10, 5, 2.5, 1.25, 0.625, 0.3125, 0.156250, 0 ngµL^{-1}) with SYBRgreen (LifeTech; S7563). Samples were diluted to 100 ngµL^{-1} and sheared into ~250 bp fragments using a Biorupter® Pico ultrasonicator with integrated cooling module (Diagenode; B01060010) with 6 cycles of 30 s at 90 s intervals. cDNA fragment size was assayed on a 2200 TapeStation (Agilent) using High Sensitivity D5000 screentapes and reagents according to manufacturers' instructions. Between 300-350 bp paired-end Illumina compatible libraries were prepared using the KAPA HyperPlus Library Preparation Kit (Roche; KK8504) without amplification. Post library construction clean-up was performed again using Agencourt AMPure XP Beads (88 µL) and eluted in EB buffer (100 uL). Library sizing and initial quantification was assayed again on 2200 TapeStation. All 10,000, 100,000, and 1,000,000-fold library dilutions were quantified using by qPCR using the NEBNext® Library Quant Kit for Illumina® (New England Biolabs) and LightCycler® 96 (Roche) according to kit instructions modified with an extended (2 min) amplification step. Libraries were pooled to a final multiplex (20 nM) and paired-end sequenced (100 bp) on two lanes of a HiSeq4000 (Illumina) by BGI Europe (Copenhagen, Denmark).

2.13.2. Illumina data processing

Quality of raw sequencing reads was assessed using FastQC v0.11.7 (Andrews S, 2010). Sequenced reads were trimmed using Nextflow computational pipeline

v0.33.5 (DI Tommaso *et al.*, 2017) implementing Trimmomatic v0.32 to remove sequencing adapters, and lead and tail bases that fell below Q30 phred score. Only reads greater than 48 b in length were retained, and read quality was again assessed post-trimming using FastQC. Sequences were aligned to the *D. magna* reference transcriptome (Orsini *et al.*, 2016), using both primary transcripts (29,121) and alternative splice variants (84,898) available from wFleaBase (http://arthropods.eugenes.org/EvidentialGene/daphnia/daphnia_magna/Genes/early access/). Read alignment was performed using Salmon v0.8.2 (Patro *et al.*, 2017), and included sequencing and GC content bias correction, using a k-mer value of 19, specifying library type (inward-facing non-stranded) and discarding read pairs that were inconsistent with this mapping orientation. Resulting Salmon *.quant.f* files were prepared for differential expression (DE) analysis using *tximport* to summarise transcript expression to gene level (Soneson, Love and Robinson, 2016).

2.13.3. Differential gene expression analysis

Reads with low counts (< 10 across all samples) were removed. Count normalisation and DE analysis was performed using *DESeq2* (Love, Huber and Anders, 2014) using R (v3.5.1) within RStudio© (v1.1.383, RStudio, Inc.). For the assessment of alternative development (Chapter 4), the design formula included embryo type, time (h) post-ovulation and the interaction thereof. A likelihood ratio test was performed against a reduced model excluding one of the three terms to identify DE genes. For the investigation of sexual data only (Chapter 6), the design formula included only time (h) post-ovulation. For the assessment of DE genes at specific time points (Chapter 5) direct pairwise comparisons were performed. Unless otherwise stated, significant DE was determined based on a false discovery rate (FDR) < 1% (< 0.01).

This arbitrary cut-off is considered very stringent by comparison to the standard level of 5% but chosen to reduce the Type I error (to falsely infer significant) and to help interpret the process of diapause by focusing on a smaller set of DE elements. Resulting gene count matrices were log-normalised prior to principal component analysis, generation of heatmaps and any clustering.

2.13.4. *Daphnia magna* functional gene databases and gene ontology (GO) over-representation analysis

In order to attribute biological functions to DE genes two custom annotation files were obtained. The first file, assigning putative protein functional annotations to the longest transcript of each *D. magna* locus (29,121 genes; Orsini et al., 2016) using *patherScore* (Mi *et al.*, 2017), utilized the library of protein classifications curated in the Protein ANalysis THrough Evolutionary Relationships (PANTHER) database. This tool predicts the functions of uncharacterised genes based on sequence similarity to gene sub-families as represented by statistical hidden Markov models (Mi *et al.*, 2013). Only 19,244 genes of the initial 29,121 (66%) input were successfully annotated using this tool. 235 unique biological processes, 182 molecular functions and 64 cellular component entries were identified. A second annotation was performed assigning genes expressed in the *D. magna* transcriptome to orthologous groups using OrthoDB v9 (Zdobnov *et al.*, 2017). In this instance genes related to each other by decent from a common ancestor are grouped together. 16,852 genes of the initial 29,121 (59%) were assigned to 8,692 unique groups among the Metazoa.

Gene ontology (GO) annotation over-representation analysis was performed using PANTHER gene list analysis tools (accessible at: http://www.pantherdb.org) on DE

gene lists (FDR < 0.1). The arbitrary threshold was set to be less stringent so to reduce the Type II error (to falsely infer that a set of differentially regulated genes is comprehensive) and to not artificially inflate the stringency of the over-representation analysis. The background gene lists consisted of all genes detected (counts across samples > 10) in the final count matrix. Significant over-representation was determined by an FDR of < 0.05 based on Fisher's exact test with Benjamini-Hochberg *p*-value, false discovery rate (FDR) correction (Benjamini and Hochberg, 1995; Thomas *et al.*, 2006). Patterns of gene expression were represented in heatmaps using *pheatmap* v1.0.10, (where included) clustering of heatmap rows (genes) or columns (samples) was based on Euclidean distance.

2.14. Direct infusion mass spectrometry (DIMS) metabolomics analysis

Samples collected for both the biomass optimisation study (Section 2.10) and final analysis (Section 2.12) were analysed in the same manner with the exception that QC samples for the biomass optimisation study were prepared from a pool of 50 embryo samples only.

2.14.1.1. *Sample extraction*

Biphasic solvent extraction of embryo samples was performed as previously described to obtain polar and non-polar fractions containing hydrophilic and hydrophobic metabolites respectively (Bligh, E.G. and Dyer, 1959; Wu *et al.*, 2008). Samples were homogenised following addition of MeOH (320 µL; 4°C) and H_2O (68 µL; 4°C) using Precellys™ 24 homogeniser (6400 rpm; 2 x 10 s; Stretton Scientific Ltd, UK). Homogenate was transferred to glass vial and further aliquots of

H_2O (160 µL) and $CHCl_s$ (320 µL) were added, the latter using a glass syringe with metal plunger. Phases were separated after vortexing (30 s), incubation (10 min; on ice) and centrifugation (4000 rpm; 10 min; 4°C) with minimal disruption to the interface layer using a glass syringe. Two equal aliquots (150 µL polar, 100 µL non-polar) were taken from each phase. Polar extracts were dried overnight using a centrifugal concentrator (Savant SpeedVac concentrator, Thermo Scientific); non-polar extracts were dried under anoxic conditions using a nitrogen sample concentrator (Techne, Bibby Scientific, UK). Extractions were performed in multiple batches of biological samples and included extraction 'blank' samples that contained no biological material. All extracts were stored (-80°C) prior to analysis.

2.14.1.2. *Sample preparation*

Polar extracts for positive ion analysis were resuspended in MeOH: H_2O, 4:1 with formic acid ((0.25%), 35 µL) and for negative ion analysis MeOH: H_2O, 4:1 with ammonium acetate ((20 mM), 35 µL), mixed thoroughly (30 s) and centrifuged (15,000 rpm; 15 min; 4°C) to remove particulate contaminants. Non-polar extracts were resuspended in MeOH:$CHCl_3$, 2:1 with ammonium acetate ((5 mM), 60 µL) mixed thoroughly (30 s) and centrifuged (4,000 rpm; 15 min; 4 °C). Two types of quality control (QC) samples were formed from pooled aliquots (10 µL) of resuspended samples from either all biological samples or from sexually derived samples only.

Biological, QC and blank samples were loaded (8 µL (non-polar); 7 µL (polar)) into a cooled 384-well plate in quadruplicate using an automated multi-dispensing pipette. The order of samples was randomized across two plates with the inclusion of QC samples evenly spaced every 9 biological samples. Four technical replicates

per biological sample were acquired consecutively. Plates were sealed immediately with 20 µm foil using a Thermo plate sealer (Agbene, UK) to minimise sample evaporation.

2.14.1.3. *Acquisition of mass spectra*

Resuspended sample extracts were analyzed without prior separation using a nESI-DIMS platform comprised of a Triversa NanoMate (Advion, UK) coupled to an LTQ-Orbitrap Elite mass spectrometer (Thermo Fisher Scientific). Analysis was performed on polar extracts in positive and negative ion modes, while negative ion mode only was performed on non-polar extracts. Data were acquired as transients across variable *m/z* ranges (Table 2-4) using SIM-stitch methodology reported previously (Southam *et al.*, 2007, 2017).

Table 2-4 Electrospray ionisation settings for analysis of *Daphnia magna* embryos with nESI-DIMS

	Setting		
Parameter	**Polar extract**		**Non-polar extract**
Polarity	Negative	Positive	Negative
Scan mode	SIM	SIM	SIM
AGC target resolution	5.00E+05	5.00E+05	5.00E+05
Resolution	240000	240000	240000
Window width (Da)	75	75	75
Range (m/z)	50-620	50-620	190-2000
No. of windows	10	10	18
Acquisition delay (s)	30	30	30
Number of microscans	10	10	10
Injection time (s)	1	1	1
Maximum inject time (s)	10	10	10

2.14.1.4. *Processing of mass spectra*

Data were processed as described fully elsewhere (Southam *et al.*, 2017). Data were acquired as transients and summed, processed and ‘stitched’ together using a

custom written SIM-stitching method (Southam *et al.*, 2007) to yield a single intensity matrix per technical replicate, peaks with signal-to-noise (SNR) ratio of > 3 were excluded. A further replicate filter was applied to retain peaks present in a minimum of 2-out-of-3 technical replicates (2.0 ppm tolerance), aligning the three technical replicates and thus generating a single matrix of samples, peak *m/z* and intensity values. Peaks were removed from this matrix that had an intensity < 10-fold higher than extract blank samples, and occurred in < 80% of all biological samples. Biological samples with greater than 20% missing values after the above filtering stages were also removed from matrices. Data were then normalised using probabilistic quotient normalisation (PQN; Dieterle *et al.*, 2006). Batch correction was applied utilising a LOESS model for each peak in QC samples to normalise sample peak intensity across the two analytical batches (i.e. plates; Dunn *et al.*, 2011) and peaks with a relative standard deviation within QC samples of > 30% were also removed from further analysis. Finally, any missing values in the matrix were assigned an estimated intensity utilising k-nearest neighbour imputation (k = 5). For multivariate analysis a generalised-log (g-log) transformation was applied (λ optimised using QC samples).

2.14.1.5. *Metabolite identification*

For metabolite annotation, data matrices (pre log-normalised) were analysed using a beta version (v0.1.3) of Birmingham Metabolite Annotation for Mass Spectrometry (BEAMS) METABOlite annoTATOR (Metabotator) implemented in python using miniconda which operates in a similar manner to the previously described MI-Pack by the same authors (Weber and Viant, 2010). BEAMS generates putative empirical

formulae and metabolite names for *m/z* values detected by matching against known compounds and common ions (e.g. ($[M-H]^-$, ($[M+H]^+$) in metabolite databases (Human Metabolite DataBase (HMDB); Lipid Metabolites and Pathways Strategy (LipidMAPS) database), an error window of 2.5 ppm was used.

2.15. Statistical analysis of transcriptomic and metabolomic data

2.15.1. Univariate

Significant fold changes in mean intensity of metabolite peaks between respective time points were assessed using Student's t-test with Benjamini-Hochberg *p*-value, false discovery rate (FDR) correction (Benjamini and Hochberg, 1995). Any monotonic relationships between peak intensity and/or gene expression were determined by calculation of Spearman's non-parametric correlation coefficient (ρ).

2.15.2. Multivariate

ANOVA-simulaneous component analysis (ASCA) (Smilde *et al.*, 2005) was performed in MetaboAnalyst (v4.0)(Chong *et al.*, 2018), significant peaks were selected based on a leverage threshold of 0.9 and SPE alpha level of 0.05, and permutation testing (50-300 permutations) was performed to validate model significance. Principal component analysis (PCA) was performed on log-transformed gene counts and peak intensity matrices, using the *prcomp* function from *stats* package. Significant differences in the PC scores of biological sample classes with time post-ovulation and embryo type (i.e. subitaneous or diapausing) were assessed using analysis of covariance (ANCOVA; type III sum of squares). Sparse principal component analysis (sPCA) was performed on the same matrices using *spca* function in *sparsepca* package (v0.1.2) (Erichson *et al.*, 2018). The sparsity

controlling parameter (alpha) used for analysis of polar data was $1x10^{-3}$ and $8x10^{-3}$ for non-polar and gene expression data.

3. Optimisation of protocols for the use of *Daphnia magna* as a model system for embryonic diapause

3.1. Introduction

Daphnia magna represents numerous possibilities as a model species for the investigation of diapause, notably the potential to explore the mechanisms of extended lifespans given its extraordinarily long diapause intervals. With any model system, a well-characterised biology is paramount in the interpretation of omics datasets (Feder and Mitchell-Olds, 2003). Furthermore, a robust experimental design where technical and systematic variation is minimised is also critical, particularly as the transcriptome and metabolome are highly influenced by subtle differences in environmental conditions (Churchill, 2002; Macel, van dam and Keurentjes, 2010; Southam *et al.*, 2017). Thus far, investigations of diapause in *D. magna* have predominately had a strong ecological focus based almost exclusively of studies *in situ* or from field-collected ephippia (Boersma, Boriss and Mitchell, 2000; Arbačiauskas and Lampert, 2003; Vandekerkhove *et al.*, 2005; Conde-Porcuna, Ramos-Rodríguez and Pérez-Martínez, 2014; Navis *et al.*, 2015; Kaupinis *et al.*, 2016). Very few studies have utilised laboratory reared ephippia but have been instrumental in partitioning out genetic versus environmental variation in ephippia pigmentation (Gerrish and Cáceres, 2003), in addition to the transgenerational transmission of photoperiodic and food quantity information, which impacts ephippia production (Alekseev and Lampert, 2001). Whilst a protocol for the successful hatching of diapausing *D. magna* has been provided (Cuenca Cambronero and

Orsini, 2018), no similar procedure for the stable production of laboratory-induced diapause is available. In this chapter, I address this limitation.

D. magna isolate “L7” when maintained in stock culture under ‘standard’ laboratory conditions shows a high propensity to produce male and ephippial offspring and therefore was chosen as an appropriate isolate for the study of *D. magna* diapause. Isolate “B2” is routinely used within the University of Birmingham *Daphnia* research group for ecotoxicological studies as it consistently reproduces asexually under standard conditions without the production of males or ephippia. Herein the reproductive behaviour of L7 is characterised and contrasted to B2 under a set of experimental conditions to optimise the reproducible induction of diapause in *D. magna.* Including confirmation that L7 fulfils specific criteria involving the absence of obligate parthenogenesis (Innes, Fox and Winsor, 2000), the production of fertile male offspring, and viable ephippia in sufficient quantity.

Induction of sexual reproduction in *Daphnia sp.* has been reported in response to a variety of biotic and abiotic stressors. Whilst juvenile hormones and endocrine disrupting analogues have been established to reliably induce male production in *Daphnia sp.* (Olmstead and Leblanc, 2002; Tatarazako *et al.*, 2003), no study has yet confirmed the fertility of these induced males. Furthermore, these compounds have been shown to illicit a concentration-dependent negative effect on neonate output (Oda *et al.*, 2006). Thus, chemical induction of males was dismissed as inappropriate for use in this study as adequate information is lacking as to how much this mimics true environmental sex determination in *D. magna*. Instead, I focus on biotic stressors known to induce sexual reproduction that were easily reproducible in the laboratory. Food concentration and quality has been reported to induce sexual

reproduction in *Daphnia sp.* (Deng, 1996; Koch, von Elert and Straile, 2009). Concentration of algal food was selected as an appropriate stress; for reproducibility and, in contrast to photoperiod and temperature, test cultures can be maintained in the same environment as maintenance and control cultures removing potential systematic environmental variation.

Next, I assess the total lifespan of the chosen clone, L7 under the given laboratory conditions and I investigate the hatching of diapausing embryos produced during my experiments over a comparatively short time scale of 2 years. This both addresses whether an extension of lifespan is observed and furthermore confirms the viability of the diapausing offspring I produced because successful fertilisation, as demonstrated by the presence of embryos in the ephippia (Zaffagnini, 1987), does not necessarily guarantee that embryos will go on to successfully break diapause and develop into healthy, mature adults.

I also establish a staging system for the morphological landmarks of ephippia production in *D. magna* in relation to ecdysis and offspring release in line with the similar information available for *D. pulex* (Hiruta and Tochinai, 2014), so to both widen the current description of sexual reproduction in *D. magna,* in addition to defining the key events of interest to the direct molecular investigation of diapause in this species.

Finally, I optimise the sample biomass required for the successful application of nESI-DIMS metabolomics for use of low biomass *D. magna* embryo samples.

3.2. Selection of isolate and culturing parameters for the laboratory induction of diapause in *Daphnia magna*

To determine appropriate culturing conditions for the induction of sexual reproduction and consequently diapause in *D. magna*, in addition to confirming the selection of L7 as a suitable isolate for use in further studies, two isolates of *D. magna* (L7 and B2) were reared under "full" (0.08 mgC Daphniid^{-1}) and "reduced" (0.03 mgC Daphniid^{-1}) *C. vulgaris* food ration at a density of 20 *Daphnia* L^{-1} for 21 d (Section 2.5). Their reproductive output during this time, including number of neonates and ephippia are shown in Figure 3-1. Under a full food ration, both isolates maintained a high level of neonate output for the duration of the experiment (Figure 3-1; A). B2 cultures (n = 3) produced an average of 598.7 (SD = 9.8) neonates, whereas the L7 cultures (n = 4) produced a significantly higher average of 755.3 (SD = 77.6; F (1,13) = 7.38, p = 0.02) neonates over the course of 21 d. When provided with a reduced diet ration, both L7 and B2 cultures produced fewer neonates, with L7 cultures producing an average of 76.8 (SD = 35.9) and B2 an average of 94.7 (SD = 10.0) neonates over 21 d. At a full food ration, no ephippia were produced by the B2 isolate, whereas the L7 isolate produced an average of 7.5 (SD = 5.1) ephippia over the course of 21 d (Figure 3-1; B). Ephippia were absent in one L7 replicate. A diet of reduced algae ration elicited an increased incidence of ephippia in B2 cultures, from absent to 6 ephippia (SD = 2.0), and from 7.5 (SD = 5.1) to 25.5 (SD = 6.8) in L7 cultures. Although the two isolates responded in a significantly different way, with the induction of ephippia in B2 being much greater at approximately 6-fold than L7 at approximately 3-fold (F (1,13) = 5.51, p = 0.041). However, L7 consistently produced a significantly larger number of ephippia under both full and reduced food ration

(F (1,13) = 27.90, p = 0.0004). The age of first release of offspring, both ephippia and free-swimming neonates, is shown in Figure 3-1 (D).

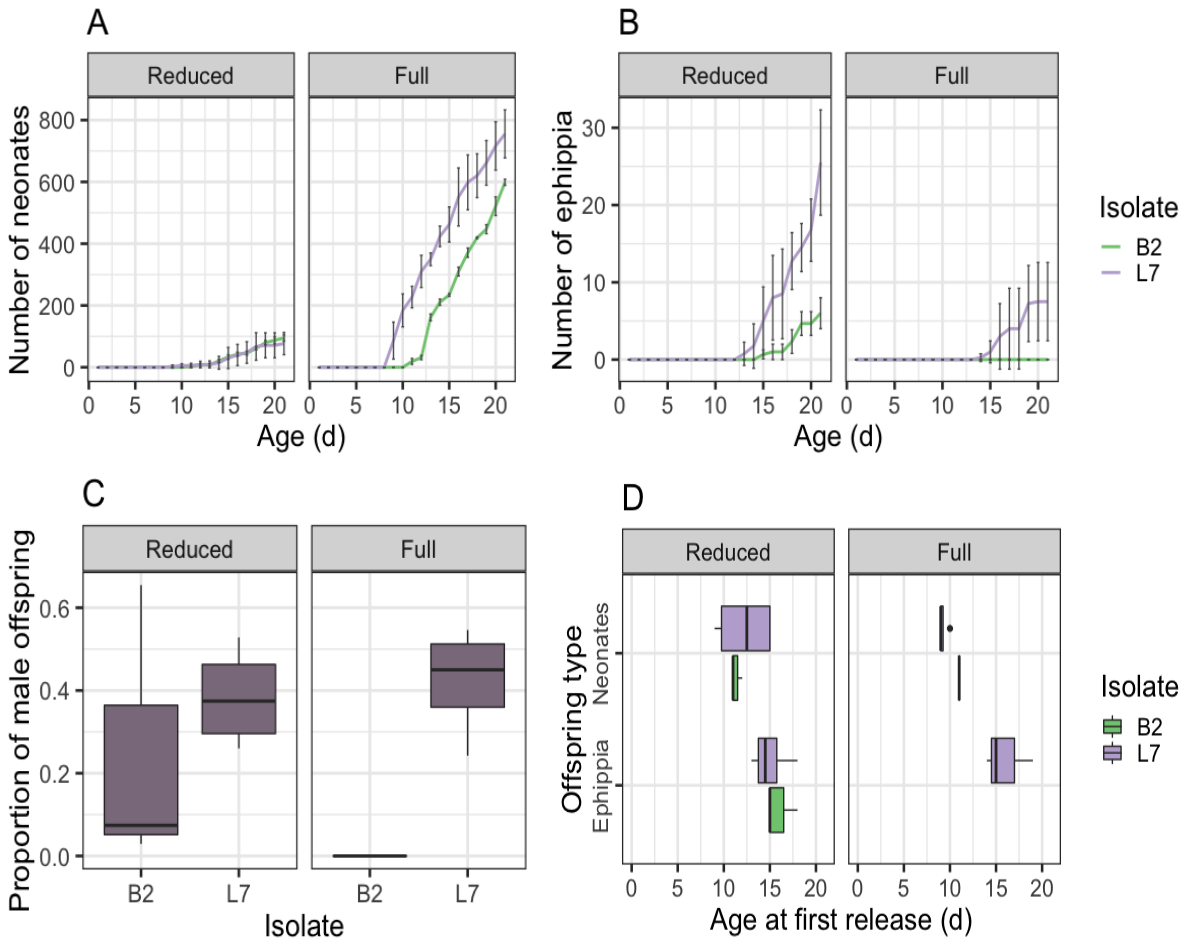


Figure 3-1 The reproductive output of L7 and B2 isolates of *Daphnia magna* over 21 d under "full" (0.08 mgC Daphniid^{-1}) and "reduced" (0.03 mgC Daphniid^{-1}) *Chlorella vulgaris* food ration at a density of 20 *Daphnia L*$^{-1}$ (L7, n = 4; B2, n = 3). Mean cumulative number of neonates (A) and ephippia (B) produced by total cultures (20 *Daphnia*) over 21 d. Error bars denote standard deviation. Proportion of total offspring produced by cultures over 21 d that were male (C). (D) Maternal age (d) at first release of respective offspring type.

Under reduced food conditions, age at first reproduction was slightly later and more variable, particularly for L7 cultures, which presented with a first brood of neonates at around 9 d of age under full food ration and significantly later at approximately 12 d under reduced food conditions. Although there were slight differences between isolates in the age of first reproduction under both food conditions, these were not statistically significant. Delayed onset of sexual maturity in response to food limitation

has previously been reported and is thus in line with expectations (Giebelhausen and Lampert, 2001). *Daphnia* are capable of producing ephippia at the onset of reproductive maturity without prior parthenogenetic broods (Booksmythe *et al.*, 2018). However, because here I observed whole cultures, rather than individual animals, it is difficult to conclude whether individuals completed a previous round of asexual reproduction prior to beginning ephippia production.

The proportion of male offspring in each culture varied with food conditions and between isolates (Figure 3-1; C). Under full food conditions the L7 cultures produced offspring with a mean proportion of 0.42 (SD = 0.13) males. In contrast, no male offspring were observed in B2 cultures under a full diet ration. Reduced food conditions had a significant effect on the proportion of male offspring produced by each isolate but in a clone specific manner (z = -3.93, p = 0.0001). Male production by the B2 isolate increased from absent in full conditions to 0.25 (SD = 0.35) under reduced food ration. In contrast, L7 cultures produced a lower proportion of males than full food ration of 0.38 (SD = 0.12). Regardless, isolate L7 produced a significantly higher proportion of male offspring than B2 in both conditions (z = 16.87, $p = 2x10^{-16}$).

These results confirm that food concentration, as previously reported, had a significant effect on *Daphnia* reproduction by affecting the number of neonates and ephippia produced, although the effect on production of males appears to be complex, potentially impaired by the use of an isolate already inclined to produce males. As previously stated, the elucidation of the mechanism of sexual reproduction in *Daphnia* is not the topic of this thesis, nor are these experiments sufficient to address the remaining questions. For example, whilst reduced food has previously

been reported to increase the production of male offspring and ephippia, there is controversy over whether the two events are independent, perhaps in response to the same stimulus but at different concentrations (Booksmythe *et al.*, 2018). Therefore, a reduced food ration of 0.03 mgC Daphniid^{-1} was used in all further experiments to produce ephippia, as more ephippia were produced under the reduced food condition, facilitating a higher throughput for further experiments. The observation that asexual and sexual reproduction still occurred suggests that this chosen concentration was not limiting in terms of growth and reproduction (Gliwicz and Lampert, 1990; Boersma and Kreutzer, 2002).

Based on the empirical evidence presented here by comparing the reproductive behaviour of the B2 and L7 isolates, L7 was chosen for use in all further experiments presented in this thesis. First, the L7 isolate consistently produced a high proportion of ephippia under both reduced and full food conditions. As these diapausing offspring will form the bulk of the samples used within this thesis, a consistent, adequate supply was required. Second, in order to ensure the viability of ephippia that require fertilisation, a large number of male *D. magna* were also needed. And irrespective of food condition, L7 produced a high proportion of male offspring. However, my results indicate notable inter-clonal differences in neonate, ephippia and male production to food limitation and are in line with previous reports (Glazier, 1992; Giebelhausen and Lampert, 2001; Lampert, Lampert and Larsson, 2012). These data outline an important limitation of the results obtained herein – notably that extrapolation to the *D. magna* species as a whole will require further validation across clonal isolates.

To ensure ephippia viability, the same conditions of full and reduced food ration were repeated with an additional 5 female or 5 male individuals added to the initial 20 *Daphnia L*$^{-1}$ culture. Over the course of 21 d, ephippia were collected from each culture and manually decapsulated to ascertain viability. As mature males are required to fertilise haploid diapause-destined eggs, ephippia viability can be defined by the presence of at least one egg within the final ephippia (Conde-Porcuna, Ramos-Rodríguez and Pérez-Martínez, 2014). Additionally, the inclusion of 25 female *Daphnia L*$^{-1}$ cultures was performed to control for the additional number of animals per beaker, as population density is known to influence incidence of sexual reproduction (Burns, 2000).

The number of ephippia corrected per adult female for cultures of 20 *Daphnia L*$^{-1}$ and 25 *Daphnia L*$^{-1}$ (25 female or 20:5 female:male) are shown in Figure 3-2; A. As in my previous experiment, a reduced diet ration significantly increased the number of ephippia produced by all culture types (F (1,18) = 30.04, p = 0.0001). However, the presence of adult males in the culture had no significant effect on the number of ephippia produced. The mean ephippia per adult under reduced food was greater than 1 for all culture types although, as output was considered at culture rather than individual level, it was not possible to ascertain whether every female completed one round of sexual reproduction or a sub-set participated in multiple rounds over the 21 d. At full food ration, ephippia production was much more variable, although in general, fewer ephippia were produced.

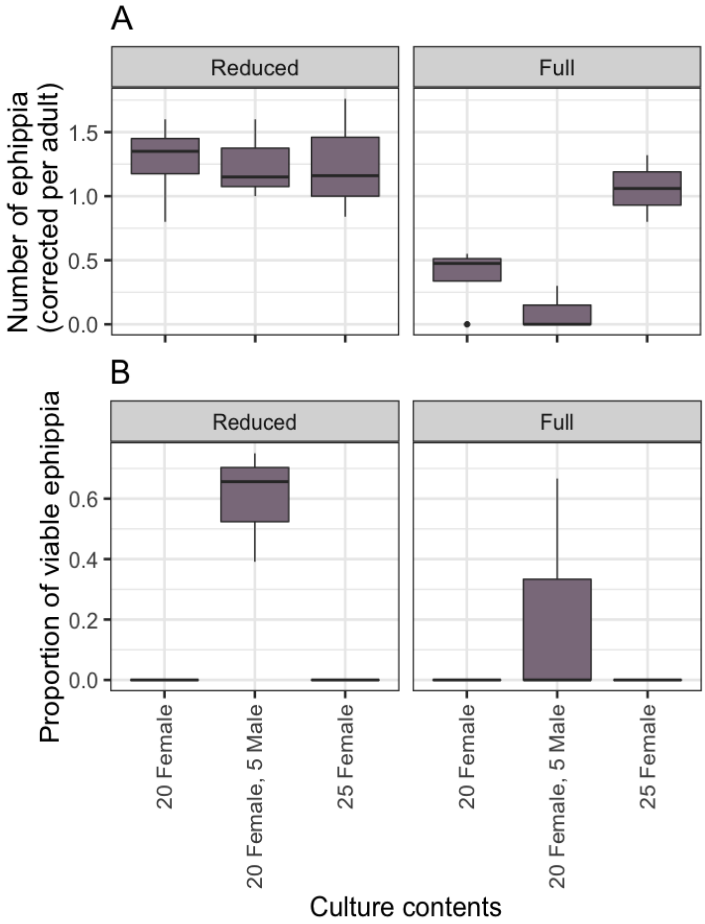


Figure 3-2 The effect of adult male presence on viable ephippia production over 21 d under "full" (0.08 mgCDaphniid^{-1}) and "reduced" (0.03 mgCDaphniid^{-1}) *Chlorella vulgaris* food ration at a density of 20/25 *Daphnia L*$^{-1}$. A) Total number of ephippia produced is corrected for the number of adult females present in each culture, B) Proportion of viable ephippia produced (ephippia that contained at least one egg upon release were considered viable).

The proportion of ephippia produced that were viable is reported in Figure 3-2(B). Viable ephippia were only observed in cultures that contained adult males. The mean proportion of viable ephippia produced by the 20:5 female:male culture at a reduced food ration was 0.60 (SD = 0.19). At a full food ration, only one of the cultures produced ephippia, of which 67% were viable. This observation is of critical importance as it suggests that the L7 isolate is not an obligate asexual; in other words, they do not produce viable diploid diapausing embryos in the absence of males as has been previously reported in *D. pulex* (Innes, Fox and Winsor, 2000).

In the presence of males, only ~60% of eggs produced were viable. In order to increase ephippia yield, culture density was increased to 62 *Daphnia* L^{-1} comprising 50 female and 12 male individuals. The reproductive output of these cultures is reported in Figure 3-3. Significantly more ephippia were produced under reduced food conditions; t(6.68) = 6.19 , p = 0.00054. At a full food ration, mean total culture ephippia output was 12.0 (SD = 8.2), whereas a greater than 6-fold induction was observed when food concentration was reduced with a mean total ephippia production of 79.5 (SD = 13.43). By comparison with my previous experiment involving males, the corrected ephippia output per individual was only slightly higher at the increased culture density. At full ration, 0.24 per daphniid (SD = 0.16) were produced at 62 *Daphnia* L^{-1} and 0.1 at 25 *Daphnia* L^{-1} (SD = 0.17), similarly at reduced food 62 *Daphnia* L^{-1} cultures produced 1.59 ephippia per adult (SD = 0.27) and 25 *Daphnia* L^{-1} 1.25 (SD = 0.31).

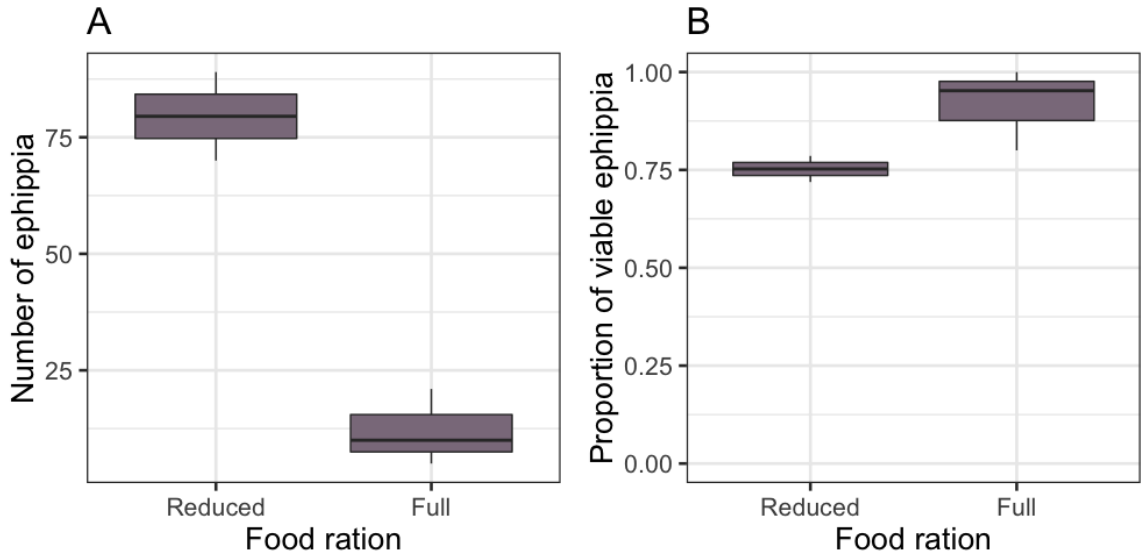


Figure 3-3 Viable ephippia production over 21 d under "full" (0.08 mgCDaphniid^{-1}; n = 4) and "reduced" (0.03 mgCDaphniid^{-1};n = 3) *Chlorella vulgaris* food ration at a density of 62 *Daphnia* L^{-1}. A) Total number of ephippia produced, B) Proportion of ephippia produced that contained at least one egg upon release.

No significant differences in the proportion of viable ephippia produced were observed between the two food rations at 0.75 (SD = 0.05) reduced, 0.92 (SD = 0.10) full. However, this proportion of viable ephippia is greater in comparison with the 20female:5male culture (0.60; SD = 0.19 reduced and 0.22; SD=0.38 full). Although the ratio of male:female *Daphnia* was equivalent in both experiments at 1:4, it is possible that, with the increased culture density, there was a greater chance of physical contact and thus mating.

There is growing evidence that *Daphnia sp.* respond to increased density of cogners and conspecifics mediated by a chemical factor of unknown origin, leading to changes in various life history traits including reproduction (Burns, 2000; Lürling *et al.*, 2003). For example, crowding increases ephippia production in *D. pulex* (Fitzsimmons and Innes, 2006; Schwarzenberger, Christjani and Wacker, 2014). Furthermore, media derived from crowded cultures reduced neonate brood size in several *Daphnia* species, but not *D. magna*, however densities used were greater, at 85–200 *Daphnia* L^{-1} than those I investigated (Burns, 2000). My data do not sufficiently address the issue of the effect of crowding on the induction of sexual reproduction in *D. magna.* Although my results do suggest an increased incidence of ephippia production at higher densities, further experimental investigation is needed. My experiment was designed to achieve an improved experimental yield of viable ephippia, which was achieved at the investigated density of 62 *Daphnia* L^{-1} with a male:female ratio of 1:4 and a reduced algal diet ration of 0.03 mgC *Daphnia*$^{-1}$. Therefore, these culturing parameters were applied throughout the remainder of this thesis, for the production of all experimental ephippia.

3.3. Assessment of the viability and longevity of laboratory-derived diapause offspring in *Daphnia magna*

There is surprisingly limited information regarding the life expectancy of *D. magna,* and typically such studies are conducted with a toxicological focus. Furthermore, authors used slightly different culturing conditions (Pietrzak, Bednarska and Grzesiuk, 2010; Schwarzenberger, Christjani and Wacker, 2014; Li *et al.*, 2017). Different culture media, photoperiods and nutritional content of algae are all known to affect *Daphnia* life history, therefore comparisons should be made carefully (Baer and Goulden, 1998; Zhang and Baer, 2000; Muñoz-Mejía and Martínez-Jerónimo, 2007; Alekseev and Lajus, 2009). I investigated the lifespan of both male and female *D. magna* under both full and reduced food conditions. Survival curves are reported in Figure 3-4; A. There is a significant difference in the lifespan of male and female *Daphnia* under both food conditions (Table 3-1), with females living longer than males. Under reduced food conditions, the mean lifespan of female *Daphnia* was 67 d (SD = 2.3), whilst males lived for 55 d (SD = 16.2), although there was considerably more variation in the male cultures with a 32 d range. By contrast, under full food conditions, females lived significantly longer with a mean life expectancy of 102 d (SD = 1.6) and 64 d (SD = 4.5) for males.

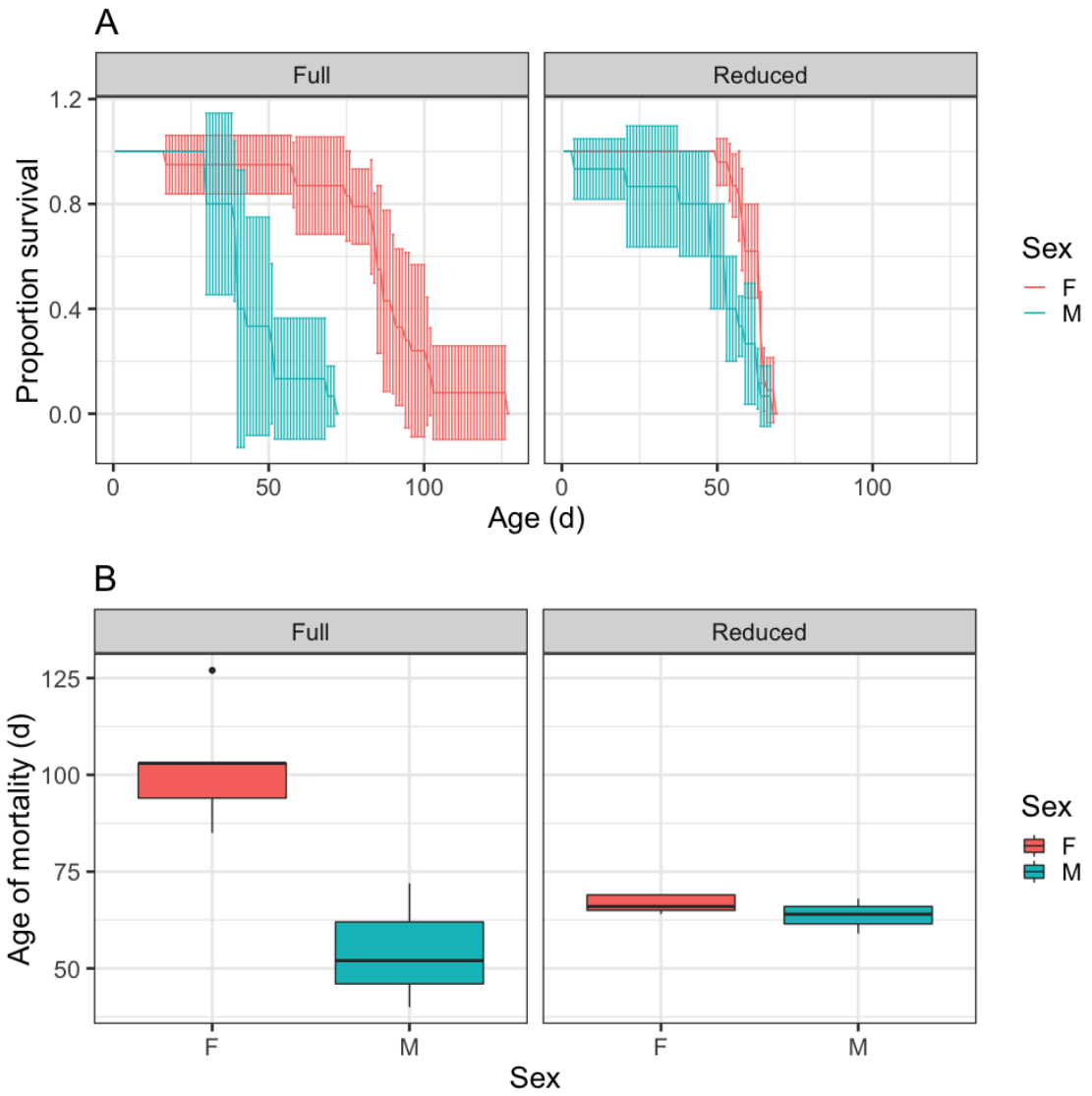


Figure 3-4 Longevity of male (M; n = 3) and female (F; n = 5) L7 *Daphnia magna* under "full" (0.08 mgC Daphniid^{-1}) and "reduced" (0.03 mgC Daphniid^{-1}) *Chlorella vulgaris* food ration. A) Mean proportion of surviving individuals in cultures of 5 *D. magna* (error bars denote standard deviation). B) Age at total culture mortality after all 5 culture individuals had expired.

Li *et al.* (2017) reported a life-span for *D. magna* of 94 days under conditions permissive for parthenogenetic reproduction, which is comparable with the 102 d that I observed. Schwarzenberger *et al.,* (2014) also examined the longevity of both *D. magna* sexes and reported that females live significantly longer (ca. 20 d) than males, which is consistent with my observations, although they report shorter life expectancies of 60–70 d and 40–50 d respectively. By contrast, no significant sex differences in longevity were reported in multiple clones of *D. magna,* but the authors

did report longer life expectancies of 80–120 d, which are more in line with observations for female *Daphnia* at non-limiting food concentrations (Pietrzak, Bednarska and Grzesiuk, 2010).

Table 3-1 Analysis of variance table investigating the effect of gender and *Chlorella vulgaris* food ration on age (d) of *Daphnia magna* at 100% culture mortality

Source	*Df*	MS	*F*	*p*	Effect size (partial η2)
Sex*	1	4272.1	32.78	9.52E-05	0.6062
Food ration*	1	3204.1	24.59	3.32E-04	0.4801
Sex *x* Food*	1	1881.6	14.44	2.53E-03	0.5461
Error	12	1563.7	–	–	–

My results also indicate that food concentration had a significant effect on lifespan, but that the life expectancy of the two sexes responds differently (Table 3-1). Female life expectancy was longer under full food conditions, however male daphniid life expectancy was longer in reduced food conditions (Figure 3-4; B). Extension of life-span under caloric restriction is documented in many but not all species, and arises due to a complex interplay between natural selection and the allocation of resources to survivorship or reproduction (Shanley and Kirkwood, 2006). Caloric restriction has been observed to significantly extend the life expectancy of *D. pulex* whilst having no significant effect in *D. pulicaria,* although both these species are shorted lived –ca. 26 d and 48 d respectively– than the larger *D. magna* (Latta IV, Frederick and Pfrender, 2011). The authors suggest that in shorter-lived *Daphnia* species that inhabit temporary ponds, allocation to reproduction rather than somatic maintenance is perhaps likely to be favoured (Latta IV, Frederick and Pfrender, 2011). In the current experiment the effect of only two food concentrations were observed, so it is unclear whether the reduced food ration I utilised that negatively affected longevity

was sufficiently low to represent a stress severe enough to induce mortality over long term exposure. As reproductive output and growth rate was not assessed, it is not possible to comment on allocation of resources to reproduction versus somatic maintenance. The conflicting and sparse evidence surrounding *Daphnia* longevity suggests that more research is needed to unravel sexual, clonal and conditional (e.g. caloric restriction) effects, which cannot be addressed by the current experiment.

In order to confirm the viability of my experimentally derived ephippia, variously aged ephippia (14 d, 30 d, 60 d, and 728 d) that had been collected from cultures and stored in darkness at 4°C (Section 2.7) were manually decapsulated and exposed to full spectrum light to induce diapause termination and hatching. The proportion of hatching from ephippia and survival of hatchlings to reproductive maturity was investigated. Successful hatching was observed across ephippia of all ages, confirming that the ephippia produced from the experimental conditions are viable and suitable for use in further experiments (Figure 3-5). Successful hatching was observed for 146 ephippia out of a total 392. Whilst the majority of ephippia contained 2 embryos, there were only 6 instances when both embryos hatched, and of these only one, if any, of the offspring reached maturity. Therefore, for clarity, hatching success as discussed herein relates to the successful hatching of one embryo within the ephippium.

Hatchlings were observed in all groups of ephippia following decapsulation and transfer to diapause-breaking conditions. However, there was a slight delay of several days, which likely accounts for time taken for the perception of stimuli and post-diapause development to occur. There are limited observations of development of ex-ephippio offspring. It has been reported that development proceeds over a

similar timescale of asexual development (ca. 48 h; Navis *et al.*, 2015), and therefore, immediate hatching was not anticipated. The hatching profiles of all ages of ephippia demonstrated similar intermittent bouts of mass hatching and sporadic individual activation (Figure 3-5; A). Whilst embryos were observed for up to 60 d after decapsulation, no hatching was observed after 30 d post-decapsulation, leading me to believe that by 30 d, maximal hatching had been achieved. The largest proportion of hatching occurred in youngest (14 d) ephippia at 51% whilst ephippia that had been stored for 728 d (2 y) displayed the lowest proportion of successful hatching at 22%. The hatching of 30 d and 60 d ephippia was intermediate at 30% and 44% respectively. Furthermore, there was no significant association with age of the ephippia and the proportion of hatching (r^2 = 0.576), although there is a weak trend towards older ephippia showing lower proportion of hatching (Figure 3-5; B).

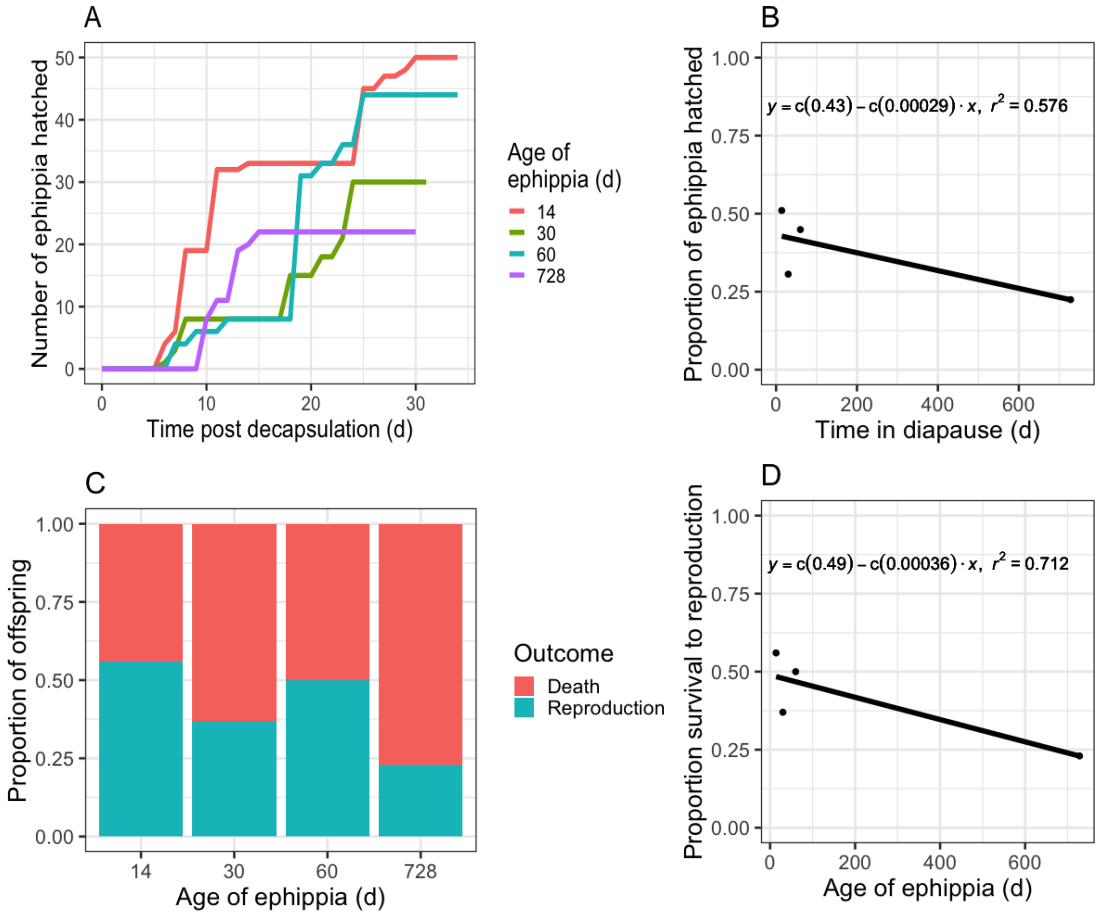


Figure 3-5 Hatching rates and survival to reproductive maturity of laboratory-produced diapausing embryos of *Daphnia magna* from ephippia of ages 14 d, 30 d, 60 d, and 728 d. A) Proportion of neonates hatching over 30 d post decapsulation, B) relationship between ephippial age and total proportion of hatching after 30 d, C) total proportion of hatched offspring to reach sexual maturity, D) relationship between proportion of hatching to reach sexual maturity and age of ephippia.

Previous reports on the laboratory hatching of *Daphnia sp.* ephippia have shown that temperature, light and conditions of storage all play a role in determining the number of successful hatchlings (Davison, 1969; Vandekerkhove *et al.*, 2005). Therefore, my treatment of prior storage in darkness at 4°C for a minimum of 14 d, followed by manual decapsulation, and exposure to 20°C, and broad-spectrum, continuous light was assumed to have elicited the maximum amount of diapause termination. Furthermore, the manual decapsulation of embryos from the ephippial casing applied here is shown to improve hatching success (Paes, Rietzler and Maia-Barbosa, 2016).

Incomplete hatching of *Daphnia sp.* ephippia has previously been reported, and typical laboratory hatching conducted under conditions similar to mine range from 6-60%. Therefore, my results were consistent with expectations (Cáceres and Tessier, 2003; Drillet, Hansen and Kiørboe, 2011; Cuenca Cambronero and Orsini, 2018). Of the ephippia that hatched, offspring were followed through to first reproduction or death to further confirm their viability (Figure 3-5; D). Consistent with previous observations, all ex-ephippio offspring were female (Hobaek and Larsson, 1990). Furthermore, all offspring that reached reproductive maturity presented with an asexual brood of neonates, although the number and sex of the brood was not determined. A large proportion of offspring that successfully hatched from ephippia died before reaching reproductive maturity–typically within the first few days after hatching. The percentages of hatched offspring that survived to first reproduction where highest for 14 d ephippia at 56%, 37% for 30 d ephippia, 50% for 60 d ephippia and just 23% of offspring that hatched from 728 d ephippia survived to produce their first brood (Figure 3-5; C). There was no significant association with age of ephippia and the proportion to reach successful reproduction (r^2 = 0.712; Figure 3-5; D).

Offspring of *D. magna* originating from sexual reproduction have previously been reported to exhibit no differences in fitness compared to their asexual counterparts (Arbačiauskas, 2004). Although robust data on the proportion of ex-ephippio *D. magna* that reach sexual maturity is scarce, clonal differences in the proportion of survival to reproductive maturity have been reported, with some isolates achieving just 53% survival (Navis *et al.*, 2013). Therefore, my results are consistent with limited previous reports, yet the reasons behind such low survivorship are unclear. I

suggest several plausible hypotheses for why this may arise. First, eggs were hatched in small volume 96-well plates and may have been without food for up 48 h between plate observations and transfer to culturing conditions. Therefore, conditions may have been sub-optimal for survival. Indeed, post-ephippia offspring have been shown to exhibit faster mortality under starvation than asexual offspring (Arbačiauskas and Lampert, 2003). Second, in my experiment, males and females of the same clone were mated (i.e. selfed) and the resulting inbreeding depression has previously been shown to reduce hatching and survivorship in *D. magna* (De Meester, 1993). Third, *Daphnia* are suggested to display staggered hatching of ephippia in order to maximise survivorship. Thus, it may be that the timescale (60 d) over which I observed embryos was not sufficiently long (Cáceres and Tessier, 2003). Last, viability was confirmed only by means of hatching, as there are currently no established protocols for determining whether the diapausing embryos of *Daphnia* are viable beyond the observation of successful completion of development and hatching. It may be that a portion of the embryos produced had expired at some earlier point due to unknown factors. Furthermore, the conditions experienced and length of diapause can also have a profound effect on later life stages of the organism. With increasing time spent in diapause, an increase in mortality, a decrease in fat reserves and a subsequent reduction in the size and fitness of emerging adults has been observed in the parasitioid *Asobara tabida* (Ellers and Van Alphen, 2002). These observations may partially explain the slight trend towards lower hatching and survival in older ephippia.

In conclusion, the successful hatching of offspring originating from ephippia 2 years in age and the maximum life expectancy observed of 127 days, confirm that

laboratory-reared diapausing offspring of *D. magna* may display greatly extended longevity. My data also confirm the success of the earlier experiments to determine suitable conditions for the induction of diapause in *D. magna* in that, the offspring originating from the chosen conditions are viable.

3.4. Temporal definition of the major morphological landmarks of embryonic diapause in *Daphnia magna*

To direct later experiments, a robust characterisation of the landmarks of sexual embryo production was required to anchor sampling time points to and define the timescale of interest for the identification of diapause initiation. Whilst, a description of the events of ephippia production is available for *D. pulex* (Hiruta and Tochinai, 2014), there is no equivalent for *D. magna*. *D. magna* were raised under the conditions outlined in Section 3.2. Prior to the onset of reproduction (ca. 7 d) *D. magna* were transferred to 24-well plates in single mating pairs, observed and imaged every 12 h.

The morphological maternal landmarks of sexual and asexual embryo production *D. magna* are described in Figure 3-6. Two major events, ovulation (Figure 3-6; white triangles) and the release of offspring from the mother, occur and are inherently linked to ecdysis. Thus, the maternal-based events of *D. magna* reproduction may be divided into two phases; an ovarian development stage encompassing the last time point at which ovaries are discernibly empty to ovulation (Figure 3-6; A and C) and a brood pouch development stage: the period between ovulation and release of offspring or ephippium from the maternal carapace (Figure 3-6; B and D). To visualise embryo morphology masked by the maternal carapace, embryonic

development outside of the maternal brood pouch was also determined and is shown in Figure 3-7.

Ovarian Development

D. magna have two ovaries, located along the length of the female thorax close to the intestine (one on either side). Prior to the onset of reproduction, *Daphnia* ovaries are difficult to visualise and it is unclear which form of reproduction a daphniid will undertake (Figure 3-6; A-i and C-i) until egg ovarian development initiates, at which point notable differences become apparent. Asexual reproduction in *D. magna* results in multiple offspring per round of reproduction, and at the onset of the reproductive cycle, several oocytes are visible along the entire length of the ovary (Figure 3-6; A-i (➔)). By contrast, each ovary produces a single amphigonic oocyte, meaning a maximum of two sexual eggs per round of reproduction. With the onset of reproduction, each oocyte originates from a germarium at the female posterior near the tail (Figure 3-6; C-I; ➔). As development progresses, the oocytes increase in size filling the ovary (Figure 3-6; A-v and C-ix (**---**)). The period of ovarian development is longer for sexual eggs, with 84 h between the first appearance of eggs and ovulation in contrast to 48 h for asexual eggs. A longer developmental period for sexual oocytes was previously reported and may be due to several factors (Hiruta and Tochinai, 2014). This may reflect a period of preparation for diapause, including the sequestration of resources required for the extended period of dormancy (Hahn and Denlinger, 2007). Furthermore, during sexual reproduction, increasing pigmentation of the maternal carapace becomes apparent 24 h prior to ovulation, demarcating the region that will eventually become the ephippium (Figure 3-6; C-vii; (*)). It was

demonstrated that the carapace of sexually reproducing *D. pulex* significantly increases in thickness prior to ovulation, and the ephippial region of the carapace has a unique elemental composition (Kawasaki, Sakata and Namiki, 2004; Hiruta and Tochinai, 2014). Additional time may be required to complete these extensive carapace-remodeling processes during ephippia production.

Brood Pouch Development

During parthenogenesis, several eggs are released with each moult cycle. Over the first 12 h (Figure 3-7; C), they remain largely unchanged and are roughly spherical, although there is a slight furrowing potentially associated with the formation of appendages roughly equivalent to stage 6, as described by Mittmann *et al.*, (2014). By 24 h, the extension of the first antenna are visible (Figure 3-7; C (➔)) and at 36 h post-ovulation, the egg morphology has altered significantly by displaying a visible loss of symmetry and closely resembles stage 9 embryos as described by Mittmann *et al.*, (2014). By 60 h, offspring appendages and body plan become obvious; the development of the eyes is now visible (Figure 3-7; C (>)). Prior to the final moult of the reproductive cycle, the now free-swimming offspring will be released from the maternal brood pouch marking the beginning of the first juvenile instar. Around 72 h after ovulation, with the maturation and release of offspring from the brood pouch now complete, the maternal carapace is once again shed, thus marking the end of the cycle of reproduction.

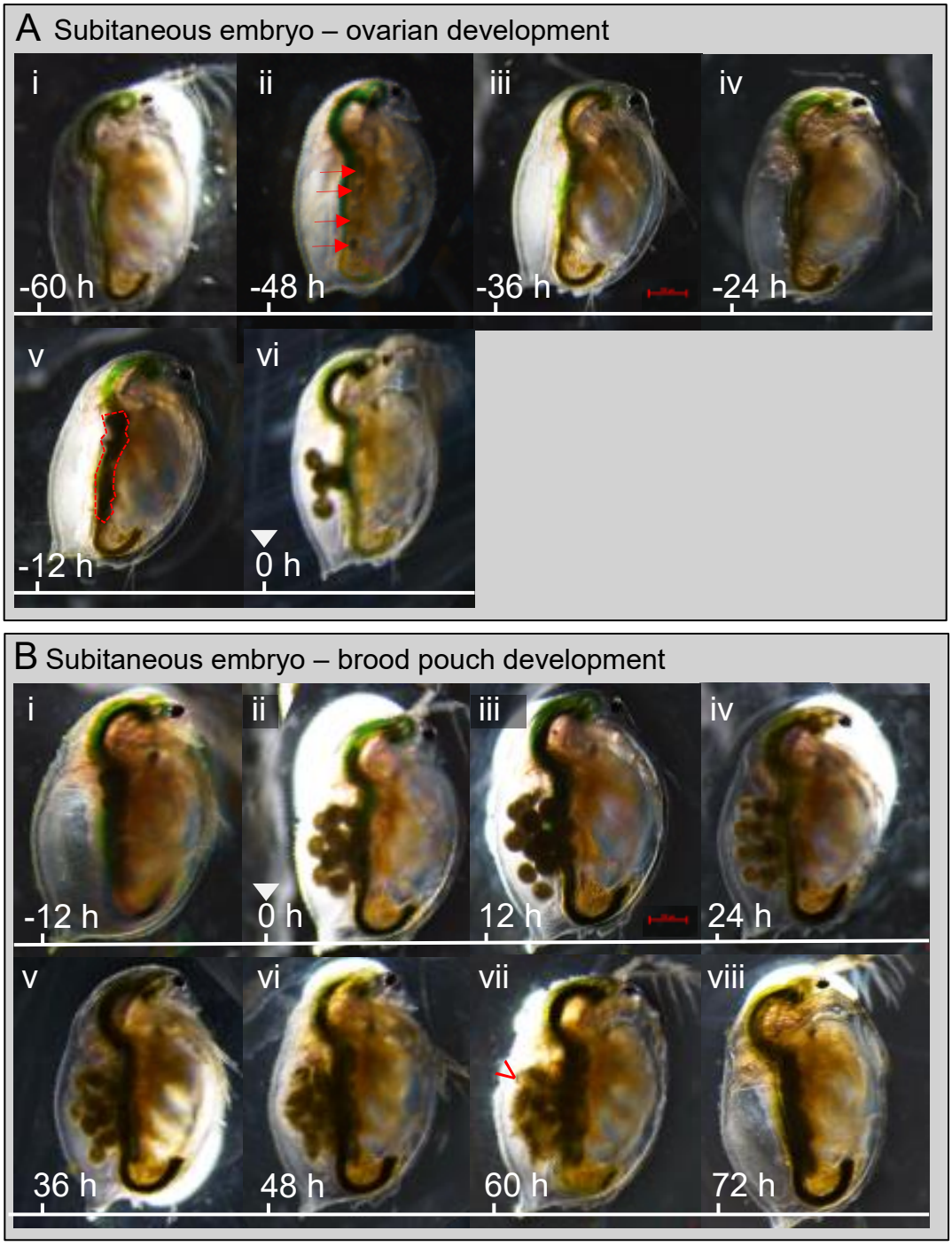
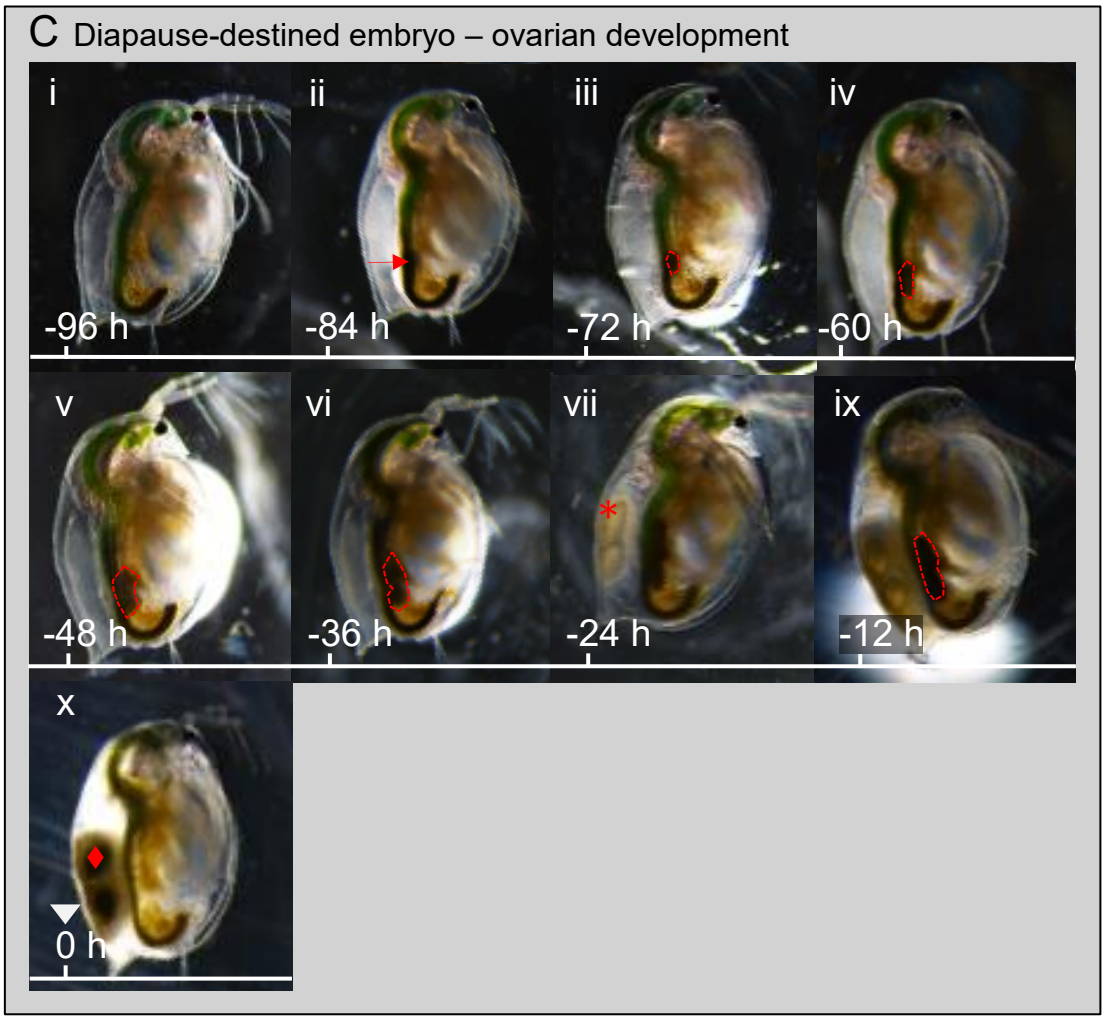


Figure 3-6 The maternal production of *Daphnia magna* asexual (A-B) and sexual embryos (C-D) imaged every 12 h. Scale bar = 500 µm. White bars and annotation indicate time (h) post-ovulation when imaging was performed. White triangles indicate ovulation which was defined as 0 h. Annotations in red indicate morphological points of interest including; the onset of oogenesis (➔), the position of oocytes within the maternal ovary are outlined (---), compound eye formation (>), increasing pigmentation of the maternal carapace(*), release of sexual embryos into maternal brood pouch (♦).


C Diapause-destined embryo – ovarian development
i
ii
iii
iv
-96 h
-84 h
-72 h
-60 h
v
vi
vii
ix
-48 h
-36 h
-24 h
-12 h
x
0 h
D Diapause-destined embryo – brood pouch development
i
ii
iii
iv
-12 h
0 h
12 h
24 h
v
vi
36 h
48 h

During sexual reproduction, a maximum of two eggs are released into the brood pouch (one from each ovary), initially appearing green and ovoid (Figure 3-7; B). Over the next 36 h, the change in egg morphology is negligible. However, the area of the maternal carapace that will eventually form the ephippium becomes increasingly melanised and the eggs move into their final positions in the indented sections (Figure 3-6; D-ii – D-v (♦)). At 36 h after ovulation, the ephippial structure is black in colour and completely opaque (Figure 3-6; D-v); over the next 12 hours, the two halves joined at the carapace spine become increasingly pinched together, closing around the eggs within, before the mother finally moults. The ephippium is then released along with the rest of the carapace.

These observations confirm the establishment of a maternal timeline for ephippia production in *D. magna,* based on major morphological landmarks. They indicate the key events of ovulation and ephippia release, which is consistent with previous reports of ephippium formation elsewhere (Hiruta and Tochinai, 2014). There is significant variation in the timing of these reproductive events across biological replicates and between reproductive strategies. The fine detail of asexual embryogenesis is not reported, yet is found in excellent detail elsewhere (Mittmann *et al.*, 2014). However, the morphological events observed herein where comparable to previously described. Differences in egg morphology between the two offspring types are already apparent at early stages (12 h) post ovulation suggesting divergent developmental processes (Figure 3-7). Amphigonic egg morphology remains unchanged. However subitaneous eggs show a visible developmental progression suggesting that the onset of diapause may occur prior to the release of ephippia (Mittmann *et al.*, 2014).

This timeline represents a framework onto which further studies can be built, including omics sampling, and defines a specific region of interest in determining when sampling to investigate pre-diapause and diapause events (i.e. before and after ephippia release respectively). As limited changes in diapausing embryo morphology were observed, I opted to anchor further sampling to ovulation (0 h) in a time-based manner rather than to a morphological system. Thus, I anticipate that any comparisons between sexually and asexually produced embryos at equivalent time points will certainly represent different developmental stages.

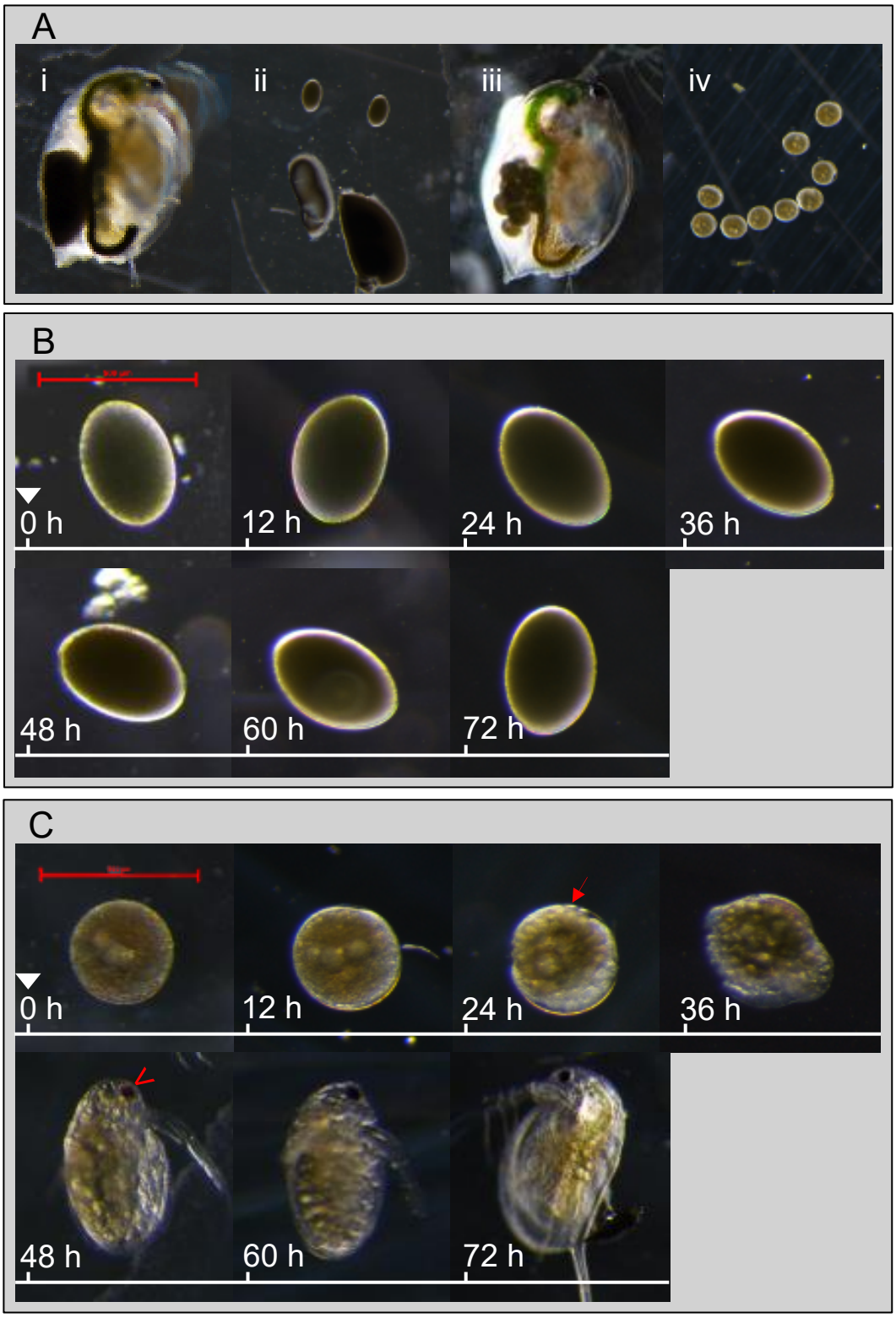


Figure 3-7 The embryonic development of *Daphnia magna* upon removal from the maternal brood pouch (A) of asexual (B) and sexual (C) offspring. White triangles indicate embryos immediately (< 15 min) following ovulation which was defined as 0 h. White bars and annotation indicate time (h) post-ovulation when imaging was performed. Scale bar = 500 µm**.** Annotations in red indicate morphological points of interest including the formation of antennae (➔) and compound eye formation (>).

3.5. Biomass optimisation study for the application of nanoelectrospray ionisation-direct infusion mass spectrometry (nESI-DIMS) to diapausing embryos

Nanoelectrospray ionisation-direct infusion mass spectrometry (nESI-DIMS) has already been successfully applied to extracts obtained from *D. magna* neonates and adults in various metabolomics studies (Taylor *et al.*, 2009; Toyota, Gavin, *et al.*, 2016; Taylor, White and Viant, 2017). However, its application to embryo samples has yet to be investigated. As embryos are significantly smaller than neonates, a considerable collection effort would be required to generate the equivalent biomass previously utilised. Here, I outline the optimisation of nESI-DIMS for the application to *D. magna* embryos based on the number and reproducibility of detected metabolites (spectral peaks) to assess whether reliable, information-rich spectra can be obtained using reduced embryo biomass.

The dry mass of a pool of *D. magna* sexual embryos was first determined in order to estimate the dry mass of single embryos (Table 3-2). My results yield an estimated dry mass of a single embryo as 5.3 µg (SD = 0.342 µg). The dry mass of a single neonate was previously reported in the region of 7.46 µg, which were obtained from neonates <24 h after release (Taylor *et al.*, 2008). Therefore, my results are in line with previous reports (ca. 66% of dry mass of one neonate) taking into account somatic growth post-hatching.

Table 3-2 Estimated dry mass of pooled Daphnia magna sexual embryos. SD = standard deviation.

Replicate	Dry mass 200 embryos (µg)	Estimated dry mass 1 embryo (µg)
1	1010	5.05
2	1170	5.85
3	1050	5.25
4	1060	5.3
Mean	1.073	5.363
SD	0.068	0.342

The current lower limit of detection within my laboratory workflow for metabolomics, thoroughly optimised for use in *D. magna* neonate and adult samples is regarded as 30 neonates (~250 µg). The equivalent biomass obtained from sexually produced embryos would require samples in the region of 50 embryos per pool. I performed a biomass optimization study with the systematic reduction in the number of embryos per sample from 50 down to single embryo samples (Table 3-3) to definitively establish a lower limit of detection for this sample type using current methodologies (Southam *et al.*, 2017). Analysis was performed on polar fractions in positive ion mode and non-polar fractions in negative ion mode in order to obtain representative spectra from chemically diverse species.

Table 3-3 Estimated dry mass of samples analysed for the biomass optimisation of *Daphnia magna* sexual embryos utilising nanoelectrospray ionisation-direct infusion mass spectrometry.

Number of embryos	Estimated dry mass (µg)	[Resuspension] (µg/µL)	Estimated mass to be analysed (µg)
50	250	8.33	50
25	125	4.17	25
10	50	1.67	10
5	25	0.83	5
1	5	0.17	1

Data were filtered to retain only peaks occurring in 3-out-of-3 technical replicates and remove features present in blank extracts. A high number of biological replicates were lost from the 5 embryo/sample group during positive ion analysis due to failed spray leaving only n = 5 replicates. In order to maintain integrity of interpretation, these samples were removed from further investigation. The total number of peaks that occurred within at least 80% of biological replicates within each sample type (i.e., the amount of biomass) were assessed and are reported in Table 3-4. This was used as an indication of the amount of biological information captured, as number of peaks were assumed to represent the number of metabolites detected.

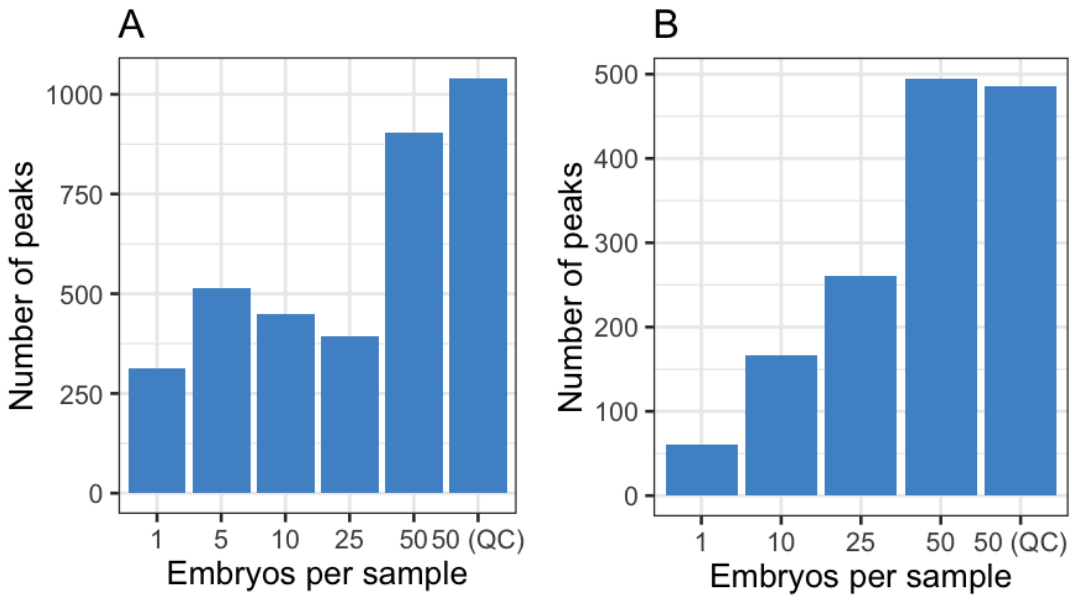


Figure 3-8 Number of peaks occurring in >80% of extract samples across embryo pools of different biomass. A) negative ion mode of non-polar fractions, B) positive ion mode of polar fractions.

Overall, the highest number of peaks in both datasets were detected in the quality control (QC) samples (Section 2.10) and the most concentrated sample which both contained 50 embryos. This was in line with my expectations, as I anticipated these groups would be the most successful as they are already in line with equivalent

sample biomass routinely investigated using DIMS within the laboratory. With a systematic reduction in the amount of biomass analysed, a reduction in the number of peaks retained in the data set was observed – presumably due to being below the limit of detection.

Peak intensity relative standard deviation (RSD) was used to assess the variability of peaks detected across the dataset as an indication of how reproducible a peak was within each sample class (i.e., the amount of biomass within the sample) (Parsons *et al.*, 2009). The RSD of all peaks occurring in at least 80% of samples within each sample class are displayed in Figure 3-9. The majority of calculated RSDs for peaks were less than 50%, however there were portion of outliers that were highly variable but these would be removed from any biological sample subjected to further analysis (Section 2.14.1.4). The median RSD of all peaks within samples classes are reported in Table 3-4.

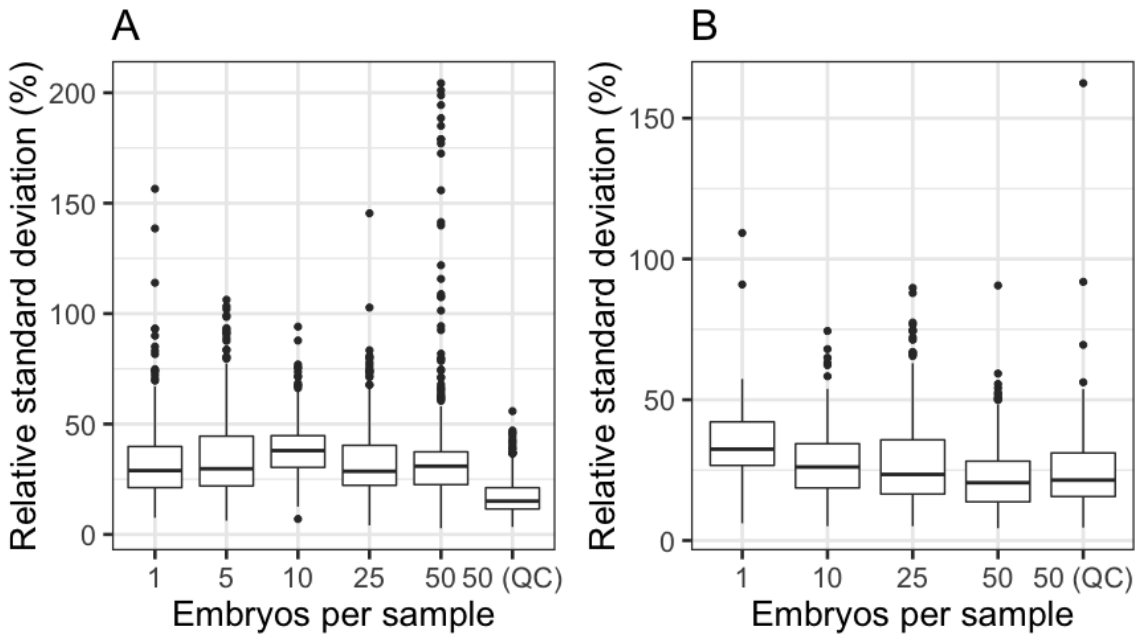


Figure 3-9 Relative standard deviation of all peaks occurring in >80% of samples across embryos pools of different biomass. A) negative ion mode of non-polar fraction, B) positive ion mode of polar fraction.

Peak variability as represented by RSD of peak intensities across biological replicates appears to largely independent of sample biomass, in negative ion mode analysis of non-polar fractions the median RSD of 1-embryo samples was slightly lower than samples containing 50 embryos (Table 3-4). Quality Control (QC) samples were more reproducible than biological samples in line with my expectations, as they are formed from a pool of extracted sample; variation in these samples demonstrates technical and instrumental variation only (no biological variation). Furthermore, eggs analysed were pooled from a variety of ages (duration of ephippia storage in darkness at 4°C). Therefore, unknown biological variation may account for some of the variation observed in terms of peak RSD.

Table 3-4 Table of median peak relative standard deviation (RSD) and peak count obtained by nano-electrospray ionisation-direct infusion mass spectrometry analysis of polar and non-polar extracts from different numbers of *Daphnia magna* sexual embryos. QC = quality control samples.

	Median Peak RSD (%)		Peak count	
Number of embryos	**Polar extract (positive ion mode)**	**Non-polar extract (negative ion mode)**	**Polar extract (positive ion mode)**	**Non-polar extract (negative ion mode)**
1	27.24	26.22	60	314
5	–	27.20	–	513
10	22.77	34.24	166	450
25	24.36	29.84	261	392
50	20.25	29.70	495	904
50 (QC)	20.89	15.23	486	1041

To conclude, when considering peak counts and the inherent biological variability of samples, biomass did not appear to have an appreciable effect on spectra reproducibility across biological replicates but did reduce the number of detectable peaks and therefore lead to a potential loss of biological information. Nevertheless, the practical requirements for the production of biological samples favour a reduction in the number of embryos per sample to increase the feasibility of larger high-throughput experiments, particularly if this permits the inclusion of further biological

replication. Therefore, as a compromise between culturing effort and number of features detected, a sample size consisting of 10 embryos per biological sample was selected for all further experiments. This represents an appreciable 5-fold reduction in the biomass used in nESI-DIMS metabolomics studies conducted previously, significantly increasing the feasibility of the application of metabolomics to low biological input studies.

3.6. Conclusions

A reduction in *C. vulgaris* food ration from 0.08 mgC Daphniid^{-1} to 0.03 mgC Daphniid^{-1} resulted in the induction of male offspring and ephippia production in the two clones of *D. magna* investigated and was thus selected as an appropriate, biologically relevant stressor for the laboratory induction of diapause in *D. magna.* The clone 'L7' showed a consistently high propensity for producing diapausing embryos and the number of male offspring required to ensure their fertilisation and subsequent viability. This isolate was therefore used for experiments throughout this thesis. A final culture density of 62 *D. magna* L^{-1}, with a male:female ratio of 1:4 and a reduced algal diet ration of 0.03 mgC Daphniid^{-1}, was subsequently chosen because of a demonstrated, consistently high and reproducible, yield of diapausing offspring. The lifespan of clone L7 under the food concentrations utilised was investigated, and, in combination with the confirmation of successful hatching of reproductive-capable offspring from ephippia up to 2 y-old, demonstrates a significant expansion of lifespan under defined environmental conditions previously observed in field-collected ephippia.

The morphological landmarks of ephippia production were characterised and based on the lack of observable changes in morphology of diapausing embryos during incubation in the maternal brood pouch, further investigations around the initiation and molecular events of early diapause were focussed on the hours post ovulation. Lastly, the biomass required for successful nESI-DIMS metabolomics investigation of *D. magna* embryos was assessed and suitable protocols for further experiments were derived based on a sample of 10 embryos, representing a significant reduction in previously analysed biomasses.

4. Cellular and molecular investigations reveal alternative rates of development and points of departure in subitaneous and diapause-destined embryos

4.1. Introduction

There is currently limited information available regarding diapause in *Daphnia sp.* beyond investigation of environmental incidence and induction from an ecological perspective (Boersma, Boriss and Mitchell, 2000; Arbačiauskas and Lampert, 2003; Vandekerkhove *et al.*, 2005; Conde-Porcuna, Ramos-Rodríguez and Pérez-Martínez, 2014; Navis *et al.*, 2015; Kaupinis *et al.*, 2016). Indeed, it is unknown at which point *D. magna* enter developmental arrest, although *D. pulex* is reported to arrest around 1000 cells, in the early gastrula stage (Baldass 1941). Herein I address this limitation, having established a stable induction protocol and described the major morphological stages of diapause initiation in Section 3 that suggested no observable changes in diapause-destined embryo morphology beyond the release of the ephippia from the maternal carapace. I aimed to identify the point of developmental arrest in diapause-destined embryos and compare the growth profiles of diapause-destined and subitaneous embryos in the immediate hours following ovulation, with extended sampling of sexually produced embryos after the release of the ephippium. This was accomplished by counting DAPI-stained nuclei of embryonic cells in the immediate hours following ovulation in both embryo types. A clear demarcation of the phases of diapause in *D. magna* is critical to any interpretation of data pertaining to early diapause development, in addition to the extrapolation of observations made in *Daphnia sp.* to diapause in general.

I also describe the collection of four large complementary omics datasets investigated throughout the remaining chapters of this thesis. Including RNA-Seq analysis of subitaneous and diapause-destined embryos collected at three early developmental time points (2 h, 10 h and 48 h post-ovulation) plus a further four time points (74 h, 82 h, 120 h and 408 h) of diapause-destined embryos after the release of the ephippium from the maternal carapace (see Figure 2-1 for full sampling scheme). The transcriptome data were paired with accompanying metabolomic data of the lipidome (negative ion mode) and of the polar metabolites (negative and positive ion mode) with the addition of three new time points (24 h, 98 h and 2 y) covering a total of four time points in subitaneous embryos and ten time points in diapause-destined embryos (Figure 2-1).

One avenue of investigation to understand the molecular processes of diapause is to directly compare the development of subitaneous and diapause-destined embryos. Previous studies have performed pairwise comparisons between adjacent diapause time points (Sim and Denlinger, 2009; Dong *et al.*, 2014; Koštál *et al.*, 2017) or directly compared diapausing with non-diapausing equivalent life stages (Poelchau *et al.*, 2013b; Kang *et al.*, 2016; Liu *et al.*, 2017; Y. Chen *et al.*, 2017). However, these investigations have predominantly utilised alternate photoperiods and temperatures to produce diapausing offspring while the impact of environmental conditions on molecular data has largely been ignored (Kankare *et al.*, 2016). Therefore, my experimental set up includes a strict, consistent control of the environmental conditions under which both diapause and subitaneous samples are maintained and sampled.

Specific unknowns about the process of arresting development associated with diapause motivate the molecular work presented in this chapter. (1) In embryonic diapause such as that of *Daphnia,* the induction of diapause is maternally controlled, and induction is followed by a period of 'pre-diapause' development (Koštál, 2006). What is occurring during this preparatory event that may ensure the success of diapause itself, including the sequestration of resources (Hahn and Denlinger, 2011; Monica F Poelchau *et al.*, 2013a; Y. rong Chen *et al.*, 2017)? (2) Diapause in *Daphnia* is unusual in the length of time that its embryos can remain viable during dormancy, well beyond the lifespan of subitaneous offspring (Section 3.3; Frisch *et al.*, 2014). Are there any specific molecular or physiological changes that occur during pre-diapause development or perhaps present from the outset that confers such an extreme physiological capability?

In order to answer these questions I assess the specific molecular changes occurring between subitaneous and diapause-destined embryos across a subset of the omics data shared between the two embryo types: 2–48 h post-ovulation, with particular reference to inherent differences in the metabolome or transcriptome between the two embryo types, in addition to features that change with time during early development.

4.2. Identification of the point of developmental arrest to demarcate the phases of diapause†

The earlier, preparatory, phase of my investigation (Section 3) revealed that minimal morphological changes occurred in sexually produced embryos over the 72 h post-ovulation, suggesting that developmental arrest likely occurs after ovulation yet before the ephippia is released. In order to unequivocally determine the point of developmental arrest, embryos produced by both sexual and asexual reproduction were collected from females following ovulation. These embryos were squashed and DAPI nuclear staining was applied and the number of stained nuclei counted. Nuclei were assumed to reflect the number of cells within the stained embryos and are reported in Section 9.1. Throughout, the number of cells counted was highly variable, as represented by the large standard deviations for these measurements (Section 9.1). This variability is expected to some extent, due to natural biological variation, human error during counting and the expectation that some cells were distorted, lost or incompletely stained during slide preparation. At early time points (< 30 cells) DAPI-staining was inconsistent, presumably due to poor reagent preparation or a low number of cells and therefore nuclear material and are not reported.

† Work presented in this section was performed in collaboration with Dr Linda Weiss and Luxi Chen at the Ruhr Universität Bochum, Germany. And has been accepted for publication in BMC Cell Biology: "Mitotic activity patterns and cytoskeletal changes throughout the progression of diapause developmental program in *Daphnia*" Luxi Chen; Rosemary E. Barnett; Martin Horstmann; Verena Bamberger; Lea Heberle; Nina Krebs; John K. Colbourne; Rocío Gómez; Linda C. Weiss

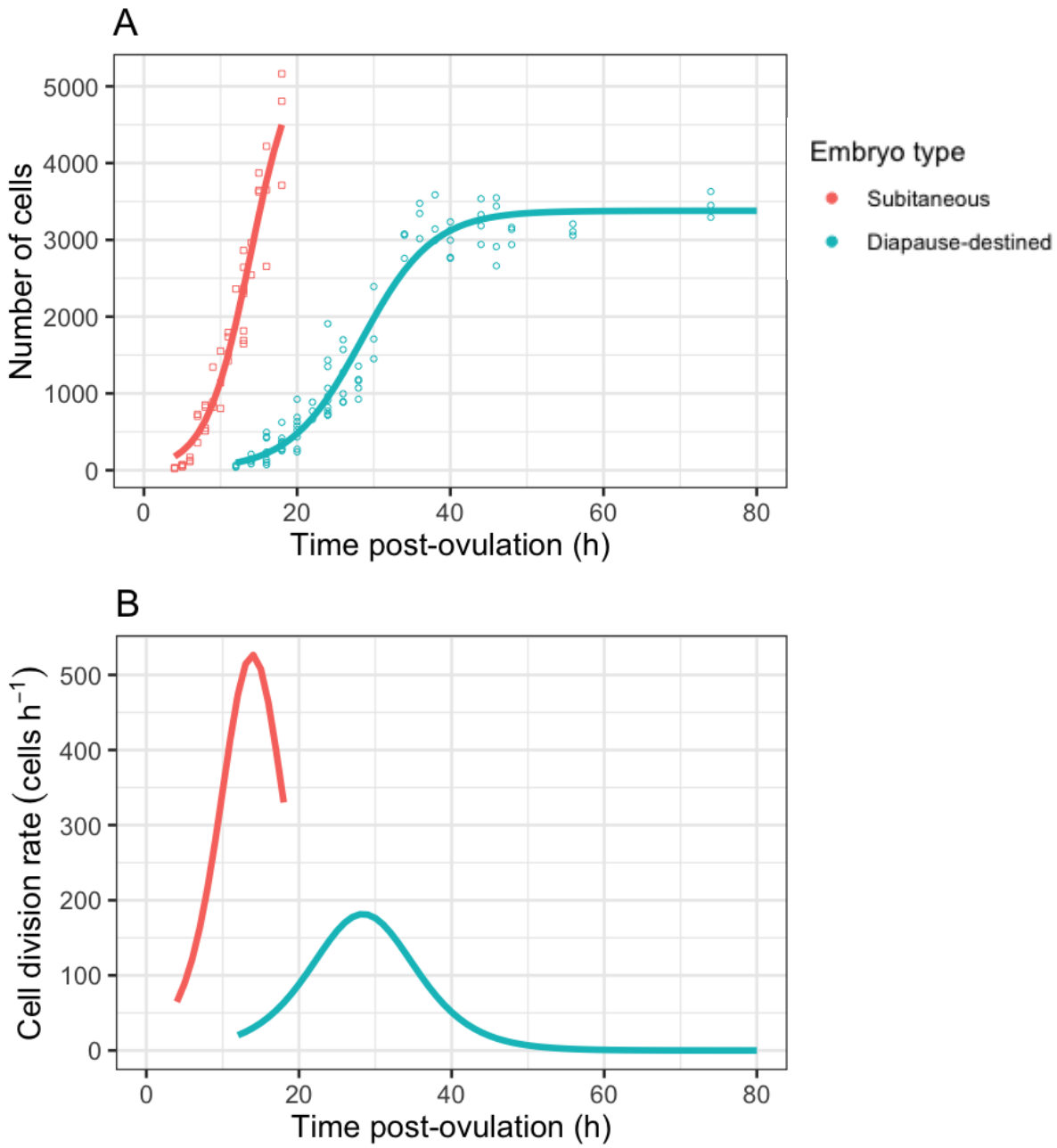


Figure 4-1 Cell count growth curves of subitaneous and diapause-destined embryos of *Daphnia magna*. A) Number of cells (DAPI-stained nuclei) in embryos sampled post-ovulation, points indicate raw counts from single biological replicates (n ≥ 3), lines represent respective logistic regression models for embryo types of cell count over time; B) Cell division rate of embryos over time estimated from first derivative with respect to time of logistic regression models of cell count over time.

The sexually produced diapause-destined embryos exhibited a cell division profile different from asexually produced subitaneous embryos. Cell number in subitaneous embryos increased with time in a classic exponential fashion over the first 18 h

following ovulation, from 27 cells (SD = 3.8) at 4 h post-ovulation to 4561 cells (SD = 756.5) 18h post-ovulation. By 25 h, it became unfeasible to accurately count the number of cells in subitaneous embryos, therefore slides were counted to a maximum of 7000 cells to confirm embryonic growth had continued but the true number of cells at this point and beyond were not determined. A decrease in the rate of cell division was observed by 21 h in subitaneous embryos (Figure 4-1; B), consistent with my expectations of the onset of cellular differentiation. In contrast, cell number in diapause-destined embryos demonstrated a sigmoidal curve with cell number increasing initially indicating direct development until ~30 h post-ovulation. A decrease in the rate of cell division then occurred indicating progression towards a stationary phase at 48 h and ~3500 cells, after which point no further appreciable increases in cell number were observed. Ephippia are shed ~72 h post ovulation, from this point onwards cell number was constant with a mean value of 3581 cells (SD = 171.5). A final sample was collected of embryos from ephippia that had been stored under the same culturing conditions for 408 h (2 weeks) and these samples contained 3705 cells (SD = 176.7).

Further differences in the early development of subitaneous and diapause-destined embryos were also apparent, notably cell division was slower in diapause-destined embryos in comparison with subitaneous (Figure 4-1; B), reaching an observable ~30 cells 8 h later (12 h and 4 h respectively). The maximal rate of cell division for subitaneous embryos was estimated to be ~500 cells h^{-1} and occurred 16 h post-ovulation, in contrast diapause-destined embryos reached a 2.5-fold lower maximal division rate of ~200 cells h^{-1} which occurred later at 28 h. Together these results suggest that cell division is slower in diapause-destined embryos.

According to current definitions, the onset of diapause is defined as the point at which developmental arrest occurs (Koštál *et al.*, 2009). There is limited evidence that the sexually produced embryos of *D. pulex* enter diapause at ~1,000 cells in early gastrula stage (Baldass 1941). The data I obtained suggest that the larger *D. magna* enter diapause at ~3,500 cells – before 48 h under the environmental conditions I utilized–however without corresponding whole mount staging I am unable to ascertain whether diapausing embryos are gastrula, but I anticipate that the point of diapause initiation will be conserved between these two closely related species.

Based on these results I am able to define the phases of pre-diapause and diapause in *D. magna.* From ovulation to 48 h post-ovulation, sexually-produced embryos are undergoing active cell division and are in the pre-diapause phase (Koštál *et al.*, 2009). The initiation of diapause in *D. magna* occurs around 3500 cells, approximately 48 h after ovulation (at 20°C), which is prior to the release of the embryo-containing ephippium from the maternal carapace–occurring ca. 72 h post-ovulation. Thus after 48 h post-ovulation, sexually produced *D. magna* embryos are in diapause and exhibit an arrest of active development.

4.3. Alternative developmental gene expression and metabolism

To investigate the molecular processes occurring during diapause in *D. magna* I collected four large complementary omics datasets. Sampling collection procedure is outlined fully in Section 2.12, but briefly included subitaneous and diapause-destined embryos collected at three early developmental time points (2 h, 10 h and 48 h post-

ovulation) plus a further four time points (74 h, 82 h, 120 h and 408 h) of diapause-destined embryos only, collected after the release of the ephippium from the maternal carapace. The transcriptome data were paired with accompanying metabolomic data of the lipidome (negative ion mode) and of the polar metabolites (negative and positive mode) with the addition of three new time points (24 h, 98 h and 2 y).

Embryo transcriptome sequencing (Section 2.13) generated over 511 M reads across 30 samples with an average of 17 M per sample; the mean proportion of reads per sample with a >Q20 phred score was 94% (Table 9-3). Following trimming and alignment of raw sequencing reads to the *D. magna* transcriptome (Orsini *et al.*, 2016) using both primary transcripts and alternative splice variants with *Salmon* (Section 2.13.2) the proportion of raw reads aligned was assessed. An average of 13% of reads per sample (Table 9-3) were binned due to low quality indicating that data obtained was of high quality, with a mean 97.26% (SD = 1.23) of trimmed reads per sample mapped to the transcriptome (Table 9-3). Following read alignment to the transcriptome, transcript counts were summarised to gene level and normalised using *DESeq2* (Love, Huber and Anders, 2014). Initial data quality assessment, removal of outliers and individual sample information is reported in Section 9.2. Considering all of the primary and alternative transcripts, a total of 114,009 unique transcripts were present within the available transcriptome corresponding to 29,134 genes. Of these, 21,399 genes (73.4%) (following the removal of 3323 with <10 counts across all samples (Section 9.2.1)) were expressed within at least one sampling time point within my dataset.

Metabolite data was collected as outlined in Section 2.14 to produce final results matrices with intensity data of peaks detected in the corresponding biological sample. Downstream statistical analysis of metabolite data requires complete matrices without missing values. In order to limit the amount of artificial variation created by missing value imputation, particularly for features that may be completely absent in specific sample classes, intensity matrices were filtered to retain features that were present in at least 80% of biological samples across all sample types. These filtering stages and an assessment of sample peak composition are presented in detail in Section 9.2. However in brief, filtration to remove peaks that were not present in at least 80% of all samples greatly reduced the number of metabolite features in the final matrix. This resulted from a combination of a large number of features unique to either subitaneous or diapause-destined embryos, in addition to a number of peaks that were detected in specific sample classes (e.g. single sampling time points). The removal of unique features represents an important limitation of the current analysis and experimental design in that conclusions regarding metabolic differences between diapause-destined and subitaneous embryos, in addition to over the course of development, can only be based on features that are shared amongst the various sample groups. In subsequent chapters, separate analyses were performed in order to capture information regarding diapause-specific features.

To visualise differences in gene expression and metabolites between diapause-destined and subitaneous embryos with time post-ovulation, principal component analysis (PCA) of log-transformed gene counts and peak intensity matrices was undertaken. As these capture the largest portions of variance within the data, the first two principal components (PCs) are shown in Figure 4-2. When considering gene

expression (Figure 4-2; A), there is a significant separation of both embryo types along PC1 with time post-ovulation, suggesting that this PC predominantly represents changes in gene expression over the course of development (Table 4-1). Both diapause-destined and subitaneous embryos cluster tightly together at 2 h post-ovulation, by 10 h they have formed two distinct clusters and separated on both PCs. Subitaneous embryos collected at 48 h post-ovulation and diapause-destined embryos collected at 48h post-ovulation and beyond show significant separation (Table 4-1) between embryo type along PC2 representing the divergent gene expression profiles of embryos destined for direct development and diapause. Cell division is at a minimum in sexually produced embryos by 48 h post-ovulation and at this point they have entered diapause. Interestingly I observed no clear separation between sexually produced embryos after this time suggesting similar gene expression throughout this (diapause) period.

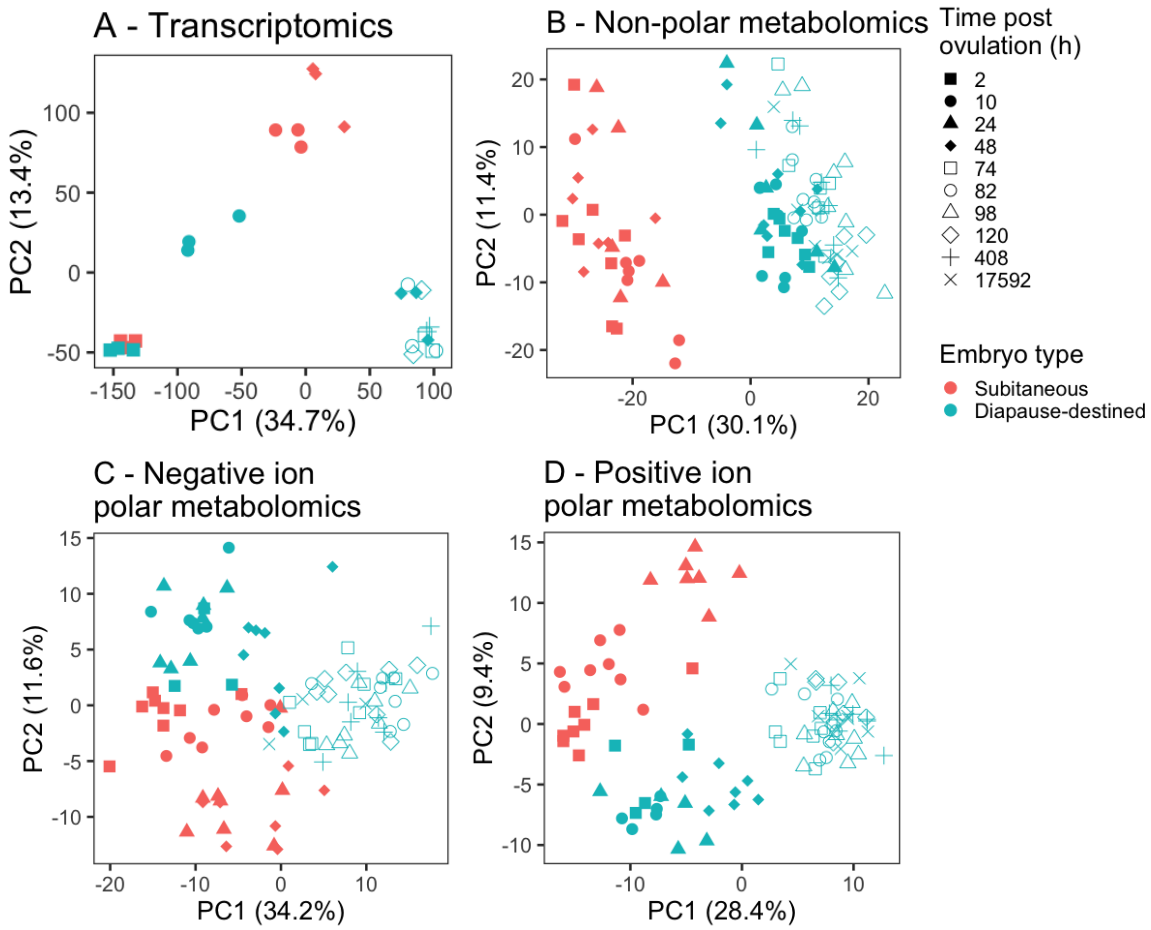


Figure 4-2 Principal component scores plots showing the first two components from PCA of log-transformed gene counts (A) and metabolite intensities (B-D) of diapause-destined (blue) and subitaneous (red) embryos of *Daphnia magna* collected between 2 h and ~2 y (17,592 h) post ovulation; B) Negative ion non-polar extracts, C) Negative ion polar extracts, D) Positive ion polar extracts.

When considering metabolite data, significant separation was observed between sexually and asexually produced embryos along the first principal component in all three studies. A separation between samples at different stages of development (time post-ovulation) was also observed along PC2 in polar extracts, however this was not statistically significant. This separation was between embryos in pre-diapause development (2 h–48 h; filled shapes) and diapause (>74 h; hollow shapes). Together this suggests that the metabolic differences between embryo types and between diapause and pre-diapause embryos are more dominant than any

dynamic changes of the metabolome between time points over the course of development. Furthermore, the portion of the lipidome of embryos that I am able to consider in my dataset appears to be less responsive during development but demonstrates clear differences between diapause-destined and subitaneous embryos.

Table 4-1 Summary of the first and second principal components and significant embryo type and time separation by mean sample group scores determined by ANCOVA (type III SS) from PCA of log-transformed gene counts and metabolite intensities of diapause-destined and subitaneous embryos of *Daphnia magna* collected between 2 h and ~2 y (17,592 h) post ovulation; Italic values = $p < 0.05$, red values = $p < 0.01$

	Principal component	Standard deviation	Proportion of Variance	*p*-values		
				Time (h)	Embryo type	Time *x* Embryo type
Transcriptomics	**1**	96.0981	0.3472	*0.0083*	0.2983	0.1284
	2	59.6928	0.134	*0.0390*	*2.26E-08*	*3.71E-09*
Negative ion non-polar metabolomics	**1**	15.4710	0.3014	0.2293	*4.34E-44*	0.5341
	2	9.3620	0.1104	0.5808	*0.0415*	0.1580
Negative ion polar metabolomics	**1**	9.6485	0.3423	0.8119	*8.50E-08*	*0.0183*
	2	5.6138	0.1159	0.1268	*3.08E-13*	*3.07E-06*
Positive ion polar metabolomics	**1**	8.9731	0.2835	*0.0161*	*1.52E-13*	*0.0029*
	2	5.1624	0.0938	*0.0103*	*4.49E-15*	*2.15E-10*

In order to further assess the differences in gene expression and metabolism between directly developing and diapause-destined embryos, data were subset into time points shared between the two embryo types only and PCA was repeated. This data covers early development between 2 h–48 h post-ovulation. Sampling of sexually produced embryos continued to 74 h and beyond into the period of diapause whereas, by this point, subitaneous embryos have hatched into free-swimming neonates and were thus not comparable. When only time points 2 h–48 h were considered in PCA, in the analysis of gene expression data (Figure 4-2; A), PC1 continued to indicate a separation of embryo samples with developmental time whilst

the second component led to a clear demarcation between subitaneous and diapause-destined embryo samples collected at 48 h. Considering the lipidome data (Figure 4-2; B) in scores plots, the overall clustering and positioning of samples changed very little and the proportion of variance explained by each PC increased minimally.

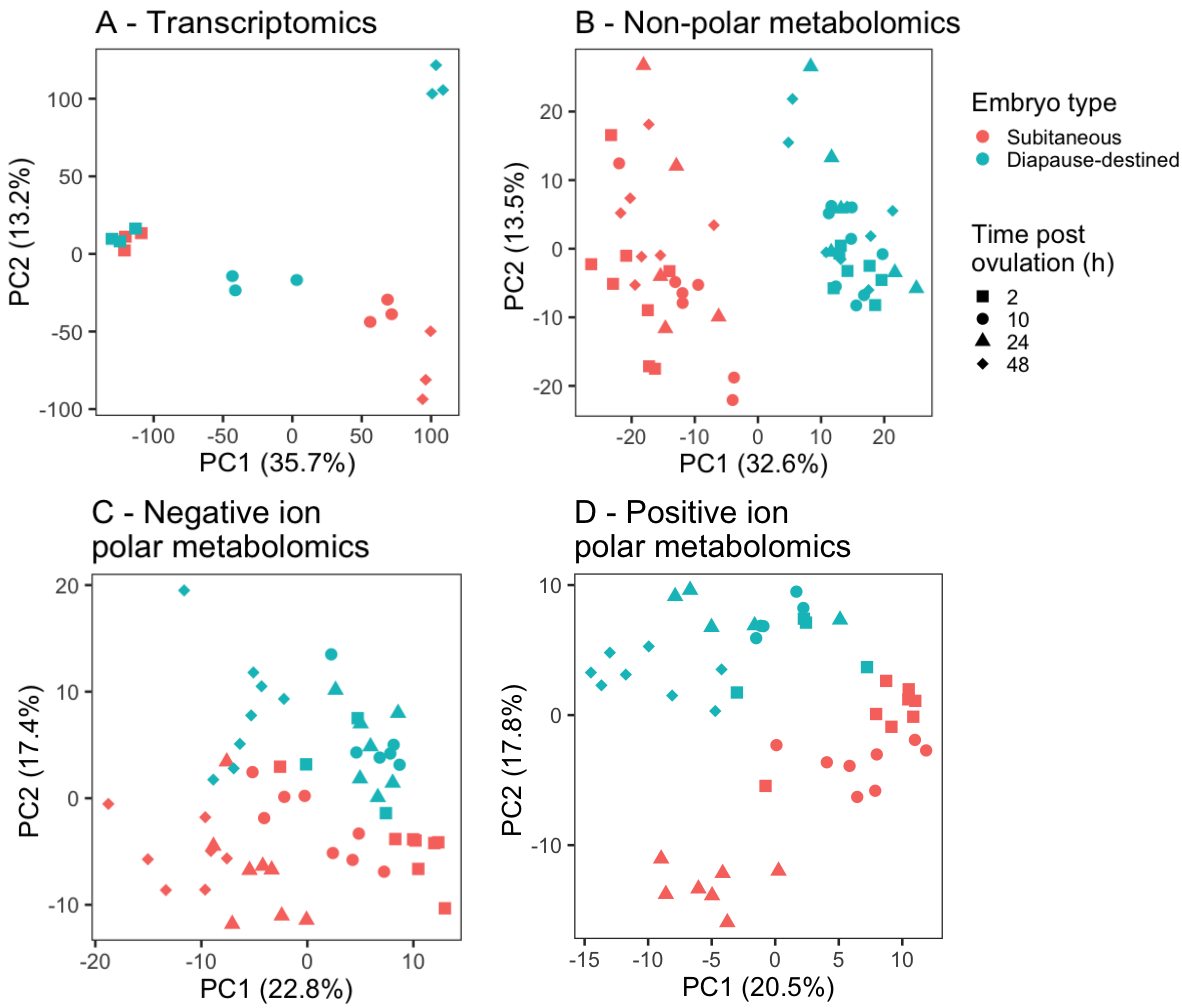


Figure 4-3 Principal component scores plots showing the first two components from PCA of log-transformed gene counts (A) and metabolite intensities (B-D) of diapause-destined (blue) and subitaneous (red) embryos of *Daphnia magna* collected between 2 h and 48 h post ovulation; B) Negative ion non-polar extracts, C) Negative ion polar extracts, D) Positive ion polar extracts.

The removal of later diapausing samples from PCA of polar analyses did yield evidence of metabolic changes with developmental time, as a significant separation of samples based on developmental time was observed along PC1 in both polar

analyses (Table 4-2) that were not captured when all sampling points were considered. For negative ion analysis, there was a more pronounced separation between subitaneous samples at different time points through development, with 2 h samples clustering tightly together, an intermediate cluster comprised of 10 h and 24 h samples and a further separate cluster of embryos collected at 48 h. Sexually produced samples collected at 48 h post ovulation, also clustered separately to earlier developmental stages although there was no clear separation between, 2 h, 10 h and 24 h sexual samples. Despite the loss of some biological replicates, including the entire asexual 48 h group, a significant relationship between developmental time and PC1 scores was also observed in the positive ion analysis of polar samples (Table 4.2). Asexually produced samples collected at 2 h clustered tightly together, 10 h samples displayed less distinct clustering falling between 2 h and 48 h, which also formed a unique cluster of samples. There was also a significant separation of diapause-destined samples with developmental time moving along PC1 in the same direction as subitaneous samples, although the two trajectories separate out along PC2 suggesting that disparate metabolic changes are occurring between the two embryo types over the course of development between 2 h and 48 h.

Table 4-2 Summary of the first and second principal components and significant embryo type and time separation by mean sample group scores determined by ANCOVA (type III SS) from PCA of log-transformed gene counts and metabolite intensities of diapause-destined and subitaneous embryos of *Daphnia magna* collected between 2 h and 48 h post ovulation; Italic values = $p < 0.05$, red values = $p < 0.01$

	Principal component	Standard deviation	Proportion of Variance	*p*-values		
				Time (h)	Embryo type	Time *x* Embryo type
Transcriptomics	**1**	98.037	0.357	*2.91e-12*	*0.0002*	*2.27e-05*
	2	59.6599	0.1322	*2.56e-05*	*1.69e-08*	*9.57e-09*
Negative ion non-polar metabolomics	**1**	16.0774	0.3256	0.5950	*4.20E-27*	0.5523
	2	10.3460	0.1348	*0.0060*	0.1953	0.7588
Negative ion polar metabolomics	**1**	7.8686	0.2276	*5.44E-15*	*0.0001*	*0.0471*
	2	6.8833	0.1742	0.5546	*2.75E-12*	*0.0379*
Positive ion polar metabolomics	**1**	7.6356	0.2053	*1.33E-10*	*0.0266*	*0.0005*
	2	7.1065	0.1778	*6.65E-10*	*2.12E-21*	*1.06E-11*

4.4. Identification of metabolic and transcriptomic differences between subitaneous and diapause-destined development

Previous studies assessing patterns of gene expression during diapause have largely consisted of multiple pairwise comparisons between adjacent time points (Sim and Denlinger, 2009; Dong *et al.*, 2014; Koštál *et al.*, 2017) or direct comparison with non-diapausing equivalent life stages (Poelchau *et al.*, 2013b; Kang *et al.*, 2016; Liu *et al.*, 2017; Y. Chen *et al.*, 2017). However, I identified alternative developmental trajectories (Section 4.1), including differences in the rate of development between subitaneous and diapause-destined embryos. Therefore, their direct comparison would not necessarily reflect diapause associated gene expression as it would also be compounded by differences in ontogeny. Therefore, differential gene expression analysis was performed using *DESeq2* (Section 2.13.3) on log-transformed gene counts to obtain lists of genes that showed differential expression between embryo types, over time and the interaction thereof, across all three sampled time points at a

false discovery rate of 1%. To assess metabolic changes, ANOVA-simultaneous component analysis (ASCA) was applied to generalized-log (glog) matrices to assess peak intensities that changed significantly with embryo type, time or the interaction thereof, based on their leverage and squared prediction error (SPE). Accompanying permutation testing was applied to validate the significance of these models. For brevity, ASCA leverage/SPE scatter plots and permutation testing results are presented in Section 9.4.

Table 4-3 Summary of the number of significantly changing metabolites or genes from ANOVA-simultaneous component analysis and differential gene expression analysis respectively, associated with time post-ovulation, embryo type (subitaneous or diapause-destined) and the interaction thereof. Significantly changing metabolites were determined based on a leverage threshold of 0.9 and squared prediction error alpha level of 0.05. Significantly changing genes were selected based on a false-discovery rate corrected *p*-value threshold of 0.01.

	Total	Time (h)	Embryo Type	Type *x* Time (h)
Negative non-polar metabolomics	794	N/A	25	N/A
Negative polar metabolomics	272	8	6	18
Positive polar metabolomics	284	12	11	7
Transcriptomics	20,142	5,720	494	2,296

Permutation testing revealed that there were no peaks detected in the negative ion analysis of non-polar extracts that demonstrated significant changes with time post-ovulation reflecting the minimal changes with time identified during PCA of these samples. Furthermore, permutation testing of the observed interaction results for both polar analyses was inconclusive ($0.05 > p > 0.01$); therefore, ASCA models for this component are interpreted with caution. Of the total number of peaks detected in each analysis, only a small percentage (>6%) changed significantly between embryo type and/or time, demonstrating that the metabolic differences between sample types are attributable to only a small number of features.

The largest changes in gene expression were predominately those with time; approximately 28% of the total 20,142 genes within the dataset were differentially expressed with developmental time (Table 4-3). There is limited information regarding how much *Daphnia magna* gene expression alters during early development. But around 15% of genes were found to be differentially expressed between <24 h neonates and 6 d animals (Li *et al.*, 2017). However, there are now numerous studies that have begun to investigate the transcriptome during early embryogenesis in many species including classical models such as *Drosophila* (Graveley *et al.*, 2011), and *Danio rerio* (Vesterlund *et al.*, 2011), as well as non-model species such as the damselfly *Ischnura elegans* (Simon *et al.*, 2017). Throughout these studies, gene expression has been shown to be highly dynamic throughout early development. For example, roughly a third of genes show increasing or decreasing expression with time post-ovulation when observing the development of *D. rerio* at a fine-scale time resolution between late blastula to mid-gastrula development (Rauwerda *et al.*, 2017). Thus, considering the major developmental processes occurring, it is not surprising that such a large portion of genes within my dataset were differentially expressed over the first 48 h of embryonic development, regardless of diapause or subitaneous development. A further 11% of genes showed some interaction between time and embryo type on their expression and just 2% of genes were differentially expressed between embryo types, amounting to 494 genes. Together, the results from metabolomics and transcriptomics suggest that a relatively small subset of molecular events is driving the alternative physiological processes that lead towards the manifestation of embryonic diapause in *D. magna*.

When comparing organisms destined for diapause or subitaneous development, other RNA-Seq studies have reported varying portions of differentially expressed genes. RNA-Seq and proteomic analysis of *Locusta migratoria* embryos revealed 29,671 transcripts and 296 proteins were differentially expressed between diapause maintenance and non-diapause embryos, representing 15% of transcripts and 14% of proteins detected (Tu, Wang, Hao, Douglas W. Whitman, *et al.*, 2015). However, in this study diapause and non-diapause embryos were produced under different temperature regimes. Which may explain why the authors report a larger proportion of differentially expressed genes than I identify in my study (where the two embryo types were produced under the same environmental conditions). Analysis of the ovaries of *B. mori* females showed that 289 genes (2% of the 14,624 genes expressed) were differentially expressed between those ovaries producing directly developing offspring and those destined for entry into diapause. Of these, 183 genes were up–regulated and 106 were down–regulated (Y. rong Chen *et al.*, 2017). In pairwise comparisons between diapause-destined and directly developing larvae of *Chymomyza costata* 7.7% of genes from a microarray subset of 1,042 genes were differentially expressed (Koštál *et al.*, 2017), in line with a previous RNA-Seq experiment in the same species that reported 6.2% of genes with differential expression (Poupardin *et al.*, 2015). By contrast, comparing adult reproductive diapause of *Drosophila montana,* authors reported about a third of transcripts were differentially expressed between diapausing and non-diapausing females with roughly half up- and half down-regulated (Kankare *et al.*, 2016). However, this may partially reflect the increased complexity and additional regulatory events that may be required to preside over the developmental arrest of a formed adult.

4.5. Discovering the function of diapause associated genes

To gain insight into the broad functions of differentially expressed gene sets, over-representation analysis of their gene ontologies (GO annotations) was performed (Section 2.13.4) after the creation of a custom GO database scoring *D. magna* transcriptome predicted protein-coding genes against the library of protein classifications curated in the Protein ANalysis THrough Evolutionary Relationships (PANTHER) database (Mi *et al.*, 2017). The number of differentially expressed genes with an assigned PANTHER ID and the number of unique functional categories assigned under each domain are reported in Table 4-4.

Table 4-4 Summary of the GO-annotation of differentially expressed genes after false discovery rate correction (*p* < 0.1) between time post-ovulation (h), embryo type (subitaneous versus diapause-destined) and the interaction thereof. Numbers represent the unique functional categories assigned under each domain. Separate "background' column indicates the total number of genes within the *Daphnia magna* transcriptome, used as a reference set for over-representation analysis.

	Time (h)	Embryo Type	Type *x* Time (h)	Background
Differentially expressed genes (*p*-adjust < 0.1)	9277	1642	4771	29244
Differentially expressed genes with PANTHER ID	6903	1400	3738	19244
Biological process	105	106	124	235
Molecular function	98	75	85	182
Cellular component	24	27	23	64
Total	227	208	232	481

No GO-annotations were significantly overrepresented after FDR correction in differentially expressed genes with time or interaction between embryo type and time. A full list of GO-annotation enrichment results is reported in Section 9.5. Considering the genes differentially expressed with time, there were only three terms that had an unadjusted *p*-value <0.1 and none below 0.05; two of these were related to DNA, 'DNA metabolic process' and 'DNA repair' containing 157 and 84 genes

respectively. This result likely represents the increased replication of DNA occurring during mitosis and is reflective of the on-going embryonic development occurring throughout the time points sampled (Simon *et al.*, 2017). A third functional annotation-term relating to transferase activity of glycosyl groups was also marginally overrepresented and contained 64 genes. Of the genes whose expression showed a relationship with the interaction between embryo type and time post-ovulation, 8 GO-annotations had a non-adjusted *p*-value <0.05 including generic terms for catalytic and transferase activity containing 840 and 279 genes respectively. Thirty genes classified under microtubule binding and 140 relating to the cell cycle were also overrepresented and may possibly reflect the alternative rates of cellular division between the two embryo types identified in Section 2.11.

Genes differentially expressed between each embryo type were divided into two groups, based on whether they were up-regulated or down-regulated in diapause-destined embryos. GO enrichment was then performed on each group separately, and significantly over-represented GO-annotations and their fold enrichment are shown in Figure 4-4.

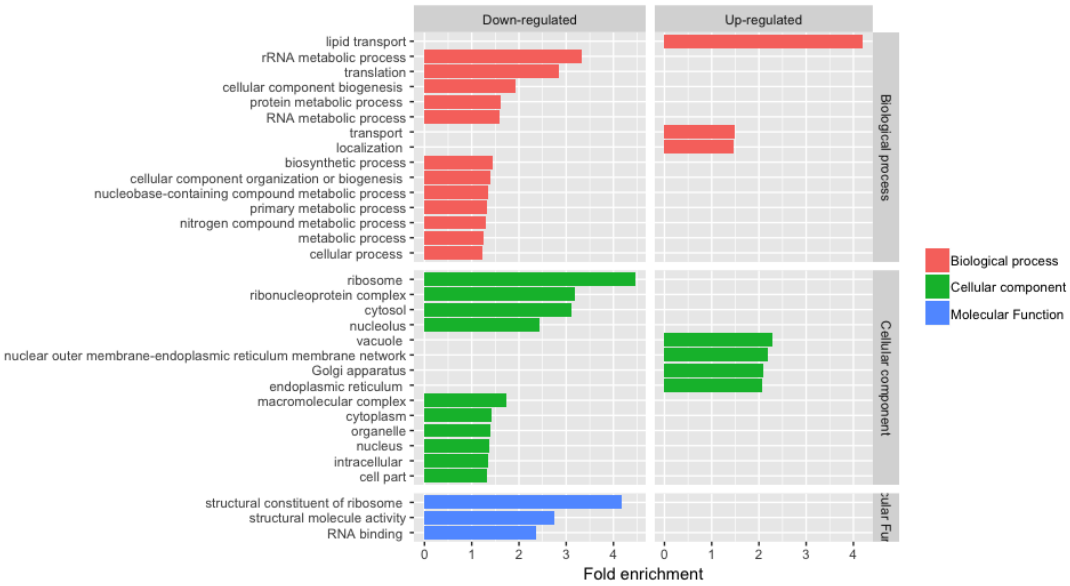


Figure 4-4 Gene ontology (GO) over-representation analysis of genes differentially expressed between subitaneous and diapause-destined embryos of *Daphnia magna*. 'Up-regulated' genes refer to those with higher expression in diapause-destined embryos and vice versa for 'down-regulated' genes.

Several of the overrepresented GO-annotations were related to protein translation and ribosomal processes. Ribosomes are the site of the translation of RNA into proteins and in eukaryotes consist of a large (60S) and small (40S) subunit with hundreds of associated proteins and co-factors (Tibshirani *et al.*, 2011; Peña, Hurt and Panse, 2017). Within the GO term 'translation', there were 45 genes that showed differential gene expression between subitaneous and diapause-destined embryos; 37 were assigned definitive membership of a protein family and a heat-map of their log-normalised expression is shown in Figure 4-5. Predominately, these genes display a dynamic behaviour throughout early embryogenesis. In asexually derived embryos, there are notable step changes of increasing gene expression with increasing time post-ovulation. By contrast, expression of these genes relating to translation is lower in sexually produced embryos, and there appears to be limited increases in their gene expression between 2 h and 10 h post-ovulation, as observed

in asexually derived embryos. At 48 h post-ovulation, there are discernible increases in gene expression in sexually derived embryos. However, these levels are still lower than that of asexually derived embryos and, furthermore, are more similar to expression levels in 2 h subitaneous embryos that of 10 h or 48 h. Many proteins associated with each of the ribosomal subunits were identified, including several transfer RNA ligases –enzymes responsible for the esterification of amino acids and corresponding tRNAs (Hopper, Pai and Engelke, 2010)– and translation initiation and elongation factors. Interestingly, the observation that these genes are comparatively higher in diapause-destined embryos at 48 h than at 2 h and 10 h, at which point cellular diapause arrest has already manifested, suggests that there are still many active molecular processed including protein synthesis occurring after entry into diapause.

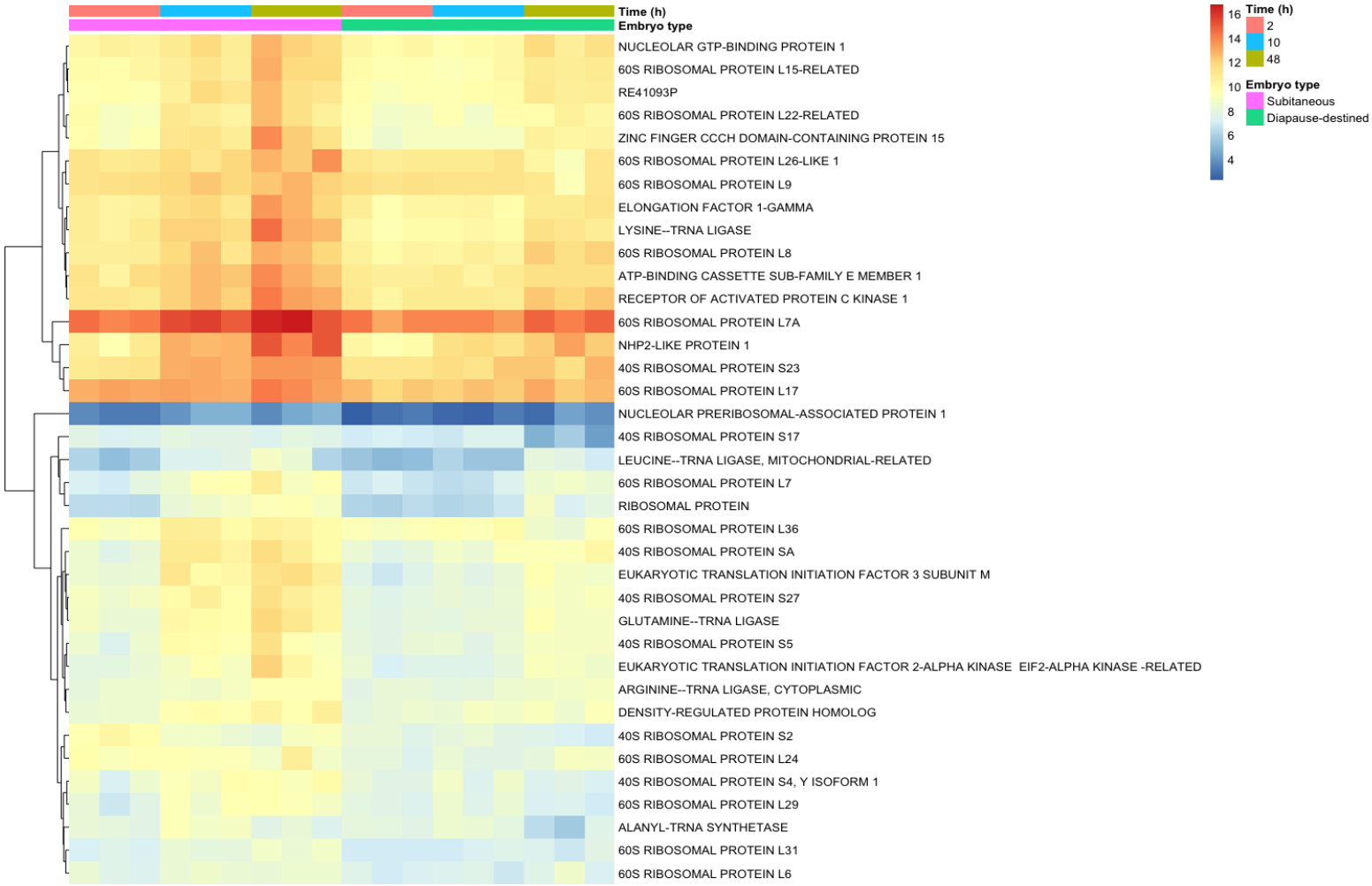


Figure 4-5 Heat map and hierarchical clustering of genes within the over-represented (p-adjusted <0.05) translation gene ontology annotation that displayed differential expression in samples of subitaneous and diapause-destined *Daphnia magna* embryos collected between 2 h and 48 h post-ovulation. Row labels indicate PANTHER-assigned gene family, coloured column top panels indicate embryo type and time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale.

The higher expression of genes relating to translation may largely reflect a higher rate of protein synthesis occurring during early development in subitaneous embryos which show a higher rate of cell division in comparison with sexual embryos (Figure 4.1). Genes associated with ribosome biogenesis and translation have previously been reported to show enriched expression during early embryogenesis of *Ischrura elegans* (Simon *et al.*, 2017). Formation of the large complex ribosome is an energetically expensive process. Therefore, protein synthesis prior to the onset of diapause may be limited to only those proteins necessary to the current development, diapause initiation, and those required for the period of diapause itself (Rinehart *et al.*, 2007; Jackson, Hellen and Pestova, 2010).

Proteomic analysis of diapausing and non-diapausing *B. mori* eggs revealed several proteins involved in translation that were only identified in the proteome of non-diapausing eggs, including eukaryotic translation initiation factor 2 (gamma subunit), and 3 (subunit 2 beta) and eukaryotic translation termination factor 1 (Fan *et al.*, 2013a). Expression of translation elongation factor 1 (delta) is down-regulated in *Helicoverpa armigera* pupae destined for entry into diapause versus directly developing pupae, although the authors did not identify any other genes related to translation (Bao and Xu, 2011). By contrast, proteomic analysis of diapausing and non-diapausing *Sarcophaga crassipalpis* pupae revealed an up-regulation of levels of translation initiation factor 5 in diapausing pupae with relative levels increasing over the 25 d duration observed (Pavlides, Pavlides and Tammariello, 2011).

Regulation of protein synthesis primarily occurs at the stage of initiation, which requires the successful assembly of the ribosome. Eukaryotic translation initiation factor 3 (eIF3) is a large multi-subunit protein which binds the 40S ribosomal subunit

and together with other factors promotes recruitment of the complex to mRNA molecules (Bommer *et al.*, 1991; Jackson, Hellen and Pestova, 2010)). My analysis revealed that *eIF3* subunit M was down-regulated in diapause-destined embryos. There is growing evidence that implicates *eIF3* in other functions beyond transcription initiation: over-expression of *eIF3* has been linked to human malignancies and *eIF3* (subunits a and h) morphant *D. rerio* embryos display visual and neuronal developmental abnormalities (Choudhuri, Maitra and Evans, 2013; Lee, Kranzusch and Cate, 2015). It has been suggested that *eIF3* acts as a regulator of translation, and has been shown to bind to mRNA, via 5' untranslated regions, of genes involved in cell cycling, differentiation and proliferation (Lee, Kranzusch and Cate, 2015). Furthermore, RNAi-mediated knockdown of *eIF3* k and l subunits, whilst having no observable effects on protein translation leads to a 40% increase in organism lifespan in *C. elegans* (Cattie *et al.*, 2016). This evidence suggests that the role of *eIF3* in diapause development in *Daphnia magna* warrants further investigation, particularly in relation to any potential role in the regulation of development and the extended lifespan identified in Section 3.3.

4.6. Discovering the function of diapause associated metabolites

For the annotation of metabolite data, peak *m/z* values and intensities from pre-glog matrices were collated for each analyses and input into BEAMS then matched against the LipidMAPS (non-polar extracts) and HMDB (polar extracts) databases using a maximum 2.5 ppm error cut off (Section 2.14.1.5) to assign putative empirical formulae and metabolite names. Annotation of mass spectra remains a major limiting factor in metabolomics experiments (Gertsman and Barshop, 2018). Consequently,

only a portion of the detected features in each dataset were assigned a putative compound ID. Additional peaks were assigned putative molecular formula without a corresponding ID. For non-polar extracts, a total of 285 out of 794 peaks were assigned at least one putative ID. Annotation of polar extracts was less successful, although this partially represents the smaller number of peaks in the initial intensity matrix with 97 out of 272 for negative ion analysis and 77 out of 272 peaks for positive ion analysis. Furthermore, numerous peaks, particularly those at higher *m/z* ranges, were assigned multiple formulae/compound IDs. Without further investigation (e.g. peak fragmentation), all putative IDs must be assumed equally possible.

The significantly changing peaks in each analysis that were successfully assigned a putative compound ID are presented in Table 4-5 to Table 4-7. For negative ion analysis of non-polar extracts, 10 of the 25 significantly changing peaks were successfully annotated (Table 4-5). For polar extracts analysed in negative ion mode, 21 peaks were successfully assigned at least one putative molecular formula with 8 of these also being assigned a putative ID (Table 4-6). And positive ion analysis annotation yielded only 8 peaks out of 30 with a molecular formula–4 of which were assigned at least one putative ID (Table 4-7).

Table 4-5 Table showing putative annotation of spectral peaks from negative ion analysis of non-polar *Daphnia magna* embryo extracts that showed significant changes in intensity between subitaneous and diapause-destined embryos based on ANOVA simultaneous component analysis (ASCA) models with a leverage threshold of 0.9 and squared prediction error (SPE) alpha level of 0.05. Observed peak *m/z* is shown in addition to ASCA model statistics (Leverage and SPE). Formula indicates putative empirical formulae assigned to each peak using BEAMS, with corresponding annotation including number of possible species attributable to the aforementioned formula, respective putative compound identity with theoretical mass. Mass error (ppm) is the difference between theoretical and experimental *m/z*, normalised to experimental *m/z.* Other putative formula without compound ID assignment are indicated, for peaks with a high number of possible formulae only number of formulae are indicated for brevity.

m/z	Leverage	SPE	Formula	Number of possible species	Ion form	Compound (or compound class) identity	Theoretical mass	Mass error (ppm)
773.51604	0.00553	2.47E-30	–	–	–	–	–	–
832.60643	0.00623	1.97E-30	C50H85N1O7	1	[M+Na-2H]-	none	832.6072726	1.02
			C43H84NO8P	48	[M+Hac-H]-	Phosphatidylethanolamine,Phosphatidylcholine	832.6073106	1.06
			C45H88NO10P	6	[M-H]-	Phosphatidylserine	832.6073106	1.06
			C47H89NO8	2	[M+K-2H]-	Glucosylceramide(d18:2/23:0)	832.6074256	1.20
			C46H87NO9	1	[M+Cl]-	Glucosylceramide(d16:2(4E,6E)/24:0(2OH))	832.6074856	1.27
887.68827	0.00775	2.47E-31	C55H96O6	175	[M+Cl]-	Triacylglyceride	887.6900916	2.06
890.70711	0.00627	4.19E-30	–	–	–	–	–	–
915.75526	0.00738	4.93E-30	–	–	–	–	–	–
935.73378	0.00701	1.97E-30	C57H96O6	172	[M+Hac-H]-	Triacylglyceride	–	–
937.74906	0.01107	7.89E-30	C57H98O6	201	[M+Hac-H]-	Triacylglyceride	–	–
940.76825	0.00973	4.19E-30	–	–	–	–	–	–
941.76963	0.00916	0	–	–	–	–	–	–
942.77340	0.00583	4.93E-30	–	–	–	–	–	–
942.78379	0.00693	8.38E-30	–	–	–	–	–	–
1084.95711	0.00554	1.23E-30	–	–	–	–	–	–
1107.95421	0.00741	4.93E-30	C69H124O6	65	[M+Hac-H]-	Triacylglyceride	1107.953644	-0.51
1108.95663	0.00732	4.93E-31	–	1	M+13C	–	–	–
1109.96996	0.00894	1.23E-30	C69H126O6	44	[M+Hac-H]-	Triacylglyceride	1109.969294	-0.60
1110.97244	0.00887	4.93E-30	–	1	M+13C	–		
1135.98646	0.01187	6.16E-30	C71H128O6	26	[M+Hac-H]-	Triacylglyceride	1135.984944	-1.34
1136.98908	0.01101	1.23E-30	–	1	M+13C	–	–	–
1137.99095	0.00765	4.93E-30	–	0	M+2(13C)	–	–	–
1157.96954	0.00626	4.93E-31	C73H126O6	3	[M+Hac-H]-	Triacylglyceride	1157.969294	-0.21
1160.98793	0.00729	6.16E-30	–	1	–	–	–	–
1162.00138	0.01005	4.93E-31	C73H130O6	10	[M+Hac-H]-	Triacylglyceride	1162.000594	-0.68
1163.00374	0.00958	7.89E-30	–	1	M+13C	–	–	–
1164.01709	0.00905	9.86E-31	C73H132O6	7	[M+Hac-H]-	Triacylglyceride	1164.016244	-0.73
1165.01942	0.00878	7.15E-30	–	1	M+13C	–	–	–

Table 4-6 Table showing putative annotation of spectral peaks from negative ion analysis of polar *Daphnia magna* embryo extracts that showed significant changes in intensity between subitaneous and diapause-destined embryos (embryo type), time (h) post-ovulation or the interaction thereof based on ANOVA simultaneous component analysis (ASCA) models with a leverage threshold of 0.9 and squared prediction error alpha level of 0.05. Observed peak *m/z* is shown in addition to ASCA model statistics (Leverage and SPE). Formula indicates putative empirical formulae assigned to each peak using BEAMS, with corresponding annotation including number of possible species attributable to the aforementioned formula, respective putative compound identity with theoretical mass. Mass error (ppm) is the difference between theoretical and experimental *m/z*, normalised to experimental *m/z.* Other putative formula without compound ID assignment are indicated, for peaks with a high number of possible formulae only number of formulae are indicated for brevity.

	Embryo type		Time (h)		Time *x* Time (h)							
m/z	**Leverage**	**SPE**	**Leverage**	**SPE**	**Leverage**	**SPE**	**Formula**	**Number of possible species**	**Ion form**	**Compound (or compound class) identity**	**Theoretical mass**	**Mass error (ppm)**
251.07883					0.030	3.29E-03	C10H12N4O4	2	[M-H]-	Deoxyinosine Nebularine	251.0785796	-1.01
							5	–	–	–	–	–
258.04755			0.035	2.33E+00			–	–	–	–	–	–
260.02391			0.018	1.22E-01			12	–	–	–	–	–
260.13316	0.020	3.69E-30					C12H23NO3S	1	[M-H]-	–	260.1325896	-2.19
							C10H19NOS	1	[M+Hac-H]-	–		
266.09012					0.029	1.84E-01	C10H13N5O4	4	[M-H]-	Adenosine Deoxyguanosine Vidarabine Zidovudine	266.0894786	-2.40
							C8H9N5O2	1	[M+Hac-H]-	6-Amino-9H-purine-9-propanoic acid	266.0894786	-2.40
							C7H16N7OP	1	[M+Na-2H]-	–	266.0900646	-0.20
							C11H22NPS	1	[M+Cl]-	–	266.0904606	1.29
							C10H23NOP2S	1	[M-H]-	–	266.0902856	0.63
282.08486					0.033	7.20E-02	C10H13N5O5	2	[M-H]-	8-Hydroxy-deoxyguanosine Guanosine	282.0843936	-1.67
							7	–	–	–	–	–
283.08831					0.050	6.27E-02	13	–	–	–	–	–
300.09522			0.016	2.48E-01			14	–	–	–	–	–
306.07663					0.052	2.31E-02	C10H17N3O6S		[M-H]-	Glutathione	306.0765326	-0.30
							C18H13NO4	2	[M-H]-	Aristolodione Hallacridone	306.0771826	1.82

							4	–	–	–	–	–
314.06235	0.016	4.11E-30					21	–				
316.05974					0.040	1.48E-02	25	–				
318.06110					0.032	1.94E-01	C11H15N5O4	4	[M+K-2H]-	1-Methyladenosine	318.0610116	-0.28
										2'-O-Methyladenosine		
										3'-O-Methyladenosine		
										N6-Methyladenosine		
							C10H13N5O5	2	[M+Cl]-	8-Hydroxy-deoxyguanosine	318.0610716	-0.09
										Guanosine		
							21	–	–	–	–	–
318.06835			0.028	1.65E+00			27	–	–	–	–	–
328.05840					0.031	1.56E-03	C10H17N3O6S	1	[M+Na-2H]-	Glutathione	328.0584776	0.23
							C18H13NO4	2	[M+Na-2H]-	Aristolodione	328.0591276	2.22
										Hallacridone		
							21	–	–	–	–	–
340.04348					0.032	7.43E-03	67	–	–	–	–	–
350.07186					0.028	8.16E-02	38	–	–	–	–	–
380.05244					0.045	7.24E-02	119	–	–	–	–	–
380.06151					0.031	7.33E-02	59	–	–	–	–	–
364.08760					0.036	1.72E-02	61	–	–	–	–	–
365.09097					0.032	7.34E-02	46	–	–	–	–	–
432.07487					0.029	4.86E-02	83	–	–	–	–	–
446.09094					0.038	2.34E-02	133	–	–	–	–	–
480.05202			0.019	8.78E-02			–	–	–	–	–	–
571.28875	0.031	5.67E-30					–	–	–	–	–	–
585.30442	0.025	6.87E-30					–	–	–	–		
597.30458	0.049	8.22E-30					–	–	–	–	–	–
598.30780	0.048	4.54E-30					–	–	–	–		

599.32003	0.029	5.10E-30					–	–	–	–	–	–
600.32336	0.027	2.83E-30					–	–	–	–	–	–
610.09695					0.031	2.69E-02	–	–	–	–	–	–
611.14473					0.053	7.78E-01	C20H32N6O12S2	1	[M-H]-	Oxidized glutathione	611.1446916	-0.06
612.14803			0.021	1.55E+00			–	–	–	–	–	–

Table 4-7 Table showing putative annotation of spectral peaks from positive ion analysis of polar *Daphnia magna* embryo extracts that showed significant changes in intensity between subitaneous and diapause-destined embryos (embryo type), time (h) post-ovulation or the interaction thereof based on ANOVA simultaneous component analysis (ASCA) models with a leverage threshold of 0.9 and squared prediction error alpha level of 0.05. Observed peak *m/z* is shown in addition to ASCA model statistics (Leverage and SPE). Formula indicates putative empirical formulae assigned to each peak using BEAMS, with corresponding annotation including number of possible species attributable to the aforementioned formula, respective putative compound identity with theoretical mass. Mass error (ppm) is the difference between theoretical and experimental *m/z*, normalised to experimental *m/z.* Other putative formula without compound ID assignment are indicated, for peaks with a high number of possible formulae only number of formulae are indicated for brevity.

	Embryo type		Time (h)		Time *x* Time (h)							
m/z	**Leverage**	**SPE**	**Leverage**	**SPE**	**Leverage**	**SPE**	**Formula**	**Number of possible species**	**Ion form**	**Compound (or compound class) identity**	**Theoretical mass**	**Mass error (ppm)**
117.05758			0.034	6.15E-02	0.047	6.21E-30	C4H9N2P	1	[M+H]+	-	117.0576124	0.29
121.08436	0.019	4.31E-30	0.024	6.97E-01			-	-	-	-	-	-
146.06024			0.077	1.62E-01			C9H7NO	1	[M+H]+	1H-Indole-3-carboxaldehyde	146.0600404	-1.35
167.08992			0.026	2.10E-01			C5H12N4O	1	[M+Na]+	-	167.0903324	2.49
183.12431	0.024	6.90E-31					-	-	-	-	-	-
192.05568	0.024	2.93E-30			0.095	7.48E-30	-	-	-	-	-	-
206.10091			0.054	1.18E-01	0.033	9.33E-30	-	-	-	-	-	-
209.09250			0.030	3.79E-01			C10H12N2O3	3	[M+H]+	4-Aminobenzoyl-(beta)-alanine	209.0920694	-2.08
									[M+H]+	Formyl-5-hydroxykynurenamine	209.0920694	-2.08
									[M+H]+	L-Kynurenine	209.0920694	-2.08
							C7H15N4P	1	[M+Na]+	-	209.0926554	0.72
							C3H12N8OS	1	[M+H]+	-	209.0927554	1.20
210.05301			0.034	1.21E-02			C11H9NO2	1	[M+Na]+	Indoleacrylic acid	210.0525504	-2.21
							C6H12NO5P	1	[M+H]+	-	210.0525884	-2.03
							C8H13NO3	1	[M+K]+	-	210.0527034	-1.48
							C4H9N7S	1	[M+Na]+	-	210.0532364	1.05
							C7H16NPS2	1	[M+H]+	-	210.0534574	2.11
243.10372	0.020	1.00E-29					-	-	-	-	-	-
249.06161			0.033	2.76E-01			-	-	-	-	-	-
258.07434	0.033	2.02E-29					-	-	-	-	-	-
276.15594	0.031	5.87E-30	0.019	3.68E-01			C11H21N3O5	2	[M+H]+	Epsilon-(gamma-Glutamyl)-lysine	276.1553984	-1.95
									[M+H]+	L-a-glutamyl-L-Lysine	276.1553984	-1.95
							C8H24N5O2P	1	[M+Na]+	-	276.1559844	0.18
							C12H25N3S2	1	[M+H]+	-	276.1562674	1.20
370.19340					0.293	7.52E-30	9	-	-	-	-	-
411.09854	0.019	5.87E-30			0.070	1.38E-29	53	-	-	-	-	-

429.13765			0.016	5.73E-01			-	-	-	-	-	-
447.35853			0.020	1.75E-02			-	-	-	-	-	-
449.06671	0.033	1.47E-29			0.082	1.38E-29	-	-	-	-	-	-
463.21952	0.031	2.02E-29					-	-	-	-	-	-
486.36895	0.021	8.63E-31	0.046	3.29E-01	0.049	4.57E-30	-	-	-	-	-	-
487.99537	0.021	1.73E-30					-	-	-	-	-	-

4.6.1. Discussion of the non-polar metabolite differences between diapause-destined and subitaneous *Daphnia magna* embryos

Considering the annotation data of the lipidomic data, the only peaks that changed significantly in intensity were associated with embryo type. Ten were successfully assigned at least 1 putative ID and molecular formula. Of these, 9 were identified as triacylglycerides (TGs). TGs consist of glycerol molecule esterified to three fatty acid chains that may vary in length and composition (Arrese and Soulages, 2010). TGs are the major components of energy reserves in cladocerans and copepods, and are present in high abundance in these species, contained within lipid storage droplets throughout their bodies (Tessier *et al.*, 2007). Cyclic changes in *D. magna* adult female TG content have been identified throughout the reproductive cycle, with evidence suggesting a large portion of maternal TG content is transferred to embryos during oogenesis (Jordão *et al.*, 2015). Therefore, their detection in *D. magna* embryos is expected. However, all of the identified putative TGs were significantly higher intensity in diapause-destined embryos versus subitaneous embryos (Figure 4-6). The accumulation of lipid-based energy reserves prior to diapause is regularly documented and believed to arise so that non-feeding animals can meet the energetic demands of sustained dormancy (Hahn and Denlinger, 2007). For example *D. melanogaster* accumulate TGs prior to the onset of adult diapause and the relative amounts of TGs accumulated also correlated with cold-hardiness (Ohtsu, Kimura and Hori, 1992). However, the relative intensity of these putative TGs showed no appreciable change over the course of the 48 h of development monitored here, suggesting their levels are of maternal origin. It is also unclear whether TGs may be used as an energy source over the diapause period, e.g., TGs represent 90% of

lipids in fat body of western corn borer *Diatraea grandiosella* that decreases by 75% over 150 days in diapause (Kuthiala and Chippendale, 1989). Alternatively, they may represent a stored energy source for post-embryonic development. Indeed, juvenile levels of TGs decrease during *D. magna* somatic growth (Bychek and Gushchina, 1999). Higher lipid reserves upon resumption of development would theoretically increase the likelihood of successful colonisation and survival upon starvation. For example, post-diapause offspring of the rotifer *Brachionus calyciflorus* exhibit much higher concentrations of neutral lipid storage droplets, which the authors suggest contribute to post-diapause offspring displaying a higher starvation tolerance and life-time fecundity in comparison with subitaneous offspring (Gilbert, 2004). However, contrasting evidence suggests that ex-ephippio *D. magna* are less resistant to starvation, which the authors attribute to higher somatic growth rate and smaller initial size (Arbačiauskas and Lampert, 2003). However, in their study, post-diapause offspring were hatched from ephippia from sediment produced under unknown conditions, which imposes a significant limitation on the interpretation of their findings (Arbačiauskas and Lampert, 2003).

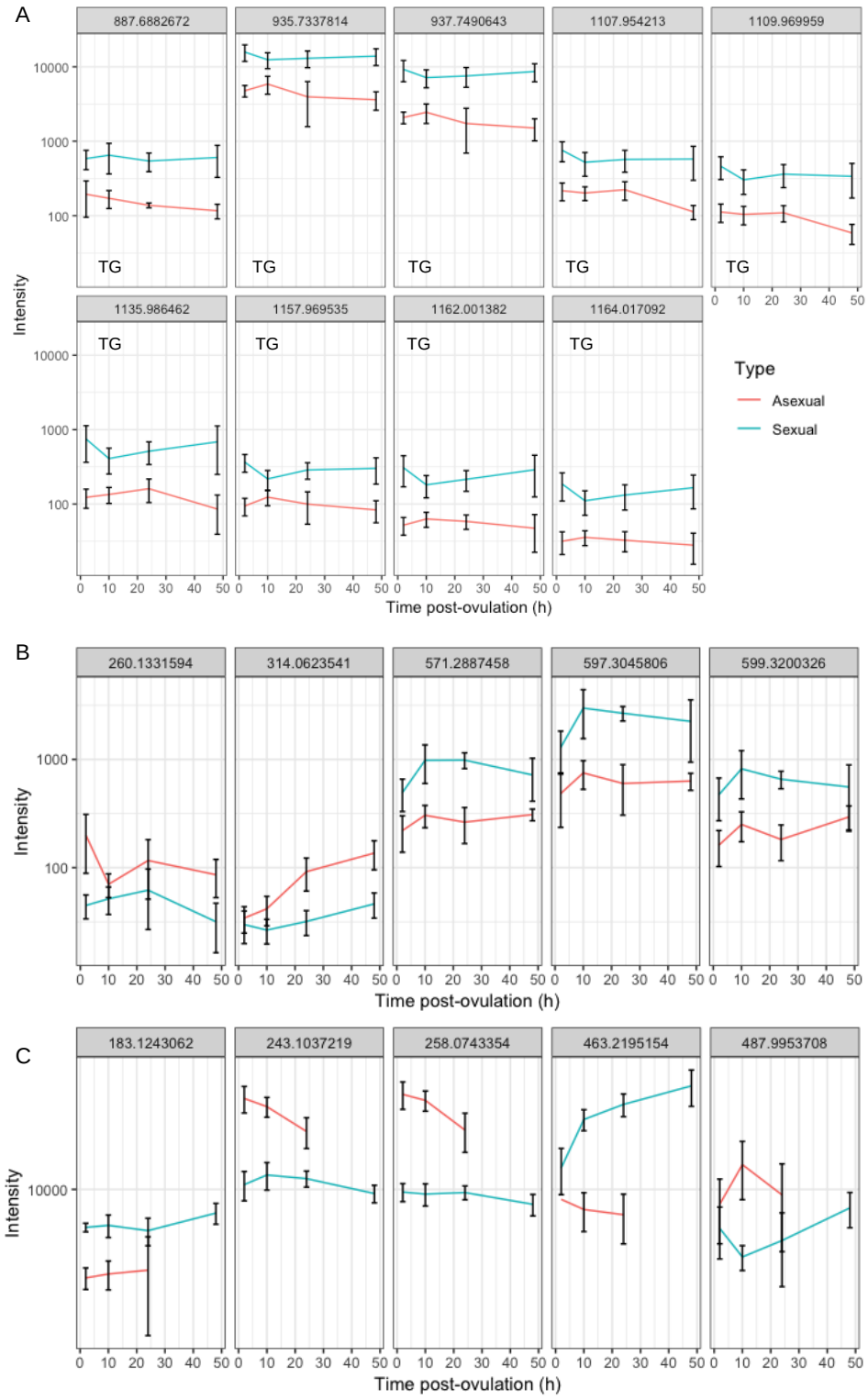


Figure 4-6 Mean intensity of peaks from A) negative ion metabolomics analysis of non-polar extracts, B) negative ion, C) positive ion metabolomics analysis of polar extracts of *Daphnia magna* embryos at different time (h) post-ovulation with significant differences in intensity between sexually produced diapause-destined (blue) and asexually produced subitaneous (red) embryos in ANOVA simultaneous component analysis (ASCA) with a leverage threshold of 0.9 and squared prediction error alpha level of 0.05. Error bars denote standard deviation. Putative identities of peaks were assigned are annotated (TG = triacylglycerides).

In the absence of further experimental data, I am unable to accurately assign a single TG compound ID to any of the TGs identified. Therefore, they may represent any one of hundreds of possible compounds, all with the same molecular formula but different arrangements of fatty acid chains and saturation, which may be instrumental in establishing a direct role for their higher levels in diapausing *D. magna* embryos. Indeed, limited knowledge of *Daphnia sp.* egg composition suggests that diapausing embryos are enriched (3.5-fold) with fatty acids (particularly C:18) and report pronounced differences in the composition of poly-unsaturated fatty acids between the two embryo types (Abrusan, Fink and Lampert, 2007). Diapausing adults of *Dolycoris baccarum* and *Piezodorus lituratus* also show higher proportions of unsaturated versus saturated lipids (Bashan and Cakmak, 2005). The authors postulate this may be an adaption to maintain energy resource fluidity at low winter temperatures (Bashan and Cakmak, 2005). There are also several other peaks identified as being significant between the two types of embryos from the polar datasets, also presented in Figure 4-6. However, I am unable to assign any putative identities to these metabolites and their nature and identity much like the TGs warrants further study.

4.6.2. Discussion of the polar metabolite differences between diapause-destined and subitaneous *Daphnia magna* embryos

Investigation of the features identified as significantly changing in intensity over time was also of limited success with only a handful of positive polar peaks being assigned a putative identity. However, the annotation of negative polar data was more successful (Table 4-6) and revealed three peaks putatively identified as

glutathione (GSH) and oxidised glutathione (GSSG) that showed a significant relationship between embryo type and time on peak intensity. The intensities of these peaks are plotted in Figure 4-7. The presence of multiple negative GSH species ([M-H]$^-$, [M+Na-2H]$^-$) as well as the oxidised form ([M-H]$^-$) all with a low parts per million (ppm) error improves confidence in the annotation over other possible metabolites. However they must still be considered as "putative". For convenience, these peaks herein are discussed in relation to their identities as GSH and GSSG.

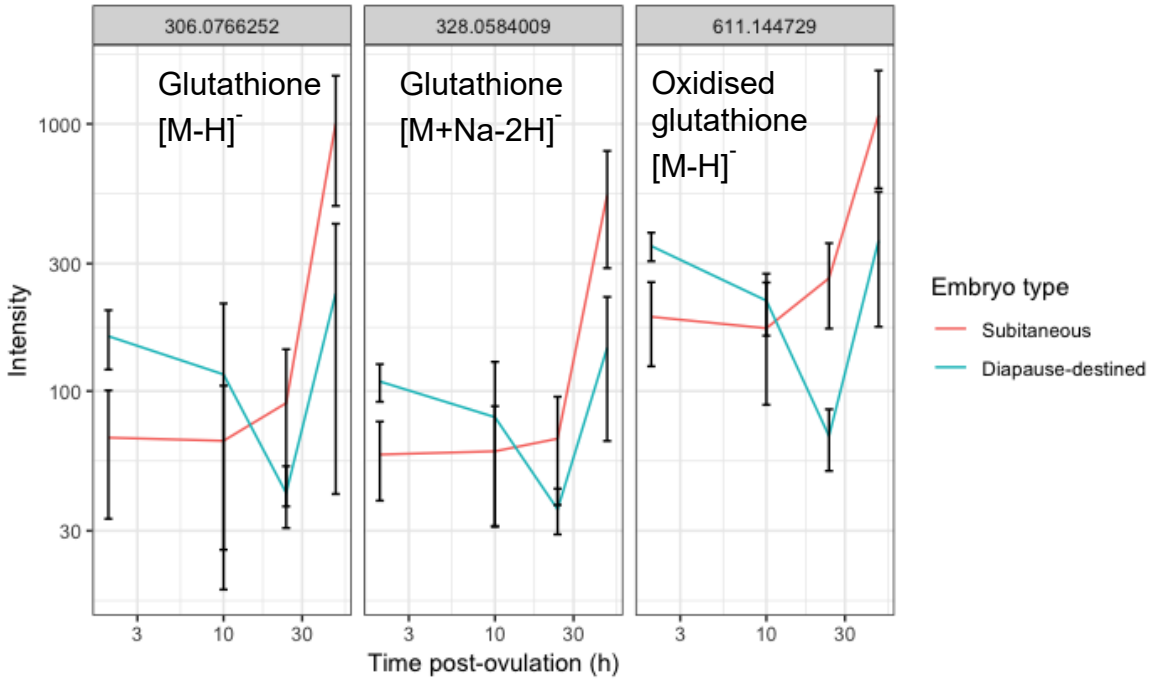


Figure 4-7 Mean intensity of putatively annotated glutathione metabolites analysed by negative ion nESI-DIMS metabolomics analysis in *Daphnia magna* subitaneous and diapause-destined embryos between 2 h and 48 h post-ovulation. Error bars denote standard deviation. Panel annotations indicate putatively assigned compound identities with corresponding ion form detected.

The relative intensity levels of GSH and GSSG showed variation across the time points, measured for both subitaneous and diapause-destined embryos. Initially at 2 h post-ovulation, diapause-destined embryos appear to have a higher content of GSH and GSSG than subitaneous embryos. The levels of GSH and GSSG appear to be relatively constant in subitaneous embryos between 2 h and 10 h, and show a

slight increase in mean intensity in 24 h samples followed by a further increase at 48 h. By contrast, after initially being higher in sexual embryos at 2 h, the intensity levels of GSH and GSSG appear to fall, reaching a minima at 24 h before rising again by 48 h. However, these remain lower in intensity that in subitaneous embryos. Following the observation of varying intensity of GSH and GSSG over diapausing and subitaneous embryonic development, differentially expressed genes were manually searched within the OrthoDB database (Section 2.13.4) based on their annotations to gene families associated with glutathione. This search included enzymes involved in its synthesis (glutamate cysteine ligase and GSH synthetase (Shelly C. Lu, 2014)), recovery of GSH from GSSG (glutathione reductase (Becker *et al.*, 2011)), and the enzymatic conjugation of GSH with lipophilic and xenobiotic substrates (glutathione-S-transferases and glutathione peroxidases (LeBlanc and Cochrane, 1987; Jovanović-Galović *et al.*, 2004; Jemec *et al.*, 2007)). A full list of these genes, including corresponding OrthoDB gene families is reported in Section 9.6. No significant differences in GSH-related genes were identified as significantly changing between subitaneous and diapause-destined embryos. However, 16 genes displayed differential expression over the time points investigated, and a further 10 demonstrated a significant interaction between embryo type and time post-ovulation on their expression (Section 9.6.). These genes were chiefly annotated as either glutathione-S-transferases and glutathione peroxidase of which, there are multiple isoforms in *D. magna* (LeBlanc and Cochrane, 1987). To identify if any of these genes displayed expression levels related to GSH and GSSG metabolite levels, Pearson correlation coefficients were calculated between the mean log-normalised gene counts and mean GSH metabolite intensities. A heatmap of genes with

expression levels that displayed a high positive or negative correlation (-0.9 > *r* > 0.9) are shown in Figure 4-8.



Figure 4-8 Heat map and hierarchical clustering of genes relating to glutathione that displayed differential expression (p-adjusted <0.01) in samples of subitaneous and diapause-destined *Daphnia magna* embryos collected between 2 h and 48 h post-ovulation and a high positive or negative correlation (-0.9 > *r* > 0.9) with glutathione metabolite levels. Row labels indicate OrthoDB-assigned gene name and coloured column top panels indicate embryo type and time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale.

Of the 26 genes, 9 were correlated with GSH levels identified in metabolomics analysis. The majority (6) of these genes had expression levels that were positively correlated with GSH levels and were all glutathione-*S*-transferases. These predominantly displayed an increasing expression over development in both embryos. Three genes (glutathione-*S*-transferase (Mu class) and two glutathione peroxidases) were negatively correlated with GSH levels. Of these, one peroxidase displays very high expression at 2 h post-ovulation, which subsequently drops. The

second shows a very similar profile to that of GSH with expression at 10 h lower than that at 2 h, then increasing at 48 h post ovulation and to higher levels in diapausing embryos than subitaneous embryos.

The tripeptide GSH is a ubiquitous antioxidant molecule, existing in reduced and oxidised (GSSG) form that acts in combination with series of enzymes to buffer cellular redox potential, offering protection from endogenous and exogenous sources of oxidative stress (e.g. reactive oxygen species) (Noctor and Foyer, 1998; Hansen and Harris, 2015). Fluctuations in GSH and GSSG have previously been identified during embryonic development, where it was postulated to play a role in protecting embryos from xenobiotic toxicity (Hansen and Harris, 2015). However, embryonic investigation of GSH levels is largely hampered in mammalian model species by practical issues accessing embryonic tissues. More recently, a compelling role of GSH in cell signalling and as a mediator of cellular proliferation was suggested (Hansen and Harris, 2015). Oscillations in the redox potential of cells with movement through the cell cycle has also been identified (Menon and Goswami, 2007).

Work conducted in mouse embryo fibroblast cell lines has documented dramatic changes in glutathione during the cell cycle, leading authors to suggest that reduced levels of GSH and consequent changes to cellular redox potential inhibit mitosis (Li and Oberley, 1998). Furthermore, this appears to be conserved, as studies in *Arabidopsis thaliana* detail recruitment of GSH to the nuclear envelope during cell division accompanied by dramatic changes to cytoplasmic levels of GSH (Vivancos *et al.*, 2010). Exactly how GSH regulates cell cycle progression has yet to be elucidated. However, many cell cycle regulatory proteins have putatively identified redox-sensitive motifs (Menon and Goswami, 2007).

The lower levels of GSH I observe in diapause-destined embryos collected 24 h post-ovulation coincide with the onset in a reduction of the rate of cell division in sexual embryos identified in Section 2.11. Therefore, the postulated role of low levels of GSH in the inhibition of the mitotic cell cycle suggests that the induction of developmental arrest of *D. magna* may indeed be modulated by changes in cellular redox potential as a consequence of reduced levels of GSH, particularly as concomitant decreases in subitaneous embryos are not observed. Indeed, GSH has previously been suggested to play a role in the induction and initiation of diapause in *B. mori* embryos (Meng *et al.*, 2011; Zhao, Hou and Sima, 2014). The authors found that at the induction stage, diapausing embryos had lower levels of GSH compared with non-diapausing embryos, whereas during diapause initiation this level had risen (Zhao, Hou and Sima, 2014). Topical application of hydrochloric acid that in *B. mori* prevents the initiation of diapause, did not change GSH levels but did lead to increased levels of GSSG leading the authors to suggest that pro-oxidant conditions drive diapause induction and reducing conditions drive diapause initiation (Meng *et al.*, 2011).

However, the role of glutathione in buffering the effects of oxidative stress cannot be discounted, particularly as evidenced by the time-dependent expression of glutathione-S-transferases and glutathione peroxidases identified as differentially expressed during embryonic development (Figure 4-8). These enzymes are predominantly involved in the role of GSH as an antioxidant, responsible for the enzymatic conjugation of glutathione with electrophilic reactive species including hydrogen peroxide (Becker *et al.*, 2011). Increased glutathione-S-transferase activity has previously been identified in the diapausing pupae of *Antheraea mylitta* however,

the authors postulate that this is due to the higher feeding rates of these pupae on phenolic compounds (Jena *et al.*, 2013). As *D. magna* embryos do not feed during embryogenesis, the same assumptions do not apply as to why the expression of glutathione-S-transferase differed between subitaneous and diapausing embryos at 48 h. By contrast, the diapausing larvae of *Ostrinia nubilalis* were found to have lower activity levels of glutathione peroxidase and glutathione-S-transferase, which is contrary to our findings of a higher level of expression of some of these genes, although it is important to consider that increase gene expression may not directly correlate with enzyme activity (Jovanović-Galović *et al.*, 2004). It may be that these enzymes are stock-piled during the pre-diapause preparatory phase in order to mediate an effective response to oxidative stress during the diapause period. During the adult diapause of *Culex pipiens* increased expression of the antioxidant enzymes *catalase* and *superoxide dismutase-2* has been observed (Sim and Denlinger, 2011). RNAi-mediated suppression of these two enzymes lead to reduced adult survival suggesting a role in promoting diapause-associated lifespan extension. However, the applicability to embryonic diapause has yet to be investigated but is also seen in dauer *C. elegans* larvae (Houthoofd *et al.*, 2002; Sim and Denlinger, 2011), suggesting that the ability to mount an effective response to oxidative stress is highly relevant during the period of diapause.

4.7. Conclusions

In the first section of this chapter, I successfully identified the point of developmental arrest and thus entry into diapause in *D. magna* based on the counting of DAPI-stained nuclei. Sexually produced embryos enter diapause at ~3,500 cells, which

occurs before 48 h post-ovulation when embryos are maintained at 20 °C. Without the corresponding whole-mount staging of diapausing *D. magna,* I am unable to confirm whether this is during the gastrula stage, as previously reported for *D. pulex* (Baldass 1941). However, I anticipate that the ontogenic stage of diapause arrest will be conserved between the two species. To the best of my knowledge, this represents the first such experimental confirmation of when *D. magna* enter diapause, thus contributing significantly to the characterization of this model species. Furthermore, whilst they are almost exclusively referred to in the literature as 'resting eggs', my data suggest that they are more correctly described as diapausing embryos. Diapause-destined and subitaneous embryos also display unique profiles of cell division. Notably sexually produced embryos are slower to divide and display a decrease in the rate of cell division after approximately 30 h. I suggest this represents the progression of cells towards arrest. However, I am unable to confirm whether this is by means of an extension to the length of cell division or, more likely, the unsynchronized arrest of cell division. Results reported in this section also forms the basis for later molecular investigation of diapause in later chapters, clearly defining the stages of pre-diapause development, and development arrest critical to the correct interpretation of results.

Previous comparisons between diapausing and non-diapausing offspring have thus far largely overlooked the effect of the different environmental effects which are utilised to induce diapause (Kankare *et al.*, 2016). My experimental design allowed me to compare the metabolic and transcriptomic profiles of subitaneous and diapause development in *D. magna* produced under the same environmental conditions. However, I avoided direct pairwise comparisons between equivalent time

points post-ovulation as the two embryo types were at different developmental stages. Furthermore, I identified unique transcripts and metabolites expressed only in one type of embryo. I was unable to adequately investigate these compounds in metabolomics data. However, transcripts were able to be included in my analysis. An important limitation of the dataset described, is that I can only draw conclusions based on shared metabolites. In order to partially address this current limitation, I later perform a separate experiment in Chapter 6 containing sexual samples with the intention of retaining and describing the nature of some of these unique peaks. But as there was previously limited knowledge regarding the molecular composition of diapause-destined embryos with only a description of a portion of fatty acids (Abrusan, Fink and Lampert, 2007), this chapter represents an important finding in that there is also disparity in the polar metabolome between these two embryo types. I identified notable dynamic changes in both gene expression and metabolite levels throughout development in both subitaneous and diapause-destined embryos. However, variation was dominated by differences between the two embryo types. Furthermore, initial investigative PCA revealed limited changes in the transcriptome 48 h after ovulation, represented by clustering of these points together. This coincides with the onset of developmental arrest, suggesting there are limited changes to gene expression after the onset of diapause. A similar pattern of separation was observed in metabolite data. However, this occurred ~24 h later, with the clustering of samples pre and post 74 h together, leading me to hypothesise that metabolic changes associated with developmental arrest take longer to manifest. Further investigation considering only the time points shared between subitaneous and diapause-destined embryos revealed that the lipidome is largely unperturbed

during development. However, I identified notable differences in the intensity of 9 peaks putatively annotated as TGs that were consistently enriched in diapausing embryos in comparison with subitaneous embryos, which may represent a potential energy source during diapause maintenance, or for post-embryonic development. My analysis also revealed a significant down-regulation of genes involved in translation in sexual embryos, which I anticipate may largely reflect a higher rate of protein synthesis occurring during the more rapidly dividing asexual embryos. However, the identification of the differential expression of *eIF3* in diapausing embryos in combination with the association of *eIF3* in various other capacities beyond the initiation of transcription, (e.g., cancer, developmental abnormalities, translation regulation and life-span extension) suggests that the role of *eIF3* in diapause development in *D. magna* warrants further investigation (Choudhuri, Maitra and Evans, 2013; Lee, Kranzusch and Cate, 2015; Cattie *et al.*, 2016).

Finally, the fluctuating levels of three peaks putatively identified as glutathione (GSH) and oxidised glutathione (GSSG) that showed significant relationships with embryo type and time in addition to the several glutathione-S-transferases and glutathione peroxidase genes that showed differential expression with time post-ovulation and between embryo type indicate that significant changes to cellular redox potential may occur throughout early diapause development. This warrants further investigation as to whether glutathione plays a role in cell division and regulation of developmental arrest or in embryonic protection from oxidative stress (Meng *et al.*, 2011; Zhao, Hou and Sima, 2014).

5. Molecular consequences of diapause initiation in *D. magna*

5.1. Introduction

The molecular events surrounding the initiation of diapause are of particular interest to my thesis, largely because it is still unknown how the reversible, coordinated arrest of development and other physiological processes is achieved. Moreover, it is becoming increasingly apparent that the molecular mechanisms behind the diapause phenotype may not always be evolutionarily conserved and will require elucidation at a species-specific level (Ragland, Denlinger and Hahn, 2010; Poelchau *et al.*, 2013b; Amsalem *et al.*, 2015; Koštál *et al.*, 2017).

In this chapter, I characterise the transcriptomic and metabolic changes that occur specifically at the onset of developmental arrest in *D. magna*. Very few researchers have previously investigated the molecular changes involved in the initiation of diapause, with reference to diapause development only. Most focus instead on comparisons with non-diapause stages (Reynolds and Hand, 2009; Lu and Xu, 2010; Bao and Xu, 2011; Pavlides, Pavlides and Tammariello, 2011; Colinet *et al.*, 2012; Ren *et al.*, 2016; Ziv *et al.*, 2017); others have focused on the maintenance and termination of diapause and excluded pre-diapause stages (Dong *et al.*, 2014; Kim *et al.*, 2015; Roncalli *et al.*, 2018). Of the two studies I identified that did conduct a comparable analysis, covering a transcriptomic and metabolomics study, both report a gradual progression of molecular changes towards diapause (Lu, Zhang and Xu, 2014; Koštál *et al.*, 2017).

Based on the data obtained in Chapter 4, I identified that the cell number of sexually produced embryos reaches a constant number of ~3,500 cells by 48 h post-ovulation

(Figure 4-1), and thus enters diapause (Koštál, 2006). Principal component analysis revealed that the variation within both transcriptomic and metabolomic dataset is dominated by differences between diapausing and non-diapausing embryos (Figure 4-2). However, there is a distinct temporal separation of sexually produced embryos before and after 48 h in transcriptomic datasets and 74 h in metabolomics polar datasets, leading me to hypothesise that a series of arrest-specific molecular changes occur, that potentially take longer to manifest in the metabolome.

To investigate the specific molecular changes in both metabolites and gene expression in relation to the initiation of diapause, sexually produced embryos only were considered. I here compare the gene expression and metabolite levels of sexually produced embryos at 48 h post-ovulation with the earlier time point at 10 h post-ovulation, with the aim to identify features that change significantly between the directly developing (10 h) and diapause (48 h) phenotype. I also compare the next successive time point sampled during the diapause period at 74 h with the 48 h time point to ascertain whether changes associated with the initiation of diapause occur immediately or continue into the stages of developmental arrest.

Not only does this study provide valuable information regarding *Daphnia* diapause, which until now has not been investigated on a molecular level, it also contributes valuable knowledge to diapause research as a whole, given that very few studies have focused on alterations within the diapause developmental trajectory alone. Genes and metabolites that show altered levels in relation to the onset of diapause are likely candidates for the molecular mechanisms of developmental arrest, be it as targets of upstream regulatory processes or by their direct action.

5.2. Results and Discussion

To quantify the extent and magnitude of metabolic and transcriptional changes occurring in diapausing embryos in relation to the initiation of diapause, the same large omics datasets introduced in Section 4.3 were utilised. However, only sexually produced embryos were considered. My experiment aimed at investigating the metabolomics data included a separate quality control sample, pooled from only sexually produced embryos without subitaneous embryos. This “diapause only” dataset, reviewed fully in Section 9.3, was performed to aid in the retention of diapause-specific peaks that would otherwise be removed during the metabolomics data filtering processes (Section 2.14.1.4). This increased the size of the resultant peak matrices with an additional 11 peaks in the lipidomic dataset and, an additional 57 and 81 peaks were present in the final intensity matrices of positive and negative ion analyses of polar extracts respectively (Section 9.3.2). Furthermore, the peak composition of the final data matrices that were used in my investigations, from either both embryo types or sexually produced embryos only, are quite different with an excess of 18% of peaks unique to the sexually produced samples (Section 9.3.2). To ascertain whether molecular changes associated with the initiation of diapause occur immediately or continue into the stages of developmental arrest, the gene expression and metabolite levels of sexually produced embryos at 48 h post-ovulation (i.e. immediately after the initiation of diapause) with the earlier time point (during active cell division) at 10 h post-ovulation were compared (48 v 10). I also compare the next successive time point sampled during the diapause period at 74 h with the 48 h time point (74 v 48).

Transcriptomic analysis compared a total number of 16,790 genes expressed at 48 v 10, and separately, 15,630 genes were interrogated at 72 v 48 (see Section 2.13.3 for methods). The relative differences in gene transcription were calculated as the average $\log_2$ fold change across three biological replicates. Statistical significance was determined based on a two-tailed Student's t-test with Benjamini-Hochberg false discovery rate (FDR) correction (Benjamini and Hochberg, 1995) adjustment. The results of these analyses as gene $\log_2$ fold changes versus -$\log_{10}$ *p*-value (FDR-adjusted) are plotted in Figure 5-1(A). Metabolite fold changes were calculated based on mean peak intensity at 48 h versus 10 h and at 74 h versus 48 h post-ovulation over 2–8 biological replicates (Table 9-8). Statistical significance was determined based on a two-tailed Student's t-test with Benjamini-Hochberg false discovery rate (FDR) correction (Benjamini and Hochberg, 1995) adjustment. The results of these analyses as metabolite $\log_2$ fold changes versus -$\log_{10}$ *p*-value (FDR-adjusted) are plotted in Figure 5-1 (B–D). An additional time point within this interval was collected in the metabolomics data sets at 24 h post-ovulation. However, as a comparable time point was not collected for transcriptomic data, these data are not considered here and are instead reported in Section 10.

5.2.1. Interrogating the transcriptome

Large changes in gene expression are observed comparing directly developing and diapausing embryos (48 v 10). 2701 genes were significantly differentially expressed (Table 5-1) at 48 v 10, of these 2294 genes displayed log_2 fold changes that were greater than two-fold (i.e. a fold-change >4), representing 13% of the total genes detected. The numbers of up- and down-regulated genes were roughly equal with 1215 up-regulated genes and 1079 down-regulated. Comparing changes from early and later diapause developmental arrest (74 v 48), a much smaller fraction of genes showed altered expression with just 67 genes (0.3%) displaying a significant fold change.

Table 5-1 Table of number of genes and metabolites that displayed a significant (false discovery rate (FDR) adjusted *p*-value) changes relative to the initiation of diapause in *Daphnia magna* embryos. FC = fold change. Total indicates the number of features detected within the two time points compared. 'large' log_2 fold change indicates the number of features with a -2 > log_2 FC > 2.

	Time point compared (h)	Total	$p <$ 0.01	large log_2 FC	FC lower range	FC upper range	% of total	Up	Down
Transcriptomics	48 v 10	16790	2701	2294	-9.7	9.2	13.66	1215	1079
	74 v 48	15630	67	67	-4.2	4.2	0.43	16	51
Negative ion non-polar metabolomics	48 v 10	805	1	0	-0.9	3.4	0.00	0	0
	74 v 48	805	0	0	-2.7	3.0	0.00	0	0
Negative ion polar metabolomics	48 v 10	353	39	9	-1.8	3.3	2.55	9	0
	74 v 48	353	72	28	-2.5	3.7	7.93	27	1
Positive ion polar metabolomics	48 v 10	341	69	4	-2.4	2.9	1.17	4	0
	74 v 48	341	83	37	-1.3	4.6	10.85	37	0

5.2.2. Interrogating the metabolome

Investigating lipidomic (non-polar) data, there were no peaks with a significant, greater than two-fold log_2 fold change in intensity in the 48 v 10 comparison. Just 9 polar metabolite peaks (less than 3% of total detected (Table 5-1)) in negative ion analysis exhibited a significant fold change between the same timepoints. Of these, 3 were specific to the diapause embryo dataset. A similarly small number of significantly changing polar metabolite peaks were identified in positive ion analysis with 4 peaks, (less than 2% of total). None of these were unique to the sex specific dataset. Comparing changes immediately after the initiation of developmental arrest (74 v 48) there were also no significant changes in intensity of peaks detected in the non-polar extracts analysed. However, changes in metabolite levels were much more pronounced in polar extracts with 8% and 11% of peaks having a significant log_2 change greater than two-fold in the 74 v 48 comparison in negative and positive ion analysis respectively (Table 5-1). The majority of peaks demonstrated a positive fold change between the two time points with only one peak in the negative dataset demonstrating a reduction in intensity. Additionally, all 28 peaks from negative ion analysis and 70% (26 peaks out of 37 significantly changing) from the positive ion analysis were found only in the diapause dataset (Section 10.2).

5.2.3. Molecular signatures of pre- and post-onset diapausing embryos

These results suggest that there are profound molecular events associated with the initiation of diapause, but more specifically demonstrates that whilst gene expression changes are closely timed with the onset of developmental arrest (occurring between 10 h and 48 h), metabolic changes occur later (between 48 h and 74 h) and appear

to be limited to polar metabolites as the lipidome showed limited changes in peak intensity. Furthermore, whilst gene expression changes were roughly equal in direction, the majority of metabolites demonstrated positive fold changes meaning their intensity at 74 h was significantly higher than at 48 h. This temporal separation between the onset of developmental arrest and wide spread diapause-induced metabolic effects may largely reflect that alterations in gene expression may manifest quicker due to the comparatively shorter half-life of mRNA versus other cellular constituents such as proteins (Suarez and Moyes, 2012). It is also important to consider that mRNA abundance does not necessarily reflect protein abundance (Feder and Walser, 2005), which is further downstream of translation and subject to additional possible post-translational modifications in addition to numerous regulatory steps–all of which could conceivably take an additional 24 hours until the consequences of diapause initiation are apparent at the level of individual metabolite abundance (Feder and Walser, 2005; Hayward, 2014).

Very few researchers have previously investigated the specific molecular changes involved in the initiation of diapause with reference to diapause development only. My observation that the majority of significantly changing peaks in relation to the onset of developmental arrest were peaks unique to embryos in the diapause developmental trajectory suggests that a large portion of biologically relevant information has been lost in prior studies that have focused on diapause versus non-diapause comparisons. Of the two studies that did conduct a comparable analysis, both report a gradual progression towards diapause rather than the drastic arrest-specific changes I describe here (Lu, Zhang and Xu, 2014; Koštál *et al.*, 2017). By investigating the pupal diapause of *Helicoverpa armigera* using gas chromatography

mass spectrometry analysis, 55 metabolites were detected in pupal haemolymph and 52 in the brain, covering a subset of processes including carbohydrate, amino acid, polyol and fatty acid metabolism (Lu, Zhang and Xu, 2014). The authors describe a gradual increase in the levels of a subset of metabolites including methionine, leucine, and *myo*-inositol in pupal haemolymph over the 10 d sampled representing diapause initiation (4 time points; 1, 2, 4, and 10 d; Lu, Zhang and Xu, 2014). Microarray analysis of *Chymomyza costata* throughout the phases of larval diapause, also revealed gradual alterations to gene expression, reflecting a gradual progression towards diapause, although the authors report that, in this species, the phases of induction and initiation of diapause are in rapid succession and often indistinguishable (Koštál *et al.*, 2017). When comparing the last directly developing 'induction' phase sample, and the first sample taken during the diapause maintenance phase, a small fraction (0.7% of 1,042) of transcripts were differentially expressed (Koštál *et al.*, 2017).

The features that demonstrated significant changes in level in relation to the initiation of diapause were further investigated in the following sections to assign putative metabolite identifications and gene annotations, to investigate the potential mechanisms of diapause initiation. Features that show altered levels in relation to the onset of diapause are potential targets of upstream regulatory processes or may directly be involved in diapause-associated developmental arrest.

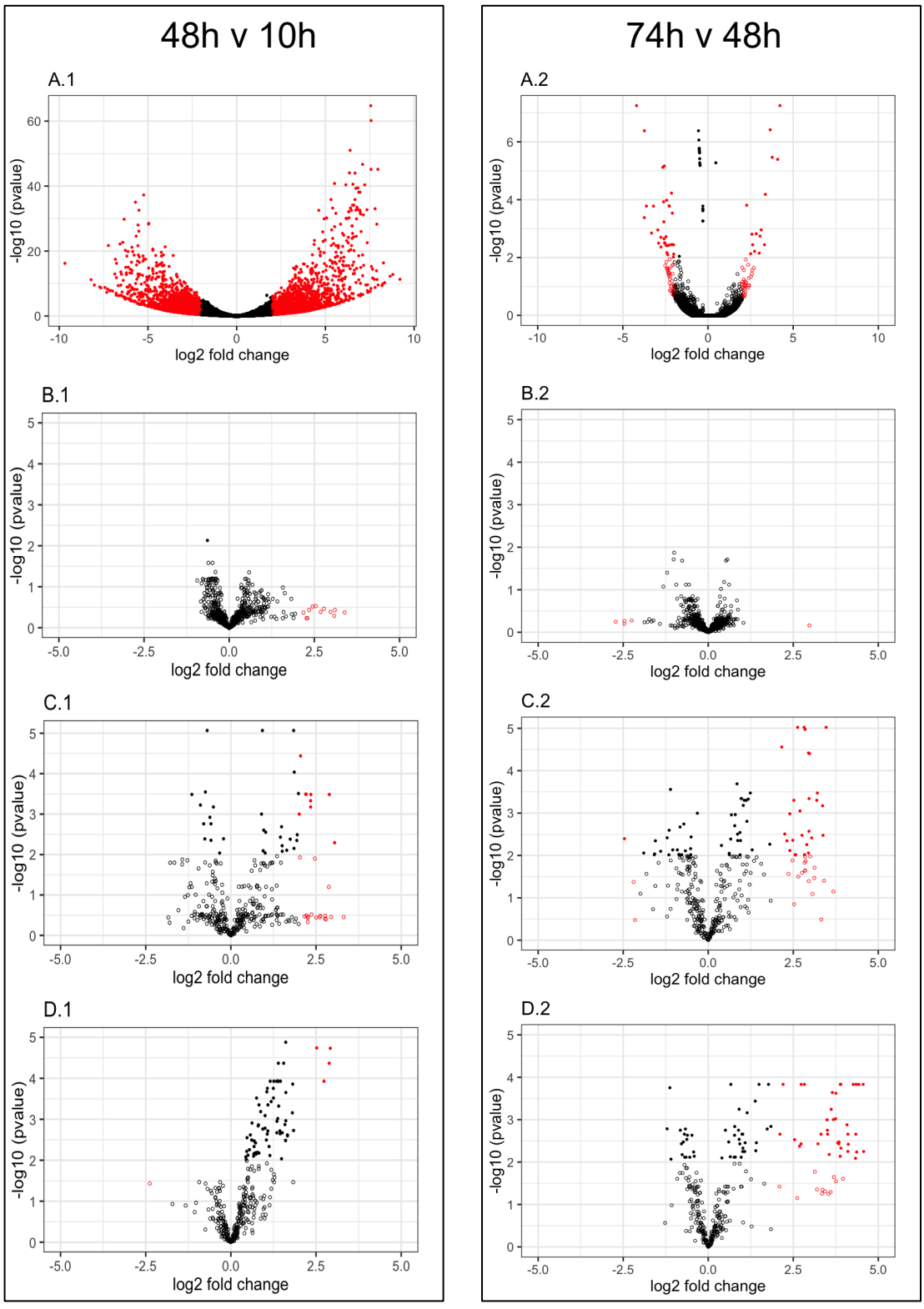


Figure 5-1 Volcano plot of differentially expressed genes (A) and metabolites (B-D) with significant intensity fold changes between pre-onset and onset of diapause (48 h v 10 h) and between onset of diapause and embryos in diapause (74 h v 48 h). Metabolomics analysis implemented nESI-DIMS analysis of non-polar extracts (B; negative ion mode) and polar extracts analysed in negative ion mode (C) and positive ion mode (D). Filled points represent $p < 0.01$ (after Benjamini-Hochberg false discovery rate correction), red points indicate $\log_2$ fold change ratio > 2.

5.2.4. Putative significantly changed metabolites

To investigate the nature of metabolite peaks changing significantly in relation to the onset of diapause developmental arrest, all peaks were assigned putative empirical formulae and identities as described previously (Section 2.14.1.5). Peak mass/charge (*m/z*) values and intensities from pre-glog matrices were collated for each analyses and input into BEAMS and matched against the LipidMAPS (non-polar extracts) and HMDB (polar extracts) databases using a maximum 2.5 ppm error cut off (Section 2.14.1.5) to assign putative empirical formulae and compound names. For non-polar extracts, a total of 300 out of 805 peaks were assigned at least one putative ID. Annotation of polar extracts was less successful, although this partially represents the smaller number of peaks in the initial intensity matrix with 115 out of 353 for negative ion analysis and 90 out of 341 peaks for positive ion analysis. This degree of metabolite annotation of DIMS data is consistent with other studies from the Viant laboratory over the last few years. Furthermore, numerous peaks, particularly those at higher *m/z* ranges, were assigned multiple formulae/compound IDs. Without further investigation (e.g. peak fragmentation), all putative IDs must be assumed equally possible.

There were no significantly changing peaks in the non-polar extracts in either of the time point comparisons and as only a handful of peaks showed significant large fold changes in the 48 v 10 polar analysis any annotations are presented in combination in Table 10-1. For brevity, only molecular formulae with a putative compound identity are shown and the number of putative molecular formulae without an assigned identity is indicated. Five of a total nine significantly changing peaks in negative ion mode and all four of the peaks in positive ion mode were assigned at least one

putative formula. However, upwards of five possible formulae were assigned to any given peak and identities included several nucleosides and modified nucleosides (Table 5-2).

Table 5-2 Table showing putatively annotated spectral peaks from the negative and positive ion metabolomics analysis of sexually produced *Daphnia magna* embryo polar extracts with signifcant ($p < 0.01$; after false discovery rate correction) large peak intensity log_2 fold change ratios (> 2) between 48 h and 10 h post-ovulation. Other putative formulae without compound identity assignment are indicated, for peaks with a high number of possible formulae only numbers of formulae are indicated for brevity.

Analysis	*m/z*	*p*-adjust	Fold change (log2)	Formula	Number of possible species	Ion form	Compound (or compound class) identity	Theoretical mass	Mass error (ppm)
Negative ion	**283.08214**	0.0003	2.2	C11H12O5	3	[M+Hac-H]-	Elenolide	283.082329	0.68
							Sinapic acid		
							cis-Sinapic acid		
				C13H16O7	3	[M-H]-	2-O-Benzoyl-D-glucose	283.082329	0.68
							D-Vacciniin		
							p-Cresol glucuronide		
				5	-	-	-	-	-
	283.08827	0.0003	2.89	C10H20N2O3S	2	[M+Cl]-	Methionyl-Valine	283.088867	2.10
							Valyl-Methionine		
				11	-	-	-	-	-
	284.08895	0.0003	2.35	C11H15N3O6	1	[M-H]-	N4-Acetylcytidine	284.0888106	-0.48
				5	-	-	-	-	-
	318.06106	0.001	2.01	C11H15N5O4	3	[M+K-2H]-	1-Methyladenosine	318.061012	-0.14
							2-O-Methyladenosine		
							N6-Methyladenosine		
				C10H13N5O5	2	[M+Cl]-	8-Hydroxy-deoxyguanosine	318.061072	0.05
							Guanosine		
				20	-	-	-	-	-
	380.05249	0.0003	2.21	60	-	-	-	-	-
Positive ion	134.02717	1.82E-05	2.52	C4H7NO2S	-	[M+H]+	-	134.0270274	-1.03
	284.09943	1.85E-05	2.92	C10H13N5O5	2	[M+H]+	8-Hydroxy-deoxyguanosine	284.098946	-1.70
							Guanosine		
				4	-	-	-	-	-
	306.08125	4.28E-05	2.89	23	-	-	-	-	-
	328.06332	0.0001	2.73	35	-	-	-	-	-

Table 5-3 Table showing putative annotation of spectral peaks from negative ion metabolomics analysis of sexually produced *Daphnia magna* embryo polar extracts with signifcant ($p < 0.01$; after false discovery rate correction) large peak intensity log_2 fold change ratios (> 2) between 74 h and 48 h post-ovulation. Other putative formulae without compound identity assignment are indicated, for peaks with a high number of possible formulae only numbers of formulae are indicated for brevity.

m/z	*p*-adjust	Fold change	Formula	Number of possible species	Ion form	Compound (or compound class) identity	Theoretical mass	Mass error (ppm)
240.08070	0.0005	2.96	4	–	–	–	–	–
344.11644	0.0027	2.96	21	–	–	–	–	–
404.13730	0.0038	3.04	38	–	–	–	–	–
415.14586	0.0000	2.98	47	–	–	–	–	–
423.09962	0.0096	2.83	86	–	–	–	–	–
425.11590	0.0044	2.49	–	–	–	–	–	–
456.17261	0.0056	2.90	51	–	–	–	–	–
457.19297	0.0007	3.36	54	–	–	–	–	–
471.07635	0.0033	2.76	157	–	–	–	–	–
472.20414	0.0088	2.95	50	–	–	–	–	–
473.59224	0.0003	3.21	–	–	–	–	–	–
474.09399	0.0005	3.19	115	–	–	–	–	–
477.03309	0.0097	2.55	336	–	–	–	–	–
484.58242	0.0000	2.95	–	–	–	–	–	–
499.23984	0.0096	2.57	56	–	–	–	–	–
552.12455	0.0005	2.52	–	–	–	–	–	–
559.03600	0.0009	2.70	–	–	–	–	–	–
563.11269	0.0077	2.40	–	–	–	–	–	–
567.14116	0.0000	3.47	–	–	–	–	–	–
568.14332	0.0000	2.85	–	–	–	–	–	–
569.09265	0.0010	2.40	–	–	–	–	–	–
578.13150	0.0033	3.38	–	–	–	–	–	–
579.05264	0.0000	2.16	–	–	–	–	–	–
583.11482	0.0031	2.25	–	–	–	–	–	–
587.02974	0.0040	-2.46	C15H24N2O17P2	2	[M+Na-2H]-	Uridine diphosphate glucose	587.029698	-0.08

						Uridine diphosphate galactose		
587.10643	0.0000	2.82	–	–	–	–	–	–
595.08441	0.0000	2.63	–	–	–	–	–	–
598.35236	0.0045	2.32	–	–	–	–	–	–

Table 5-4 Table showing putative annotation of spectral peaks from positive ion metabolomics analysis of sexually produced *Daphnia magna* embryo polar extracts with signifcant ($p < 0.01$; after false discovery rate correction) large peak intensity log_2 fold change ratios (> 2) between 74 h and 48 h post-ovulation. Other putative formulae without compound identity assignment are indicated, for peaks with a high number of possible formulae only numbers of formulae are indicated for brevity.

m/z	*p*-adjust	Fold change	Formula	Number of possible species	Ion form	Compound (or compound class) identity	Theoretical mass	Mass error (ppm)
157.53785	0.0030	2.54	–	–	–	–	–	–
182.03398	0.0056	4.11	3	–	–	–	–	–
183.05041	0.0042	2.68	2	–	–	–	–	–
191.54085	0.0047	3.90	–	–	–	–	–	–
192.05568	0.0022	4.11	1	C12H5N3	[M+H]+		192.0556234	-0.31
200.54606	0.0058	4.37	–	–	–	–	–	–
203.56370	0.0037	3.84	–	–	–	–	–	–
209.04987	0.0056	4.57	–	–	–	–	–	–
211.55254	0.0081	4.33	–	–	–	–	–	–
249.04617	0.0022	2.11	7	–	–	–	–	–
258.11060	0.0001	2.83	C13H17NO3	1	[M+Na]+	Pandamarilactam 3x	258.1100654	-2.07
			C8H20NO6P	1	[M+H]+	Glycerophosphocholine	258.1101034	-1.92
			C10H21NO4	1	[M+K]+	Miglustat	258.1102184	-1.47
			C3H17N9OP2	1	[M+H]+		258.1104084	-0.74
			C6H17N7OS	1	[M+Na]+		258.1107514	0.59
275.15157	0.0001	2.73	7	–	–	–	–	–
356.17774	0.0001	4.55	9	–	–	–	–	–
356.67941	0.0001	4.35	–	–	–	–	–	–
360.15050	0.0013	4.08	16	–	–	–	–	–
367.11010	0.0001	3.90	24	–	–	–	–	–
370.19338	0.0001	4.43	9	–	–	–	–	–
370.69502	0.0001	4.26	–	–	–	–	–	–
381.07998	0.0038	4.12	C12H22O11	26	[M+K]+	3-b-Galactopyranosyl glucose Allolactose Alpha-Lactose Cellobiose D-Maltose	381.079374	-1.60

						DEAE-cellulose		
						Epimelibiose		
						Fagopyritol A1		
						Galabiose		
						Galactinol		
						Glucinol		
						Glucose-1,3-mannose oligosaccharide		
						Inulobiose		
						Isomaltose		
						Kojibiose		
						Lactulose		
						Maltulose		
						Mannobiose		
						Melibiose		
						Neotrehalose		
						Sakebiose		
						Sucrose		
						Trehalose		
						Trehalulose		
						Turanose		
						beta-Lactose		
			39	–	–	–	–	–
382.08303	0.0035	3.81	37	–	–	–	–	–
384.12553	0.0066	3.56	25	–	–	–	–	–
387.08765	0.0001	3.88	40	–	–	–	–	–
404.17665	0.0022	4.34	22	–	–	–	–	–
411.09855	0.0034	3.84	53	–	–	–	–	–
423.06441	0.0010	3.49	70	–	–	–	–	–
433.08057	0.0022	3.51	108	–	–	–	–	–
434.09627	0.0002	3.75	46	–	–	–	–	–
449.06668	0.0073	3.88	89	–	–	–	–	–
450.06117	0.0006	3.62	79	–	–	–	–	–
450.07010	0.0037	2.73	75	–	–	–	–	–
450.08295	0.0010	3.68	65	–	–	–	–	–
455.06226	0.0022	3.32	156	–	–	–	–	–

471.03974	0.0002	3.65	126	–	–	–	–	–
489.32614	0.0001	2.20	17	–	–	–	–	–
495.05061	0.0037	3.23	148	–	–	–	–	–
495.06316	0.0010	3.75	138	–	–	–	–	–
569.06713	0.0018	3.50	–	–	–	–	–	–

Of the peaks identified in both polar analyses that changed significantly between 48 h and 74 h post-ovulation, the majority were assigned multiple empirical formulae. However, very few had plausible compound identities associated with the formula. Twenty-eight (28) peaks were identified as significantly changing in negative ion mode. Of these, 13 were assigned at least one formulae and only one with an accompanying putative compound identity of uridine-diphosphate (UDP)-glucose or UDP-galactose (Table 5-3). A similarly limited annotation was achieved within positive ion dataset; twenty-eight of the 37 significant peaks were assigned putative formulae. However, only 2 were also annotated with a compound identity. The peak with a mass/charge ratio (*m/z*) of 381.079983 was associated with a high number of putative molecular formulae (65; Table 5-4), 26 of which were assigned a putative identity. These were predominantly disaccharides. However, due to the isomeric nature of hexose sugars, they are typically difficult to resolve during mass spectral analysis without the utilisation of separative techniques or tandem mass spectrometry (Draper et al., 2013). Therefore, my discussion regarding the nature of this peak is acknowledged as being speculative. However, one of the possible identities of this peak is trehalose.

Trehalose is a disaccharide sugar formed from uridine diphosphate (UDP)-glucose and glucose-6-phosphate (Xu *et al.*, 2009; Guo *et al.*, 2015; Miao *et al.*, 2016). UDP-glucose was also identified in polar *Daphnia* extracts analysed in negative ion mode, with a significant $\log_2$ fold change at 74 v 48 (Table 5-3). The mean intensity of the putatively annotated trehalose and UDP-glucose peaks with time post ovulation in sexually produced embryos is shown in Figure 5-2 (A). Initially the intensity of trehalose is low in embryos between 2 h and 48 h post ovulation, coinciding with the

period of pre-diapause development (Section 4.2). However, the intensity of trehalose increases dramatically with a 4.12 log_2 fold increase between 48 h and 74 h (Table 5-4). Indeed, its intensity continues to increase (albeit minimally) throughout the time points that I sampled within diapause (Figure 5-2) wherein the intensity of this peak at 2 y (17592 h) post ovulation is almost 28-fold higher than the levels detected at 2 h (Section 10.3; Table 10-4). By contrast, the putative UDP-glucose peak displays decreasing intensity throughout the period of active embryo development (2 h–48 h) and reaches a minimum at 74 h and remains low throughout the period of diapause. However, the mean intensity of UDP-glucose and trehalose peaks were only weakly correlated and insignificant ($\rho = -0.6$, $p = 0.073$). This may be due to having not detected or (more likely) annotated the trehalose-6-phosphate intermediate. Furthermore, UDP-glucose also participates in many other metabolic pathways (e.g. glycolysis; (Wang, Li and Chen, 2016). Although, the identification of the $[M+Na-2H]^-$ sodium adduct only, likely means this is an erroneous annotation of UDP-glucose.

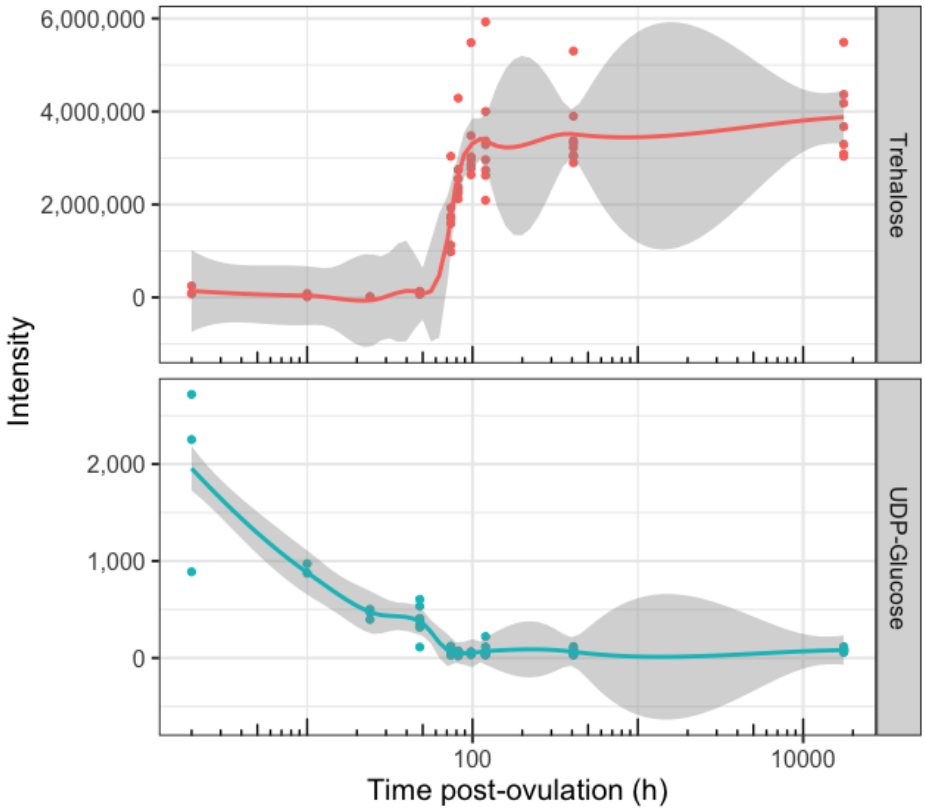


Figure 5-2 Intensity of putatively annotated trehalose (red, top panel) and uridine diphosphate (UDP)-glucose (blue, lower panel) peaks measured utilising nanoelectrospray ionisation-direct infusion mass spectrometry analysis of polar (positive and negative ion analysis respectively) extracts of *Daphnia magna* diapausing embryos. Points indicate individual biological replicates, lines fitted with locally weighted scatterplot smoothing (LOESS; span = 0.5), grey shading indicate model 95% confidence interval.

The predominant pathway for trehalose synthesis is shown in Figure 5-3 and involves trehalose-6-phosphate synthase (TPS), which catalyses the conversion of glucose-6-phosphate and UDP-glucose to trehalose-6-phoshate. This molecule is then dephosphorylated by trehalose-6-phosphate phosphatase (TPP) to yield trehalose (Guo *et al.*, 2015; Tang *et al.*, 2018). The enzyme called trehalase catalyses the hydrolysis of trehalose to glucose (Zhi-Hui *et al.*, 1994). However, an additional four trehalose synthesis pathways have also been described (Magalhães *et al.*, 2017).

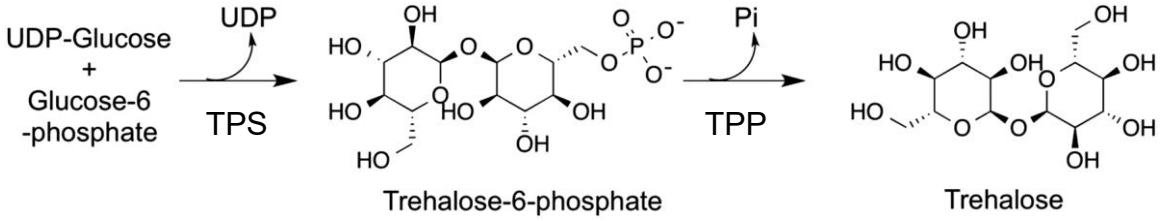


Figure 5-3 Biosynthesis of trehalose from Miao *et al*., (2016). Uridine diphosphate (UDP)-glucose and glucose-6-phosphate are condensed to form trehalose-6-phosphate catalysed by the enzyme trehalose-6-phosphate synthase (TPS). Trehalose is formed in a final enzymatic step catalysed by trehalose-6-phosphate phosphatase (TPP).

Manual inspection of the two lists of expressed genes in this study revealed two trehalase (Dapma7bEVm005872, Dapma7bEVm005873) and a total of 4 trehalose phosphatase genes (Dapma7bEVm004057, Dapma7bEVm009716, Dapma7bEVm015481, Dapma7bEVm029000), identified based on similarity with OrthoDB gene families (Section 2.13.4). The expression of these genes in sexually produced embryos across all time points post-ovulation is displayed in Figure 5-4. There was no detectable expression for one trehalose phosphatase gene in our RNA-Seq data. However, two genes (Dapma7bEVm009716, Dapma7bEVm015481), both putative trehalose phosphatases, did show patterns of expression consistent with trehalose metabolite intensity levels. Expression of these two genes is initially low in embryos immediately after ovulation, and an increase in gene expression is then observed between 10 h and 48 h. The expression of these genes was also investigated in subitaneous embryos and no concomitant increases in expression were observed (Section 10.3; Figure 10-2). Both genes displayed high fold changes in the 48 h v 10 h comparison (4.381 and 2.347 $\log_2$ fold change; Section 10.3). However, because of variation between technical replicates, Dapma7bEVm015481 was flagged as an outlier based on Cook's distance between biological replicates. Furthermore, there was no significant correlation between trehalose-related gene

expression and metabolite intensity, which is most likely due to the increase in gene expression between 10 h and 48 h being earlier than the increase in trehalose metabolite intensity between 48 h and 74 h (Table 10-5).

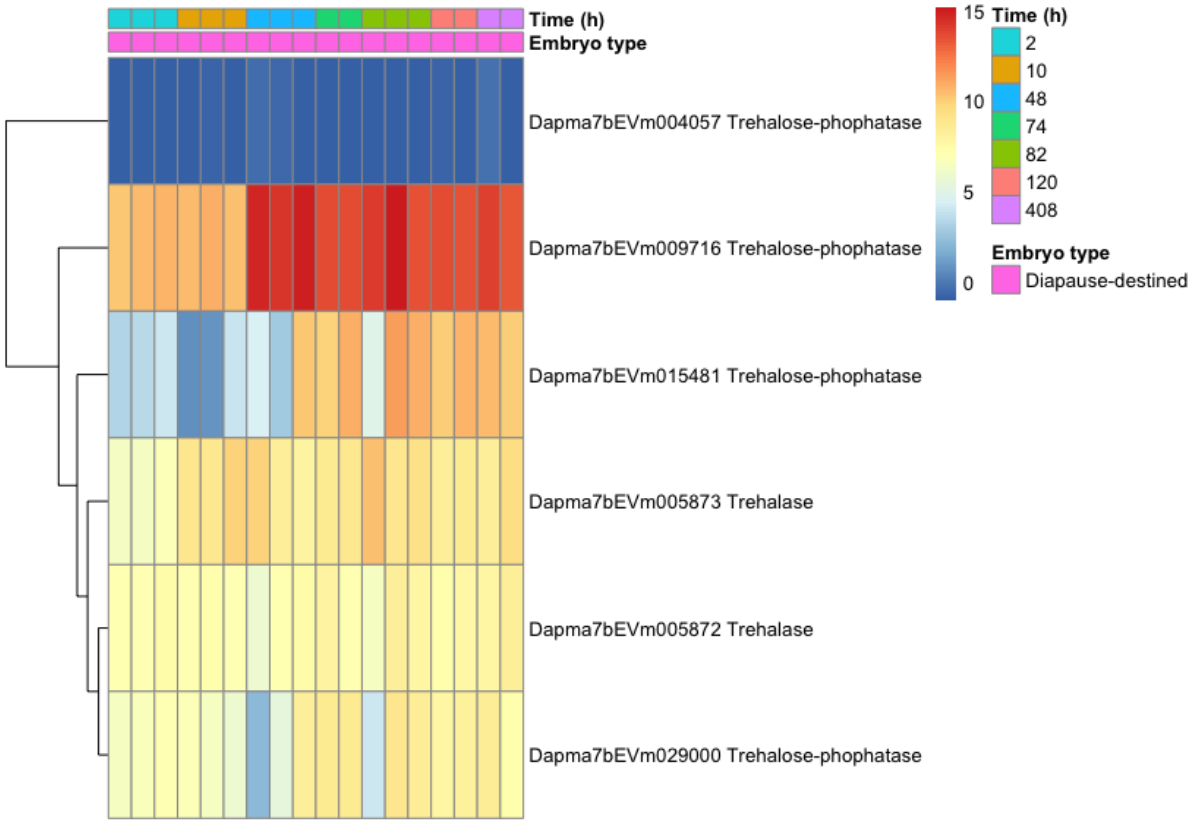


Figure 5-4 Heat map and hierarchical clustering of genes relating to trehalose in samples of sexually produced *Daphnia* embryos collected between 2h and 2 weeks (408h) post-ovulation. Row labels indicate gene name with corresponding annotation based on OrthoDB orthologous gene family group and coloured column top panels indicate time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale**.**

Trehalose is the predominant blood sugar in insects. However, it is also an important energy source implicated in stress resistance (Tang *et al.*, 2018). Trehalose has been implicated in anhydrobiosis through the stabilisation of proteins in their native state and may participate in desiccation resistance as a vitrifying agent (Hengherr *et al.*, 2011; MacRae, 2016). Trehalose may also contribute to cold tolerance as a natural anti-freeze substance offering cryoprotection during diapause through

depression of organism super cooling point (Su et al., 2017). For example, in *Drosophila montana,* increased levels of trehalose in addition to glucose, proline, and *myo*-inositol correlated with overwintering cold tolerance (Vesala *et al.*, 2012). *Helicoverpa armigera* pupae also accumulate trehalose prior to the onset of diapause and its levels correlate with TPS enzyme activity, whilst expression of TPS cDNA is also higher in diapausing individuals (Xu et al., 2009). TPS and TPP gene expression and enzyme activities were also observed to increase in line with pupal trehalose content during the onset of *Delia antiqua* diapause (Guo *et al.*, 2015). Diapausing *Artemia franciscana* accumulate trehalose during the onset of diapause, and its levels remain high throughout 4 y in anoxic conditions (Clegg, 1997; MacRae, 2003). By contrast, during the initiation of *Bombyx mori* diapause, diapause hormone stimulates entry of the embryo into diapause through increasing mRNA expression of trehalase in the maternal ovary, facilitating conversion of trehalose into glycogen in embryos (Yamashita, 1996).

Altogether, this evidence suggests that trehalose may be important in *Daphnia* diapause based on increasing intensity of a metabolite putatively annotated as trehalose upon entry into diapause, plus the detection of two enzymes involved in the synthesis of trehalose (trehalose phosphatases) showing similar patterns of gene expression. Furthermore, there were no clear trends in the expression of any of these genes in asexual embryo samples (Section 10.3; Figure 10-2), suggesting a diapause specific role. Analysis of trehalose levels in *D. magna* and *D. pulex* has previously been reported in relation to a possible role in the desiccation resistance of diapausing embryos. However, these levels were low (accounting for < 0.5% of dry weight), and no detectable changes occurred during exposure to desiccation

(Hengherr *et al.*, 2011). Two isoforms of trehalase have been described in several insects. In *B. mori* these correspond to a soluble and membrane bound form (Mitsumasu *et al.*, 2008). In several species of insects, only TPS and not TPP genes have been identified. For example, in *D. melanogaster Tps1* contains both TPS and TPP domains capable of *de novo* trehalose synthesis (Yoshida *et al.*, 2016; Tang *et al.*, 2018). Therefore, robust characterisation of these four genes in *D. magna* will be required to determine their exact function as their annotations are currently based on predicted protein similarity. Annotation of the 'trehalose' peak will require further experimental confirmation.

5.2.5. Putative significantly expressed genes

Of the total genes detected in each respective differential expression analysis, the majority were successfully annotated with PANTHER protein domains and/or OrthoDB orthologous gene families (Section 2.13.4). For the 48 v 10 analysis, of the total 16,790 genes (Table 5-1) 11,322 were assigned to an OrthoDB Orthologous Group and 12,235 were successfully annotated with a PANTHER protein family classification. For the 74 v 48 analysis, 10,860 and 11,726 out of a total 15,630 genes (Table 5-1) were successfully annotated with OrthoDB and/or PANTHER respectively. Differentially expressed genes in the two analyses that I performed (10 v 48; 74 v 48) were divided into lists of successfully annotated up-regulated or down-regulated genes (FDR < 0.1; Table 5-5); gene ontology GO annotation overrepresentation analysis was then performed on each group separately (Section 2.13.4).

Table 5-5 Table of number of genes that displayed a significant (false discovery rate (FDR) adjusted *p*-value < 0.1) change in expression relative to the initiation of diapause at 48 h v 10 h or 74 h v 48 h post-ovulation in *Daphnia magna* embryos. Up- or Down-regulated genes are based on expression at the later time point versus the earlier. Of the differentially expressed genes the number with putative PANTHER gene family annotations used for further GO-annotation overrepresentation analysis are indicated.

		Time point compared (h)	
		48 v 10	**74 v 48**
Total detected		16790	15630
Significant (p < 0.1)		5127	229
Up	**Total**	2681	135
	PANTHER annotated	2093	29
Down	**Total**	2446	94
	PANTHER annotated	1893	77

Within the 94 down-regulated and 135 up-regulated genes differentially expressed at 74 h in relation to 48 h, 77 and 29 genes respectively were annotated (Table 5-5), no significantly overrepresented GO terms were identified at $p < 0.05$ after FDR correction (full results table presented in Section 10.4; Table 10-7). Within both down-regulated and up-regulated genes, a large number of genes were unclassified (42-62%) within each GO domain. Within the down-regulated genes (74 v 48; Table 10-7), although not statistically overrepresented (*p*-adjust > 0.05), several annotations relating to metabolic processes (e.g. carbohydrate metabolism (6 genes)) and more specifically protein glycosylation (4 genes) were present, in addition to single genes in the steroid, polysaccharide, and fatty acid metabolism domains. Altogether indicating that after the point of developmental arrest, there may still be on-going transcriptional changes including those related to metabolism, which is in line with the high number of metabolites changing in intensity during this period. Manual inspection of GO-annotations for genes with up-regulated expression (74 v 48) revealed 3 genes within the ‘response to stress’ (Section 10.5). These were E3

Ubiquitin-protein ligase RAD18, 5'-adenosine monophosphate (AMP)-activated protein kinase subunit gamma-2, and a Cu/Zn superoxide dismutase based on pantherScore hidden Markov models (Mi et al., 2017). AMP-kinase is activated in response to a depletion of cellular ATP, and drives the activation of ATP-generating catabolic pathways, but also inhibits a variety of biosynthetic pathways to conserve energetic resources (Hardie, Carling and Halford, 1994). AMP-activated kinase expression has previously been observed to be up-regulated in *D. pulex* during toxicant (nanoplastic) induced stress (Hardie, Carling and Halford, 1994; Liu et al., 2018). AMP-activated kinase is consistently expressed throughout all biological samples sequenced in this study (Section 10.5; Figure 10-4). However, it is much more abundant during the diapause period and thus may serve as an indication that alterations in energy status are occurring. RAD18 participates in the post-replication repair of DNA damage, including ultra violet-induced lesions, through recruitment to lesions and monoubiquitination of proliferating cell nuclear antigen (PCNA; Bailly *et al.*, 1994; Durando, Tateishi and Vaziri, 2013). Monoubiqutination of PCNA by RAD18 facilitates the exchange of replicative DNA polymerases for non-classical polymerases (e.g. pol η) which repair the lesion (Freudenthal *et al.*, 2010). Expression of this gene was highly sporadic within biological samples (Section 10.5; Figure 10-4) with particularly high expression in one biological replicate at 74 h post-ovulation and thus may reflect active DNA-repair processes occurring in this particular biological sample rather than related to the diapause phenotype as a whole.

Further investigation yielded a total of 9 genes with putative SOD function within the transcripts expressed in this study (Section 10.6; Table 10-8). Of these, four showed

a significant fold change between 48 h and 10 h in addition to the aforementioned SOD between 74h and 48h; their expression throughout the sampled sexual embryo time points is shown in Figure 5-5. Dapma7bEVm004779 (Cu/Zn SOD) shows somewhat varied expression throughout the time course, appearing highly expressed at 74 h and 408 h (2 weeks) post-ovulation and low in 2 h samples, with disparate expression levels during the intervening time points. The other three SOD genes; two Cu/Zn SODs (Dapma7bEVm028228, Dapma7bEVm008098) and a Mn/Fe SOD (Dapma7bEVm003936) have remarkably similar expression, with low expression at 2 h and 10 h post-ovulation followed by a large increase in expression by 48 h post-ovulation. All three had $\log_2$ fold changes over 6 (48 v 10) and were also expressed at a high level throughout the other time points sampled after 48 h and thus into the period of diapause (Figure 5-5). I also investigated the expression of these genes within subitaneous samples (Section 10.6; Figure 10-5) to confirm that this expression was diapause specific instead of related to development in general and no clear patterns in gene expression were observed in asexually produced samples.

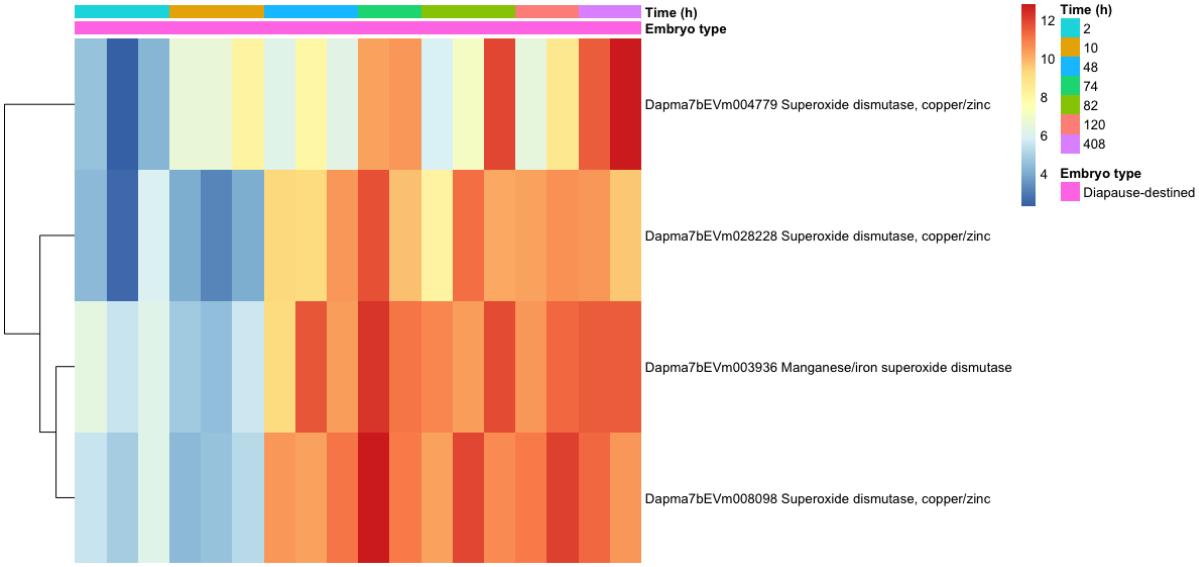


Figure 5-5 Heat map and hierarchical clustering of putatively annotated superoxide dismutase genes in samples of sexually produced *Daphnia* embryos collected between 2 h and 2 weeks (408 h) post-ovulation that had significant (false discovery rate corrected $p < 0.01$) large $\log_2$ fold changes between 10 h and 4 h or 48 h and 74 h (Dapma7bEVm004779 only) post-ovulation. Row labels indicate gene name with corresponding annotation based on OrthoDB orthologous gene family group and coloured column top panels indicate time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale.

Superoxide dismutases form part of the enzymatic cellular defences against oxidative stress induced by normal respiratory processes in addition to exogenous perturbation (Jena *et al.*, 2013). They are responsible for the detoxification of the superoxide anion radical $O_2^{-\bullet}$ to yield H_2O_2 and O_2 (Lyu *et al.*, 2014) and are a multi gene family that differ in terms of the metal cofactors at the catalytic centre and subcellular location (Zelko, Mariani and Folz, 2002). Currently, a Cu/Zn SOD and Mn SOD have been characterised in *D. magna,* the expression and enzyme activity of which have been demonstrated to increase upon exposure to a variety of stresses including anoxia, and heavy metals (Lyu *et al.*, 2013, 2014).

Several studies in different species have reported increased gene expression, protein levels and/or activity of SODs in diapausing individuals versus non-diapausing (Houthoofd *et al.*, 2002; Tu, Wang, Hao, Douglas W Whitman, *et al.*, 2015; Kučerová, Olga I. Kubrak, *et al.*, 2016). By contrast, SOD enzymatic activity is lower in the

diapausing larvae of *Chouioia cunea* versus non-diapausing (Zhao *et al.*, 2018), and is also reduced during *Osmia bicornis* obligate imago stage diapause (Dmochowska-Ślęzak *et al.*, 2015).

An increased level of antioxidant protection during diapause is logical in terms of protection from oxidative stress that may be encountered during the period of dormancy. Indeed RNAi knockdown of *catalase* and *SOD-2* in diapausing *Culex pipiens* led to a significant reduction in survival (Sim and Denlinger, 2011). Additionally, transgenic *D. melanogaster* lines with additional *SOD* and *catalase* genes displayed up to a 34% increase in lifespan (Orr and Sohal, 1994). Thus, mitigation of oxidative stress may contribute to the extension of lifespan observed in *D. magna* diapause, although direct evidence of increased SOD enzymatic activity in addition to investigation throughout the period of diapause is still required. Additionally, a vitellogenin-SOD fusion peptide has been identified in both *D. magna* and *Artemia parthenogenetica*, although in both species SOD enzymatic activity is yet to be confirmed (Kato *et al.*, 2004; Chen *et al.*, 2011). In *A. parthenogenetica,* different patterns of proteolysis occurring in the region of the SOD-like domain between diapause and non-diapause cysts are observed (Chen *et al.*, 2011). This fusion protein is highly abundant in *D. magna* embryos, and gene expression and proteomic analysis has revealed that they are over expressed in post-diapause *D. pulex* females in relation to subitaneous (Kato *et al.*, 2004; Kaupinis *et al.*, 2017). However, this orthologous gene in *D. magna* (Dapma7bEVm009708) was identified within the dataset but was present at low levels and did not display any diapause specific patterns of expression.

The results from GO-annotation overrepresentation of genes up- and down-regulated at 48 v 10 are reported in Figure 5-6 (tabulated in Section 10.4; Table 10-6). There were 31 overrepresented terms within the down-regulated gene set (1893 out of 2446 genes; Table 5-5) and just two overrepresented GO-annotation within the up-regulated gene set (2093 out of 2681 genes; Table 5-5).

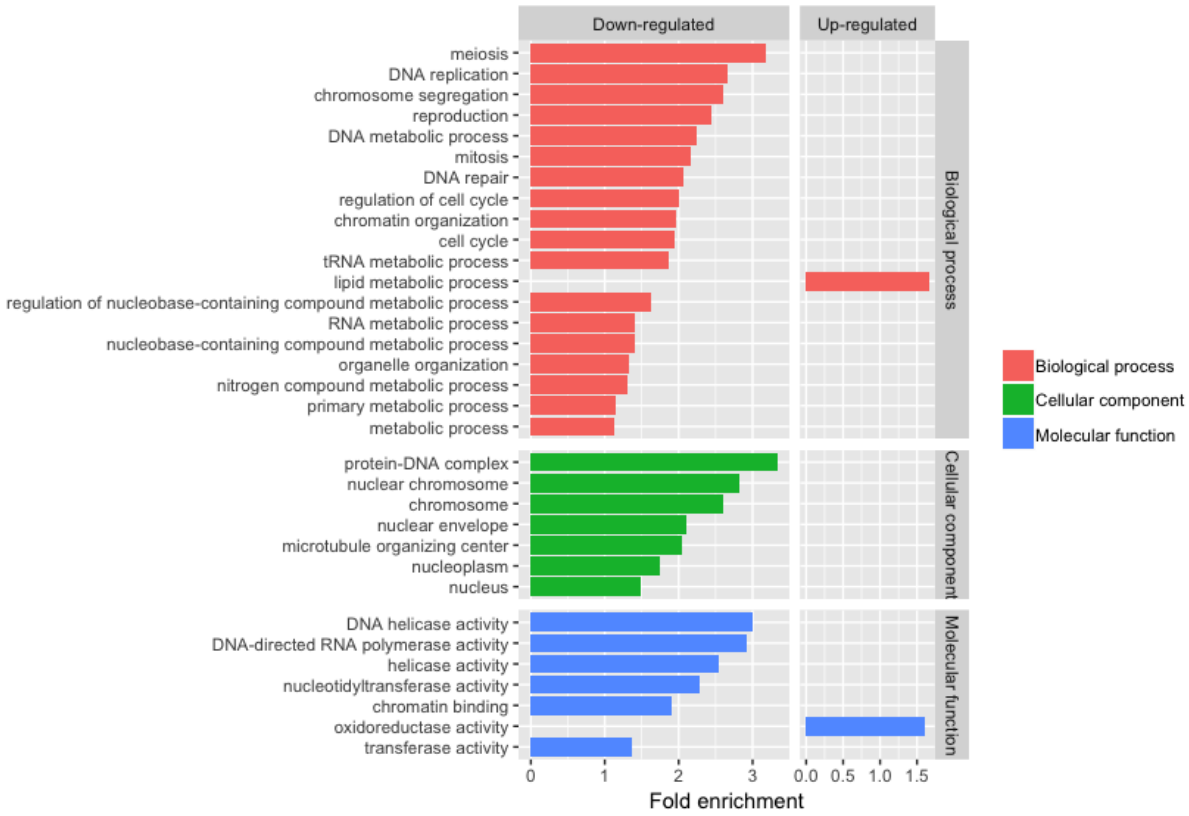


Figure 5-6 Gene ontology (GO) over-representation analysis of genes that are differentially expressed between 10 h and 48 h post-ovulation in sexually produced embryos of *Daphnia magna*. 'Up-regulated' genes refer to those with higher expression in 48 h embryos and vice versa for 'down-regulated' genes. Each gene set (i.e. up- or down-regulated) was investigated for overrepresented annotations within each GO domain covering biological processes, cellular components and molecular functions.

The two overrepresented GO-annotation terms within the up-regulated gene set included 'lipid metabolic process' and 'oxidoreductase activity'. The full list of genes assigned to each annotation term are reported in Section 10.7 (Table 10-10), including 34 genes that were assigned to both GO categories. Oxidoreductase activity covers a diverse range of enzymes that participate in cellular redox reactions

including dehydrogenases, reductases and transferases. Genes assigned this annotation included the SODs already identified in this chapter, in addition to the glutathione-related genes already discussed in Section 4.6.2 (Figure 4-8). Furthermore, several genes that appear to be related to the mitochondrial electron transport chain (e.g. NADH-ubiquinone oxidoreductase (Complex I) and Ubiquinol-cytochrome C reductase (complex III)) were also annotated (Guo *et al.*, 2018). The up-regulation of mitochondrial electron transport genes and the overrepresentation of lipid metabolic processes in the set of genes up-regulated upon entry into diapause was surprising in that, throughout this study, minimal changes in the lipidome have been observed and in general a down-regulation of metabolism is anticipated as a hallmark of diapause (Section 4.4) (Hand and Podrabsky, 2000).

As already established, another well-characterised hallmark of diapause is developmental arrest. Many of the overrepresented GO annotations (Figure 5-6) in the set of genes down regulated at 48 h (onset of developmental arrest) versus 10 h (during active development) reflect biological processes involved with cell division including ‘DNA replication’ and ‘cell cycle’. Heatmaps of the expression (log-norrmalised counts) of genes assigned to these two terms are shown in Figure 5-7, and the full lists of putative gene functions and labelled heatmaps are presented in Section 10.8. The overall pattern of gene expression of genes within these two categories is remarkably consistent and demonstrates a clear clustering of embryo samples before (2 h and 10 h) and after (48 h to 408 h) the onset of diapause developmental arrest. In general, the expression of these genes is elevated prior to the onset of diapause, and in line with on-going cellular division occurring during pre-diapause development. Then upon entry into diapause, by 48 h there is consistent

down-regulation of these genes that is maintained up to 2 weeks (408 h) into diapause.

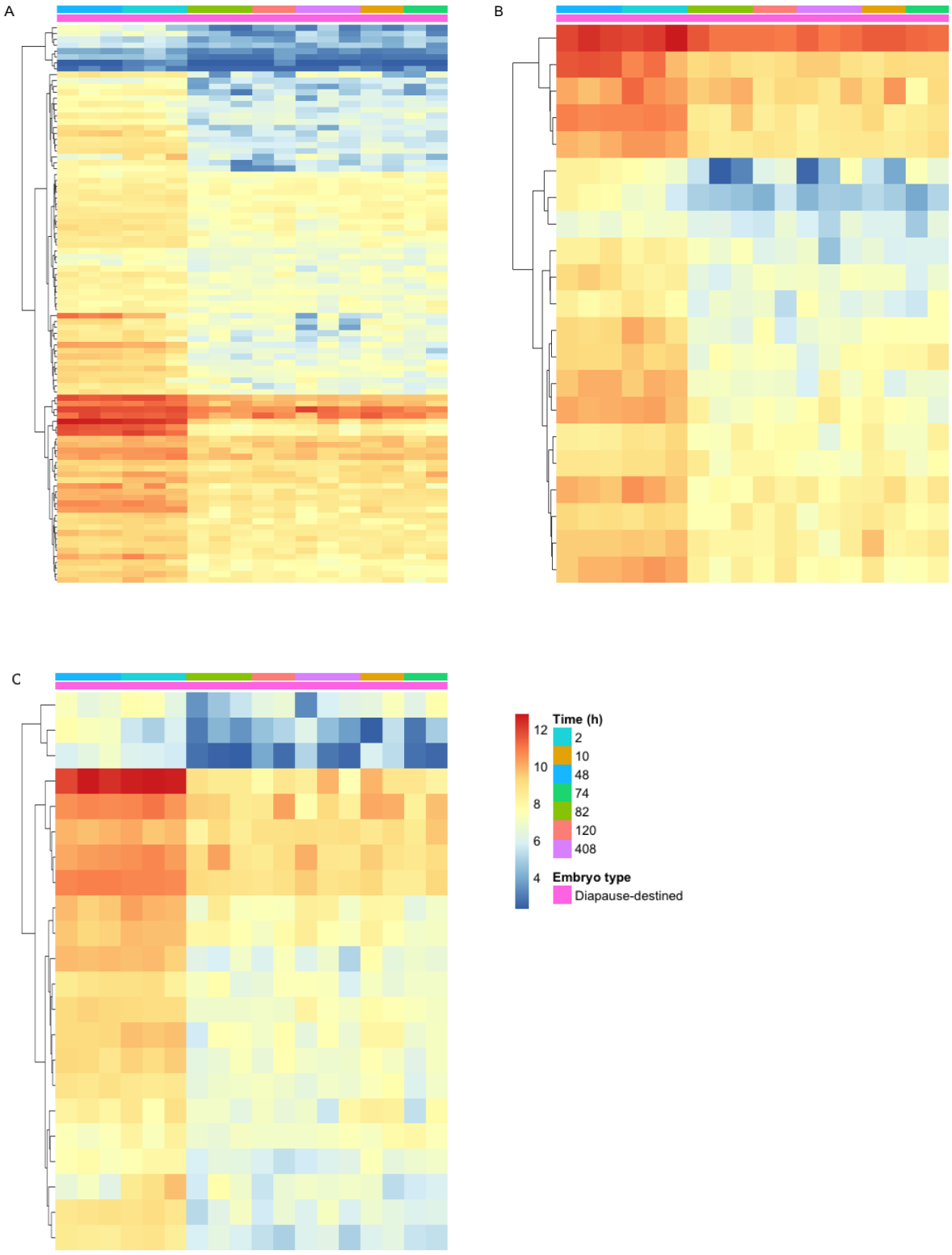


Figure 5-7 Heatmap and hierarchical clustering of genes annotated with the gene ontology terms 'cell cycle' (A; 95 genes), 'DNA replication' (B; 21 genes), and shared between the two terms (C; 22 genes) that were significantly (false discovery rate corrected $p < 0.1$), differentially expressed between 10 h and 48 h in sexually produced embryos of *Daphnia magna*. Coloured column top panels indicate time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale.

43 genes annotated with functions involving ‘DNA replication’ (listed fully in Table 10-12) include various DNA polymerase subunits. Other factors required for the formation of complexes required for the initiation of DNA replication such as those associated with the origin of replication complex and minichromosome maintenance (Méndez and Stillman, 2000), in addition to *pcna* are also annotated.

The 117 genes annotated to the ‘cell cycle’ GO term (Table 10-11) include various cyclins (e.g. *cyclin B*, *D*, and *E*) which were identified as also being down-regulated upon entry into developmental arrest. Progression of eukaryotic cells through the respective phases of the cell cycle during cell division is, amongst other processes, regulated by the action of cyclin-dependant kinases (cdks), which are serine/threonine kinases. Cdks are in turn regulated by cyclin levels that fluctuate throughout the cell cycle, driving or inhibiting cell cycle progression (Vermeulen, Van Bockstaele and Berneman, 2003). The behavior of various cyclin genes was previously reported in other studies investigating diapause, particularly events surrounding initiation. Reduced transcript levels of *cyclin D* and *cyclin D2* were observed during diapause maintenance in *D. antiqua* (Hao *et al.*, 2016). Levels of *cyclinD, cyclinE* in addition to *pcna* mRNA were lower in *Nasonia vitripennis* diapausing larvae versus non-diapausing (Shimizu, Mukai and Goto, 2016). Protein levels of cyclinK are reduced in diapausing *A. parthenogenetica*, and authors propose that cyclinK is involved in the regulation of RNA polymerase II upon the termination of diapause. Curiously however authors note that *cyclinK* transcript levels are not reduced in diapausing cysts (Zhao *et al.*, 2012). Cell cycle progression is also dependent on a variety of other regulatory proteins. Wee1 is a protein kinase that inhibits cyclin-cdk complex activity, *wee1* transcript levels increased upon initiation of

diapause in *C. costata* larvae (Koštál *et al.*, 2009). By contrast however, a wee1-like protein kinase (Dapma7bEVm004149) was also down-regulated upon entry into diapause within this study.

The successful detection and identification of genes discussed above that have previously been identified in other studies investigating diapause represent an important 'positive control' for this investigation, in that I am able to detect genes anticipated to be involved in diapause-associated developmental arrest. This means that further interrogation of the dataset and any conclusions derived, can be undertaken with increased confidence. Furthermore, it demonstrates that there may be conservation across species with regards to the mechanisms of developmental arrest as has been previously hypothesised (Poelchau *et al.*, 2013a). And therefore, that conclusions drawn from considering observations in the *D. magna* diapause model system regarding genes, pathways, and/or regulatory pathways are likely to be relevant to diapause in other species (Ragland, Denlinger and Hahn, 2010; Poelchau *et al.*, 2013a).

5.3. Conclusions

The aim of this chapter was to characterise the transcriptomic and metabolic changes that occur specifically at the initiation of diapause in *Daphnia magna* through a comparison of metabolite intensities and gene expression levels before (10 h post-ovulation) and immediately after (48 h post-ovulation) diapause initiation, in addition to a further point in early diapause (74 h).

I identified profound molecular changes in relation to the initiation of diapause at both the mRNA and metabolite levels. A total of 2231 genes were significantly differentially expressed at 48 h versus 10 h with roughly half up- and down-regulated (Table 5-1) representing 13% of the total genes detected. However, in contrast at 48 h versus 10 h only 4 and 9 metabolite levels changed significantly in the positive and negative analysis of polar *D. magna* extracts in the metabolomics analysis whilst, no peaks showed significant alterations in intensity in the non-polar extracts. Between 48 h and 74 h, a much smaller portion of total genes (0.3%) differed in expression, whilst large changes in metabolite intensities in polar extracts occurred with 28 and 37 peaks changing significantly in intensity in negative and positive ion analysis respectively. This may reflect that changes in gene expression can manifest more rapidly than those in metabolites due to the intermediate processes and regulatory steps between the two phenotypes (Feder and Walser, 2005; Hayward, 2014).

Interestingly, levels of the majority of metabolites increased upon the initiation of diapause, and the reason for this is unknown. There is no current evidence to suggest that *D. magna* exhibit comparable suppression of metabolism during diapause as seen in other species. For example a ninety percent reduction in respiration rate was observed in the diapausing pupae of *S. crassipalpis* versus non-diapausing pupae, and *A. franciscana* embryos also show a rapid decrease in respiration rate over the first five days of diapause with respiration after 20 days reducing to just 2% of original levels (Ragland, Denlinger and Hahn, 2010; Podrabsky and Hand, 2015). Thus it is unclear whether the increase in metabolite intensity upon the initiation of diapause I observed is as a result of the accumulation

of metabolic intermediates as a result of metabolic suppression or potential energy sources sequestered for use during the diapause period and post-diapause development, or whether they play a specific role in maintaining the diapause phenotype, through cryoprotective and other roles or perhaps more likely a possible combination of the three.

Of the metabolites changing significantly in relation to the onset of developmental arrest, a peak putatively annotated as a disaccharide and potentially trehalose, a molecule implicated in the diapause and concomitant desiccation resistance and cold tolerance of other organisms showed a 28-fold increase in intensity in relation to the onset of diapause in *D. magna*. Additionally, several trehalose phosphatases that catalyse the synthesis of trehalose from UDP-glucose also exhibited large positive fold changes in relation to the onset of diapause. However, based on the current level of annotation this peak may be one of a further 64 metabolites and will thus require further investigation to improve confidence in the annotation before a role of trehalose in *D. magna* diapause can be confirmed.

I also identified several SOD genes that were up-regulated at the initiation of diapause with high mRNA levels persisting throughout the period of diapause, and together with the results regarding glutathione discussed in Section 4 suggest that the diapausing embryos of *D. magna* are primed to respond to oxidative stress.

Of the numerous genes whose mRNA levels changed significantly in relation to the onset of diapause, genes assigned to the GO annotations DNA replication and cell cycle were overrepresented and include *pcna* and *cyclin* genes. These genes are of particular interest as they represent likely effectors of developmental arrest and candidates for a starting point to identify the currently unknown regulatory processes

of diapause. The previous identification of these genes (analogous to a ‘positive control’) in other studies investigating the mechanisms of diapause in several species increases confidence in the data presented. In that they are both robust (i.e. able to successfully identify ‘known knowns’ regarding the diapause process) and demonstrates that *Daphnia magna* is a suitable model system relevant to the diapause phenotype as a whole.

Furthermore, there are thousands of other genes connected to diapause in *D. magna* that were not investigated due to the time limitations of this thesis that may represent other important processes in the diapause phenotype. Notably only two GO annotations were overrepresented in the lists of genes up-regulated upon entry into diapause (48 h v 10 h) despite this containing over a thousand genes, and there were many more genes without an annotation. Further analysis and manual inspection of data will be required to elucidate the functional importance of any genes present.

Regardless, the lack of unequivocal annotation of both genes and metabolites identified in this section does not detract from the profound molecular changes described that occur in relation to the onset of diapause-associated developmental arrest. This highlights that diapausing embryos (>48 h post-ovulation) can be defined by molecular profiles different to those of pre-diapause embryos (as represented by 10 h samples) at both the level of the transcriptome and (polar) metabolome. In addition, there appear to be a distinct temporal separation between the wide-spread diapause-induced transcriptomic and metabolic effects, with changes manifesting later in the metabolome (i.e. 74 v 48) than the transcriptome (48 v 10).

6. Molecular profiling of embryonic diapause in *D. magna*

6.1. Introduction

As an alternative developmental trajectory, diapause is a dynamic process that progresses through a program of phases including pre-diapause, initiation, maintenance and termination (Koštál, 2006). To date, studies investigating the molecular underpinnings of diapause maintenance have largely focussed on comparisons of organisms in diapause with a directly developing non-diapausing stage (Monica F Poelchau *et al.*, 2013b; Kang *et al.*, 2016; Liu *et al.*, 2017; Y. Chen *et al.*, 2017). However, in Section 4 I observed that diapause-destined and subitaneous *D. magna* are inherently different on a cellular and molecular level from the outset of development, highlighting the limitation of this approach. In addition, only a handful of studies in insect species, summarised in Table 6-1, have considered more than one phase of diapause development. Of the six studies that I identified (5 investigating gene expression and 1 covering the lipidome), two performed only pairwise comparisons between diapause and non-diapause development. The remaining four studies, whilst covering multiple diapause phases were addressing pupal and larval diapauses. Thus, there is a significant knowledge gap regarding the temporal molecular changes occurring in embryonic diapause. Furthermore, where information was provided regarding the endogenous period of the diapause development considered, all were less than 1 year–with mortality observed after 7 months (Koštál *et al.*, 2017) and 100% mortality after 1 year (Lehmann *et al.*, 2016)–with the most extended molecular sampling regime

considering diapause development over a 6 month period (Dong *et al.*, 2014). Therefore, despite prolonged diapause (i.e. lasting over 1 year), having been described in a variety of organisms (Hanski, 1988; Solbreck and Widenfalk, 2012; Ploomi *et al.*, 2018), I was unable to find any studies considering the mechanistic basis for this phenomenon at a molecular level. However, the conclusion of any individual dormancy is dependent on the navigation of all successive phases of development, and logically the patterns of gene expression and metabolite changes occurring during these phases will contribute to a successful diapause outcome. As outlined in Chapter 4 diapause in *D. magna* is initiated approximately 48 hours after ovulation (at 20°C), thus I am able to define the time points when molecular data were collected for this thesis in relation to the phases of diapause as defined by Koštál (Koštál, 2006; Koštál *et al.*, 2017). Therefore, my sampling points were set in pre-diapause development, spanning immediately from conception to the stage at which direct development arrests at the initiation of diapause. Further sampling points during “early” diapause, covering the first 2 weeks after the onset of developmental arrest were collected with an additional sampling of the metabolome only at “late” diapause of embryos in a state of suspended animation for two years.

Table 6-1 Summary of published studies investigating diapause development including relevant organism and diapause life history and experimental design. Where reported, total duration indicates maximum time in diapause until onset of mortality. Data type collected and statistical design used in analysis are also indicated (PCA = principal component analysis, GLM = generalized linear models). Experimental parameters include induction of diapause method by changing photoperiod (LD = Long Day; SD = Short Day; values reported as light:dark) and sampling time points (days) are indicated by grey cells relative to organism diapause phase.

								Pre-diapause	Diapause			Post-diapause	
Study	**Organism**	**Life stage**	**Diapause type**	**Total duration**	**Data type**	**Design**	**Induction method**		**Early**	**Mid**	**Late**	**Quiescence**	**Termination**
(C Sim and Denlinger, 2009)	*Culex pipiens*	Adult	Facultative	–	a priori qRT-PCR	Relative to non-diapause pairwise	LD v SD: 15:9 v 9:15 18°C	–	7 d	–	60 d	–	14d post transfer to 25°C, LD(15:9)
(Ragland, Denlinger and Hahn, 2010)	*Sarcophaga crassipalpis*	Pupae	Facultative	~60 d	Transcriptome-wide microarray	All pairwise	LD v SD: 15:9 v 10:14 25°C	5 d	20 d	–	50 d	–	50 d (1d post topical hexane)
(Poelchau *et al.*, 2013b)	*Aedes albopictus*	Larvae	Facultative	–	Whole transcriptome RNA-Seq	Relative to non-diapause pairwise	LD v SD: 16:8 v 8:16 21°C	–	11 d	21 d	40 d	–	
(Dong *et al.*, 2014)	*Bactrocera minax*	Pupae	Obligatory	–	Whole transcriptome RNA-Seq	All pairwise	Wild capture	–	30 d	–	90 d	–	180d (topical 20-hydroxyecdysone
(Lehmann *et al.*, 2016)	*Pieris napi*	Pupae	Facultative	~120 d	Lipidomics	PCA + GLM	LD v SD: 22:2 v 10:14 20°C	–	24 d	–	–	144 d	endogenous at 148 and 158 d
(Koštál *et al.*, 2017)	*Chymomyza costata*	Larvae	Facultative	~210 d	Custom microarray	Pairwise: successive time points	LD v SD: 16:8 v 12:12 18°C	14,21,28 and, 42 d	70 d	98 d	–	–	77 d (7 d post transfer to LD (16:8))

In this chapter, I expand the investigation of transcriptomic and metabolomic molecular data to encompass the entire timeline of diapause, chiefly with the aim to assess whether there are distinct patterns of gene expression or metabolite intensity that define each of the stages of diapause. Two separate complementary analyses were performed to achieve this aim. First, sparse principal component analysis (sPCA) was implemented in the un-biased assessment of whether there are inherent differences in sample molecular profiles in relation to collection time post-ovulation and thus, diapause phase (Section 6.2.1). Next, genes and metabolic peaks assigned non-zero loadings to sparse principal components and thus contributing to the identified separation were also investigated to ascertain possible relevance to the diapause phenotype. Lastly, through direct pairwise comparisons between successive time points, I aim to examine the number and magnitude of molecular changes with time, along the diapause developmental trajectory, and to assess whether any perturbations are localised to specific stages (i.e. time interval; Section 6.2.2). Of particular interest in this analysis are possible changes in metabolites between the “early” diapausing stages and “late” diapause stage collected 2 years after the onset of diapause. These will be particularly relevant in understanding how the metabolism of diapausing organisms is shaped by the physiological requirements of extended dormancy.

6.2. Results and discussion

6.2.1. *Comparative analysis of the molecular modalities*

Using sexually produced embryo samples only, gene count matrices were compiled for all time points (T), from T = 2 h through to T = 408 h (2 weeks) post-ovulation. The transcriptome data were log-normalised using *DESeq2* (Section 2.13.3) revealing a total of 23,517 genes that were expressed in at least one time point, and 5,422 that were differentially expressed when comparing levels of transcription between two subsequent time points (i.e. T_n / T_{n-1}) at a false discovery corrected threshold of $p < 0.01$ (6,682 at p-adjust < 0.05).

To visualise any similarities in metabolic and transcriptomic profiles of biological samples with relation to diapause phases (i.e. time post-ovulation), sPCA was then performed on this normalised count matrix, then again separately on each of the log-normalised metabolite intensity matrices obtained from the polar and non-polar sexual embryo only metabolomics analyses (Section 9.3.2). The first four sparse principal components (sPCs) were considered. But as these describe the largest portions of variance within each dataset, scores plots of sPC1 and sPC2 are presented in Figure 6-1.

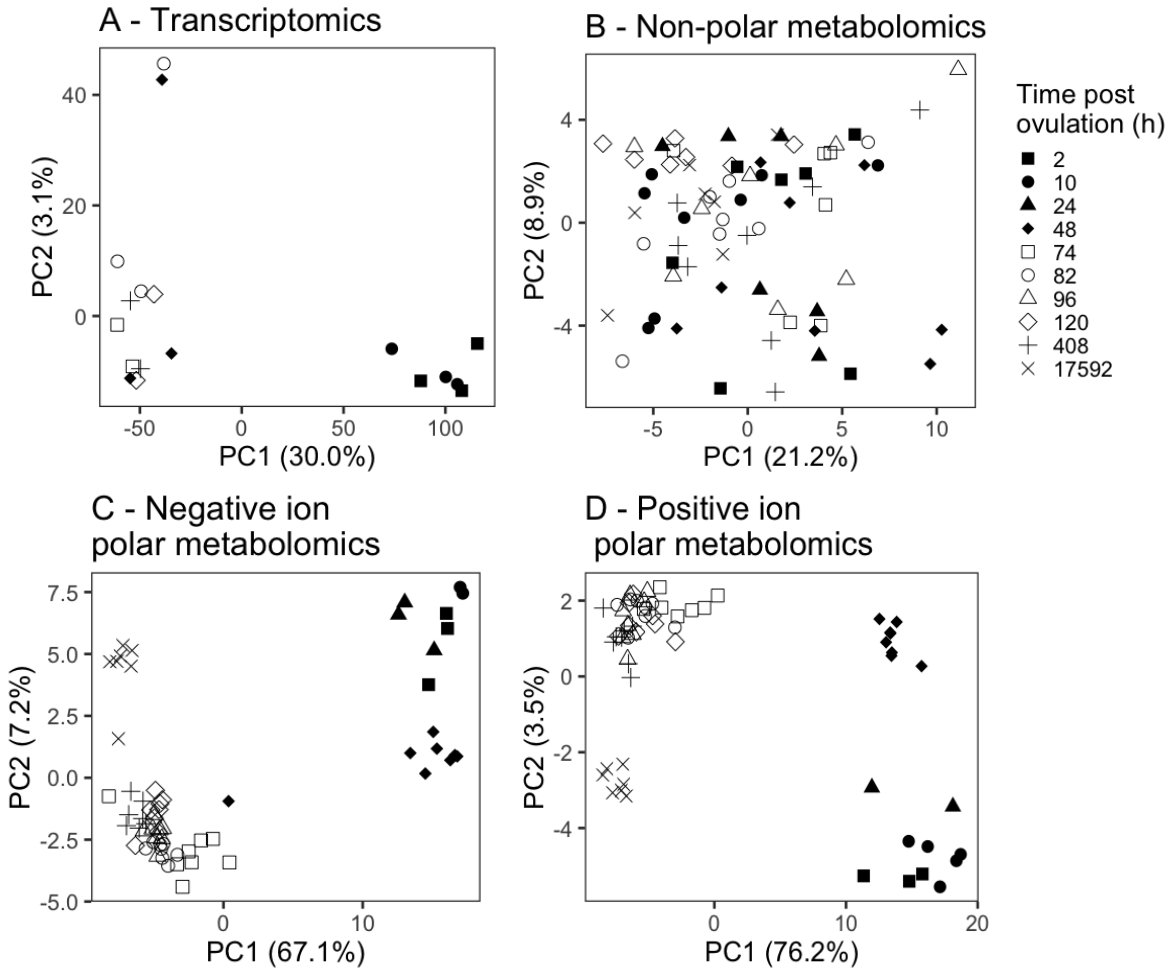


Figure 6-1 Scores plots from the sparse principal component analysis (sPCA) of log-transformed gene counts (A) and metabolite intensities (B-D) of diapause-destined and diapausing embryos of *Daphnia magna* collected between 2 h and 2 y (~17592 h) post ovulation; Metabolomics analysis of B) Negative ion non-polar extracts, C) Negative ion polar extracts, D) Positive ion polar extracts.

sPCA of transcriptomic data (Figure 6-1; A) showed a separation between pre-diapause samples at 2–10 h and the samples collected during diapause i.e. +48 h along sPC1, which accounts for the largest amount (30%) of the variance within the dataset. Considering the metabolomics data, there was no distinct clustering of biological samples classes in the non-polar extracts along sPC1 and sPC2 (Figure 6-1; B). This observation is aligned with the results obtained so far in this thesis; there are limited changes in the lipid profiles of *D. magna* embryos during diapause. However, sPCA of the polar extracts did yield potential metabolic changes in relation to the time of collection post-ovulation and thus diapause development. Specifically,

for polar extracts in both negative (Figure 6-1; C) and positive (Figure 6-1; D) ion mode, there was a clear separation between samples collected in pre-diapause at 2–48 h and those in diapause collected after 48 h along sPC1. These data suggest that there are both distinct transcriptomic and polar metabolic profiles for embryos at pre-diapause and diapause stages of development. In transcriptomic data, this distinction occurs between pre and post 48 h samples, whilst in polar metabolic samples, this distinction is later between pre and post 74 h samples. This finding is in agreement with the observations made in Chapter 5 of large fold changes in metabolic peaks occurring slightly after the onset of developmental arrest, between 48 h and 74 h samples. I anticipate that many downstream metabolic changes associated with diapause may take slightly longer to manifest (i.e. an additional 24 h) than the up-stream transcriptomic changes.

Considering sPC2, the positioning of transcriptomic samples leads to the separation of only two samples, single biological replicates of the 48 h and 84 h groups. Whilst along sPC2 in both polar metabolomics analyses, there was a separation of the 48 h group from time points collected in the pre-diapause period (2–24 h), and the embryos collected at 2 y (17,592 h) post-ovulation also partitioned out from the other samples collected in diapause. This result indicates that subsets of metabolic perturbations do occur promptly around the initiation of diapause and that there are changes in metabolism associated with longer-term diapause.

Table 6-2 Summary of sparse principal component analysis conducted on the log-normalised gene count and metabolite peak intensity matrices from the analysis of *Daphnia magna* sexually produced embryos collected between 2 h and 2 y (17592 h) post-ovulation, including the proportion of variance explained and the number of elements (peaks or genes) assigned non-zero loadings for the first four sparse principal components.

		Sparse principal component	Standard deviation	Proportion of variance	Number of elements
Transcriptomics		**1**	86.536	30.0	114
		2	27.992	3.1	17
		3	25.158	2.5	14
		4	6.272	0.2	2
Metabolomics	**Negative ion non-polar metabolomics**	**1**	5.3490	21.2	40
		2	3.4610	8.9	22
		3	3.0650	7.0	10
		4	2.5230	4.7	20
	Negative ion polar metabolomics	**1**	9.4500	67.1	49
		2	3.1040	7.2	32
		3	1.9260	2.8	7
		4	1.7880	2.4	26
	Positive ion polar metabolomics	**1**	10.4180	76.2	36
		2	2.2190	3.5	14
		3	1.5710	1.7	25
		4	1.5240	1.6	9

To further investigate the metabolomic and transcriptomic changes occurring throughout the entire time-course that may be masked by the predominant effect of the onset of diapause, the sample scores for sPC3 and sPC4 were also assessed in each dataset for biologically relevant clustering of samples. Alongside this investigation, the mean log-normalised gene counts and peak intensities of metabolites that were assigned non-zero loadings in the sPCs were calculated and plotted. This test was performed to distinguish those genes and metabolites that are associated with the separation of samples relative to diapause stages and thus, potentially relevant to the physiological manifestation of diapause. Variables that were assigned non-zero loadings for more than one sPC, were plotted for the lowest numbered sPC only.

6.2.1.1. *Transcriptomics patterns revealed*

The sPCA scores plots and expression of each gene assigned a non-zero loading to the corresponding sPC are shown in Figure 6-2. SPC1 described 30% of the variance within the dataset (Table 6-2) and partitioned the embryo samples collected in pre-diapause (2 h and 10 h) and after (+48 h) the onset of diapause (Figure 6-2; A.1). Therefore, these results mirror the magnitude of the transcriptional changes that occur upon initiation of diapause, as I have already identified in Section 5. The genes assigned non-zero loadings for this sPC accordingly exhibited dramatic alterations in expression between 10 h and 48 h. Of these 114 genes (Table 6-2), 107 appeared in the list of genes described in Chapter 5 that had large, significant $\log_2$ fold changes at 48 h v 10 h and thus were not further investigated within this chapter. Furthermore, these genes display a fairly consistent level of expression in the time points at either side of this large step change during the initiation of diapause.

The second sPC (sPC2) separates single biological replicates of the 48 h and 82 h post-ovulation sampling groups. Upon investigation of the 17 genes assigned to this component, they were found to have particularly high expression only in these two samples (Figure 11-1), of these 17 genes 10 were of unassigned homology – reported in Table 11-1. I anticipate these likely represent erroneous results. Prior to this analysis, there was no suggestion that these samples were outliers in comparison with the rest of the biological samples, although this cannot be discounted as sPCA can be particularly biased towards to the presence of outliers (Erichson *et al.*, 2018).

Fourteen genes were assigned non-zero loadings in sPC3, which largely led to the separation of embryos collected 10 h after ovulation from the other biological samples (Figure 6-2; C.1).

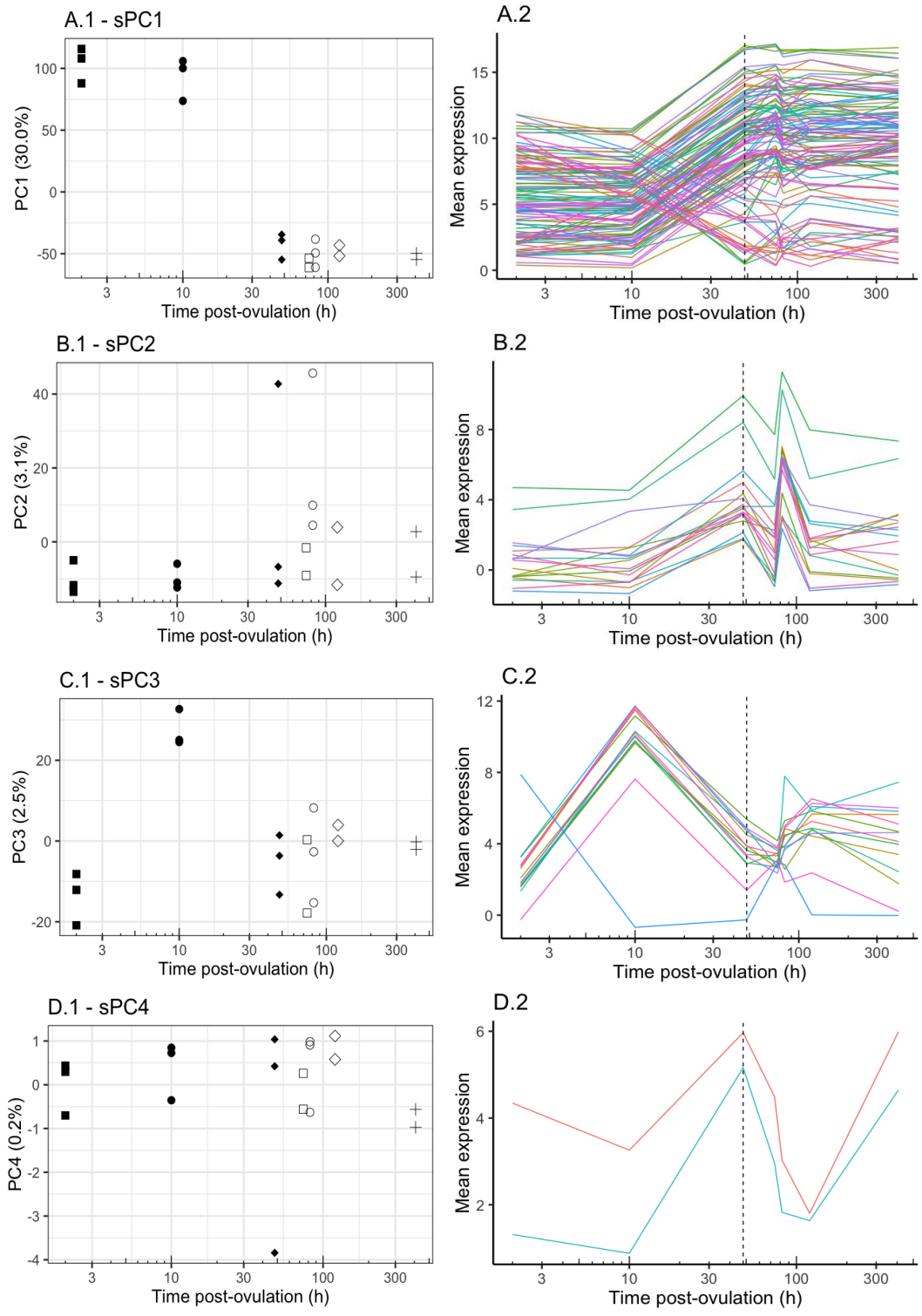


Figure 6-2 Sparse principal component analysis (sPCA) scores plots (1) and mean (log-normalised) expression of genes with non-zero loadings for associated sparse component (2) from sPCA of RNA-Seq data from sexually-produced *Daphnia magna* embryos collected between 2 h and 408 h (2 weeks) post-ovulation. Sparse components 1–4 (A–D) are shown (refer to Figure 6-1 legend for point characters). For expression plots, coloured lines indicate individual genes (mean expression at each time point calculated across 2-3 biological replicates), vertical dashed lines indicate approximate onset of developmental arrest (48 h).

A heatmap of the log-normalised counts of these 14 genes is shown in Figure 6-3. For comparison, the log-normalised counts in subitaneous embryos are also shown, although these were not used in the sPCA model. A single gene (Dapma7bEVm022793) containing a putative DNA binding fork head domain was highly expressed at 2 h with reduced expression in the remaining time points during diapause. However, a comparable level of expression was also present in asexually produced embryos at 2 h, suggesting that this gene expression may be unconnected to the diapause phenotype. The other 13 genes displayed a high level of expression at 10 h only, with low to medium expression in all other embryos collected. There were no obvious patterns of expression of these genes within asexual samples, suggesting that expression of these genes is specific and therefore relevant to diapause development.

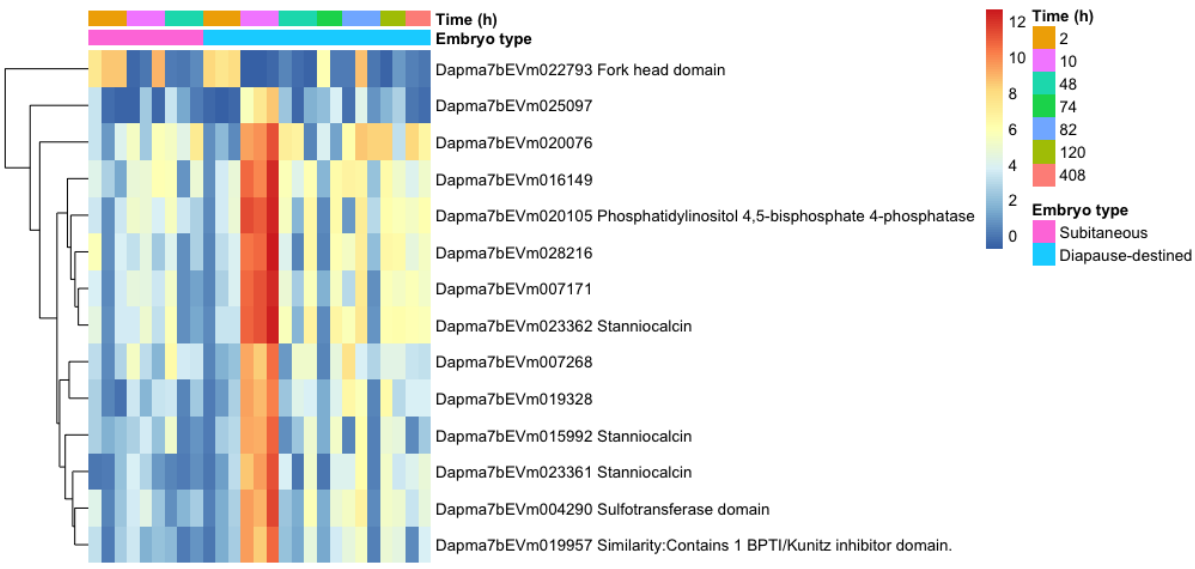


Figure 6-3 Heat map and hierarchical clustering of expression of genes with non-zero loadings in the third sparse principal component from the sparse principal component analysis of log-normalised gene counts in *Daphnia magna* sexually produced embryos collected between 2 h and 408 h post-ovulation. Log-normalised counts in subitaneous embryos are shown for comparison only. Row labels indicate gene name with corresponding annotation based on OrthoDB orthologous gene family group and coloured column top panels indicate embryo type and time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale.

Half of these 14 genes have no known homology to OrthoDB metazoan gene families (Table 11-1). However, 3 genes were identified as putative Stanniocalcins. Stanniocalcins are glycoprotein hormones, involved in calcium and phosphate homeostasis (Chang *et al.*, 2005). Since the initial isolation of stanniocalcin 1 in fish and humans, putative stanniocalcin homologs have been identified in several eukaryotic taxonomic groups including vertebrates, protostomes and sponges (Roch and Sherwood, 2011). There are two known stanniocalcin isoforms in humans and fish, and two additional paralogs have been identified in the teleost fish *Tetraodon nigroviridis* that display tissue specific localisation, although the authors believe this duplication is specific to the teleost lineage (Schein *et al.*, 2012). Our investigations revealed no previous descriptions of stanniocalcin genes in *Daphnia sp.* therefore the annotation of three putative stanniocalcin genes in this species will require further validation. Furthermore, there have been no previous studies suggesting a role of stanniocalcin(s) in diapause development. However, they have been shown to be involved in a wide range of additional biological processes including; murine osteogenesis (Yoshiko, Aubin and Maeda, 2002), in addition to various human cancers (Miyazaki *et al.*, 2014; Dai *et al.*, 2016). There is also evidence from transgenic mouse models to suggest that stanniocalcin 2 is a negative regulator of post-natal growth although the mechanism by which this occurs has yet to be elucidated (Gagliardi *et al.*, 2005; Chang *et al.*, 2008). The specific over-expression of putative stanniocalcin genes prior to the onset of development arrest at 10 h post-ovulation, in addition to the previously implied roles as a growth inhibitor in other organisms suggests that further investigation into the role of stanniocalcin in diapause development is worthy of further study.

The two genes with non-zero loadings in sPC4 appear to also represent anomalous results as in scores plots (Figure 6-2; D.1), only a single replicate of the 48 h post-ovulation group separated out, additionally inspection of the count matrices for these genes revealed particularly high expression of the two genes in this sample only and therefore, these genes were not further investigated.

6.2.1.2. *Lipidomics patterns revealed*

The sPCA of log-normalised peak intensities obtained from negative ion mass spectrometry analysis of non-polar embryo lipid extracts assigned 92 peaks with non-zero loadings to sPCs 1–4 at the chosen sparsity level (Section 2.15.2). Eighteen (5%) of these peaks were detected in diapause-destined embryo samples only (indicated in bold in Table 11-2). The first two sPCs largely appeared to describe variation within biological replicates, as exemplified by the incoherent distribution of sample scores (individual points) shown in Figure 6-4 (A.1 and B.1) and a lack of clustering of samples from the same biological group (i.e. time of collection post-ovulation). A peak (*m/*z = 1145.591560) within sPC2 showed a large increase in intensity at 120 h post-ovulation (shown in pink in Figure 6-4; B.1). However, upon closer inspection, this appeared to be due to an inflation of the mean value among replicates by the intensity value of a single technical replicate (Figure 11-2). There was a separation of biological samples with reference to time post-ovulation along sPC3 (Figure 6-4; C.1), with a gradual decrease in mean sample score from 2 h post-ovulation until 408 h. There was no appreciable separation between the 408 h and 17,592 h groups. Unfortunately, none of 10 peaks with non-zero loadings for this component were annotated with a putative compound identity. One peak

(*m/*z = 331.1890899) was assigned 24 possible molecular formulae, but none were assigned an identity (Table 11-2). Of the 10 peaks, 6 showed a general trend with time post-ovulation towards increasing intensity (Figure 6-5; A) and 4 peaks showed a decreasing intensity (Figure 6-5; B), indicating that peaks associated with this sPC describe alterations to the lipidome occurring during the early (first two weeks) diapause period.

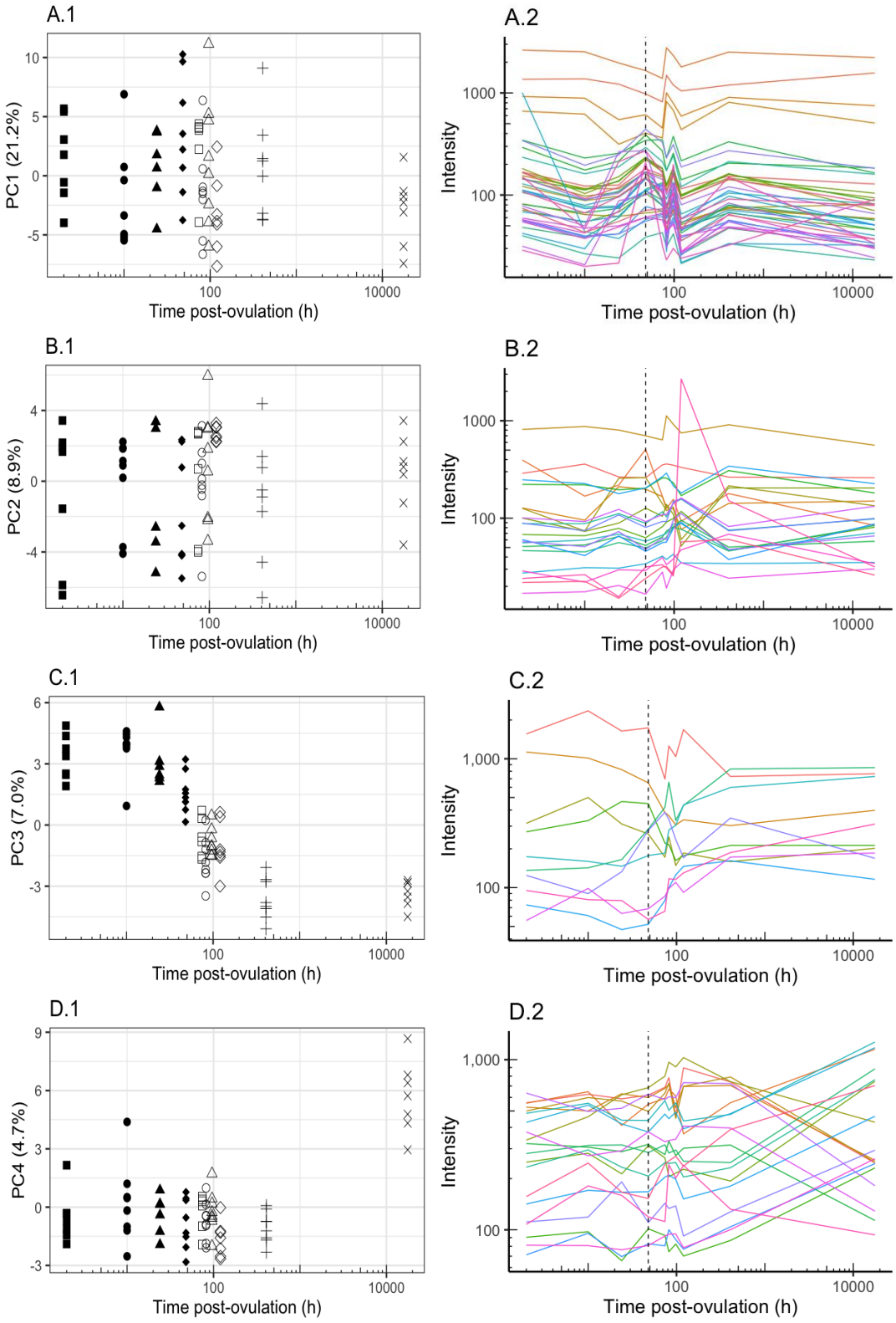


Figure 6-4 Scores plots (1) and mean log-normalised intensity of peaks with non-zero loadings for associated component (2) from sparse principal component analysis of log-normalised peak intensity from negative ion mode metabolomics analysis of non-polar extracts of sexual *Daphnia magna* embryos collected between 2 h and 2 y (17592 h) post-ovulation. Sparse principal components 1 (A) through to 4 (D) are shown. Vertical dashed line indicates approximate onset of developmental arrest (48 h). Coloured lines indicate mean intensity of individual peaks with time post ovulation.

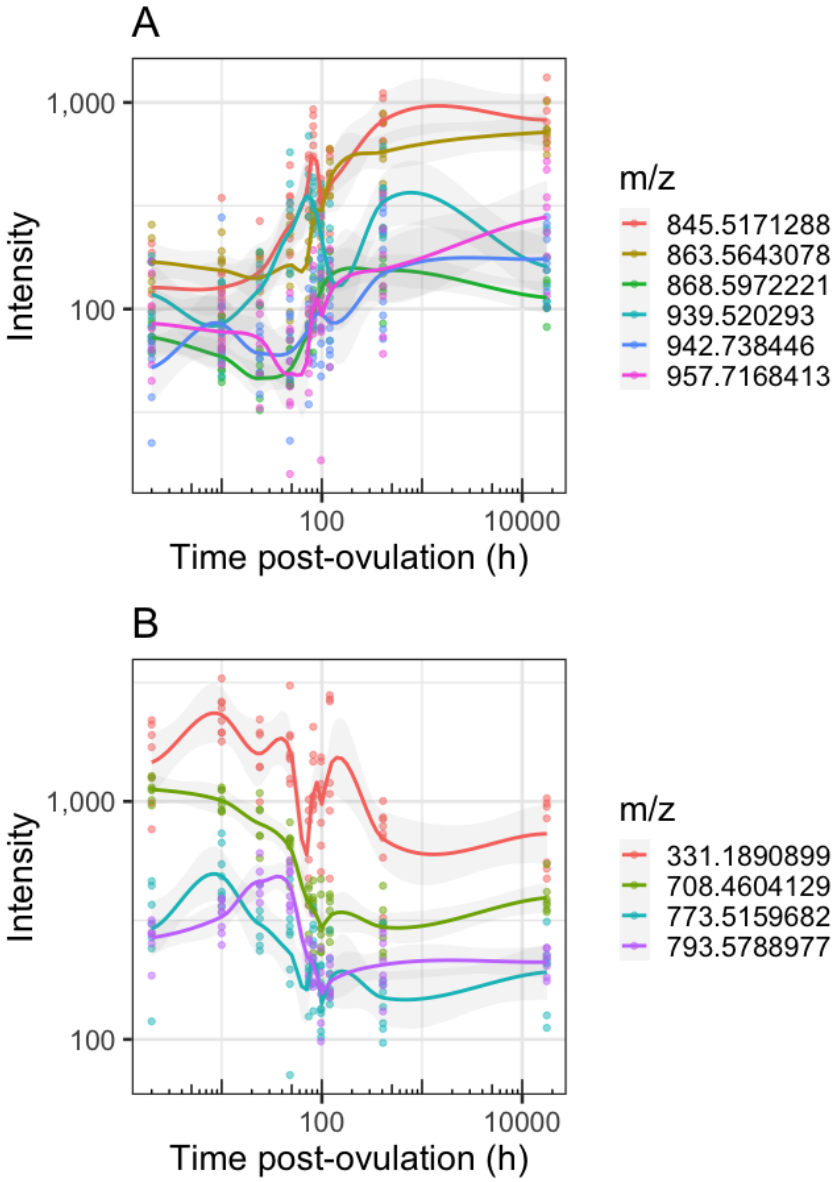


Figure 6-5 Intensity of peaks with a non-zero loading in the third sparse principal component identified through sparse principal component analysis of log-normalised peak intensity matrix obtaining follow the negative ion metabolomics analysis of non-polar *Daphnia magna* embryo extracts collected between 2 h and 2 y (17592 h) post-ovulation that displayed a general increase (A) and decrease (B) in intensity with time post-ovulation. Individual points represent biological replicates. Lines are fitted with locally weighted scatterplot smoothing (LOESS; span = 0.3). Grey shading indicates model 95% confidence interval.

There was a distinct separation of the 2 y (17,592 h) group from other biological samples along sPC4 (Figure 6-4; D. 1) and investigation of the peak intensities for those with non-zero loadings (Figure 6-4; D. 2) revealed significant changes in intensity between 408 h and 17,592 h for 13 of these 20 peaks (Table 6-3).

Table 6-3 Table of metabolite peak intensities detected during the negative ion metabolomics analysis of non-polar *Daphnia magna* sexual embryos collected at 408 h and 17592 h (2 y) post-ovulation. FC = fold change, calculated relative to 2 y. Peak *m/z* values indicated in bold denote peaks specific to sexually produced embryos. Peaks with significant (false discovery rate (FDR) adjusted *p*-value < 0.05) changes in intensity between the two time points are also indicated in bold and were determined based on a Student's t-test. Green shading denotes peaks assigned at least one putative annotation.

	Mean intensity					
m/z	408 h	17592 h	FC	$\log_2$ FC	*p*	*p*-adjust
896.6370885	722.75	181.43	0.25	-1.99	4.35E-06	**0.0006**
514.0618621	791.34	249.56	0.32	-1.66	0.008	**0.0461**
897.6408854	396.76	128.65	0.32	-1.62	5.07E-07	**0.0004**
313.0583743	739.84	247.26	0.33	-1.58	0.009	**0.0490**
872.6382304	314.95	113.47	0.36	-1.47	4.45E-06	**0.0006**
357.0729023	705.46	256.03	0.36	-1.46	0.002	**0.0206**
720.4549592	742.95	429.06	0.58	-0.79	0.002	**0.0189**
912.4849181	132.20	93.34	0.71	-0.50	0.336	0.4858
938.7074435	388.79	704.05	1.81	0.86	0.019	0.0747
327.1884086	558.45	1151.49	2.06	1.04	0.027	0.0913
889.6860285	127.65	293.71	2.30	1.20	0.057	0.1455
887.6795157	104.15	245.15	2.35	1.23	0.006	**0.0367**
884.6603481	481.70	1176.51	2.44	1.29	0.008	**0.0458**
887.6702189	180.65	464.92	2.57	1.36	0.018	0.0723
910.6750613	100.15	261.73	2.61	1.39	0.046	0.1284
883.6570781	476.66	1270.28	2.66	1.41	0.020	0.0759
868.5700342	86.57	231.10	2.67	1.42	0.007	**0.0409**
882.644765	231.67	762.65	3.29	1.72	0.005	**0.0321**
881.6417022	248.61	883.34	3.55	1.83	0.009	**0.0485**
812.5000667	193.51	746.16	3.86	1.95	0.001	**0.0140**

A total of 6 peaks displayed significant (Table 6-3) increases in intensity between 408 h and 17,592 h, and 7 peaks exhibited a decrease in peak intensity during this time, shown in Figure 6-6; A and B respectively. Three of these peaks were assigned a putative compound annotation, however there were in excess of 20 possible formulae for each peak (Table 11-2). These putative annotations were largely phosphatidylcholines (PCs) and phosphatidylethanolamines (PEs), which due to their isomeric nature are largely impossible to resolve accurately via direct-infusion mass spectrometry without additional analytical approaches. Furthermore, a diverse

number of PE and PC isomers, varying in fatty acyl chain length, position and, degree of saturation have been identified in *D. magna* thus each annotation must be considered equally likely (Stanley *et al.*, 2013).

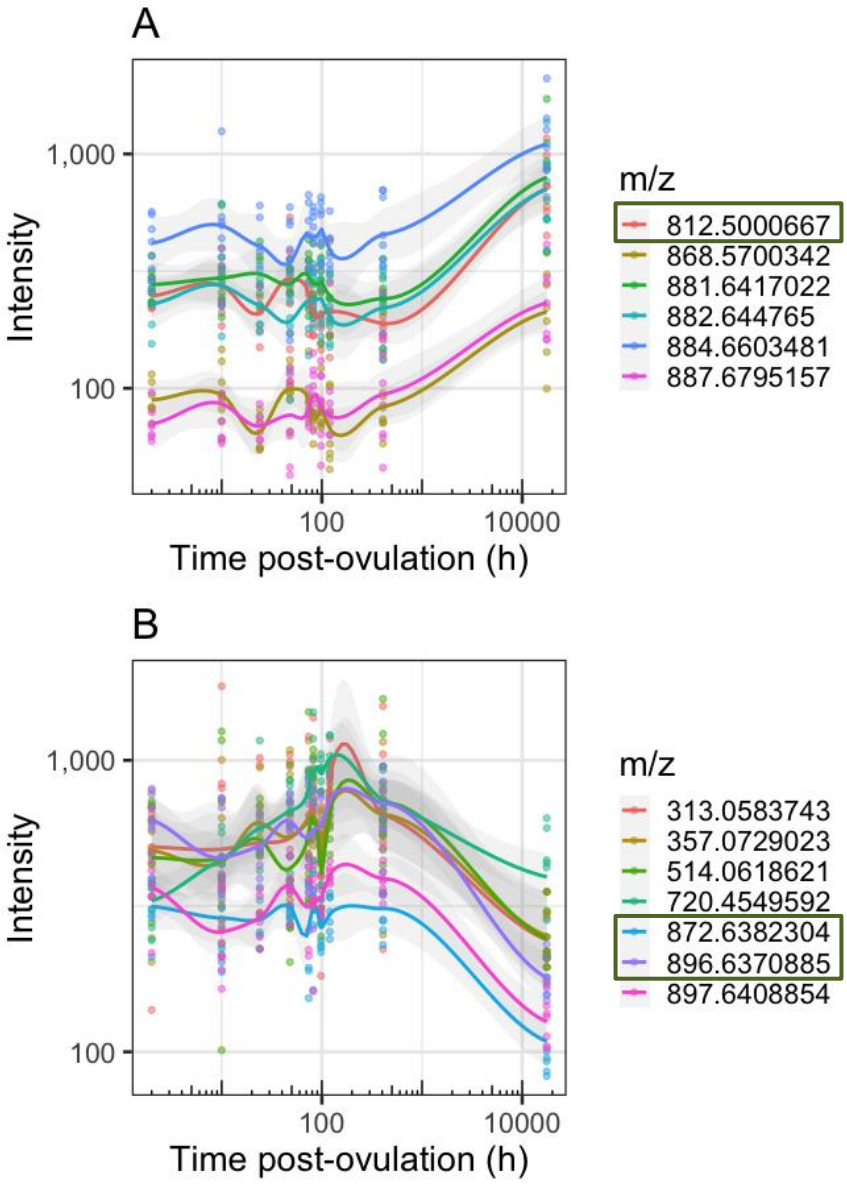


Figure 6-6 Intensity of peaks with a non-zero loading in the fourth sparse principal component identified through sparse principal component analysis of log-normalised peak intensity matrix obtaining follow the negative ion metabolomics analysis of non-polar *Daphnia magna* embryo extracts collected between 2 h and 2 y (17592 h) post-ovulation that displayed a significant increase (A) and decrease (B) in intensity between 408 h and 17592 h post-ovulation. Peaks with putatively assigned compound identities are shown in legend boxes. Individual points represent biological replicates. Lines are fitted with locally weighted scatterplot smoothing (LOESS; span = 0.3). Grey shading indicates model 95% confidence interval.

When animals do not feed for the duration of their dormancy, diapause is expected to be metabolically expensive and create starvation stress associated with a prolonged period without nutrient input (Hahn and Denlinger, 2007). Thus, increased energy reserves and reduced metabolism are physiological factors anticipated to be used in diapausing individuals to mitigate such stress (Sinclair and Marshall, 2018). Calorie dense lipids are an optimal energy store hypothesised to be important for diapause maintenance (Hahn and Denlinger, 2011). For example, prepupal diapausing *Asobara tabida* exhibit a gradual decline in total fat content in samples collected over 9 months (Ellers and Van Alphen, 2002). Considering the adult diapause of *Culex pipiens*, authors reported differences in genes related to fatty acid metabolism between early diapause (2 weeks) and late diapause (2 months) individuals (Sim and Denlinger, 2009). This included increased expression of genes related to fatty acid synthesis in early diapause whilst in late diapause those relating to β-oxidation were over-expressed, suggesting a shift from storage to utilisation with increasing time in diapause (Sim and Denlinger, 2009). However, in both cases, fold changes in expression were determined relative to non-diapause individuals rather than to successive diapause stages. Furthermore, in a lipidomic study of *Pieris napi* pupal diapause, the authors reported a lack of observable changes in lipid class and species composition, and fatty acid profiles between diapausing pupae at 6 d and 144 d (Lehmann *et al.*, 2018). But the authors did observe significant changes in all three upon termination of diapause, leading them to hypothesise that lipids largely provide necessary resources for post-diapause development (Lehmann *et al.*, 2018). Based on the data I obtained, it appears that only minor changes in non-polar *D. magna* metabolites occur during diapause development as meaningful separation

between sampling time points was only apparent when observing lower sPCs (i.e. sPC3 and sPC4), which account for small portions of the variance (Table 6-2). However, two distinct trajectories were observed within peaks assigned non-zero loadings to these sPCs. First, peaks with non-zero loadings in sPC3 displayed increases or decreases in intensity over the time points sampled during “early” diapause. Second, an additional contingent of peaks, those with non-zero loadings in sPC4, demonstrated significant increases or decreases in intensity at 2 y versus 408 h in diapause. Indicating a subset of non-polar metabolites that change in relation to extended diapause. At this stage it is unclear whether the changes to these lipids over long term diapause represents their possible utilisation as an energy source and will require additional investigation. However, an additional consideration is that the phospholipid composition of *D. magna* has previously been shown to alter in response to a reduction in environmental temperature, leading to an increase in PE content, which the authors had postulated may contribute to the maintenance of membrane fluidity at reduced temperatures (Farkas, Nemecz and Csengeri, 1984). In this study, all embryos sampled up to 408 h had been maintained at a constant 20°C prior to collection. However, the 17,592 h sample had been stored at 4°C in darkness for the 2 y period between ephippia production and embryo collection. Thus, the additional effect of environmental temperature cannot be discounted.

6.2.1.3. *Polar metabolites revealed*

The same sPCA approach was also applied to metabolomic data obtained from polar embryo lipid extracts. The results for negative ion analysis are shown in Figure 6-7 and the results for positive ion analysis are shown in Figure 6-8. The 114 peaks that were selected in negative ion analysis and assigned non-zero loadings included 76 peaks (over 60%) that were specific to sexually produced samples (Table 11-3). By contrast, 29 peaks (35%) out of the 84 peaks selected in positive ion analysis and assigned non-zero loadings to at least one sPC were specific to sexually produced samples (Table 11-4).

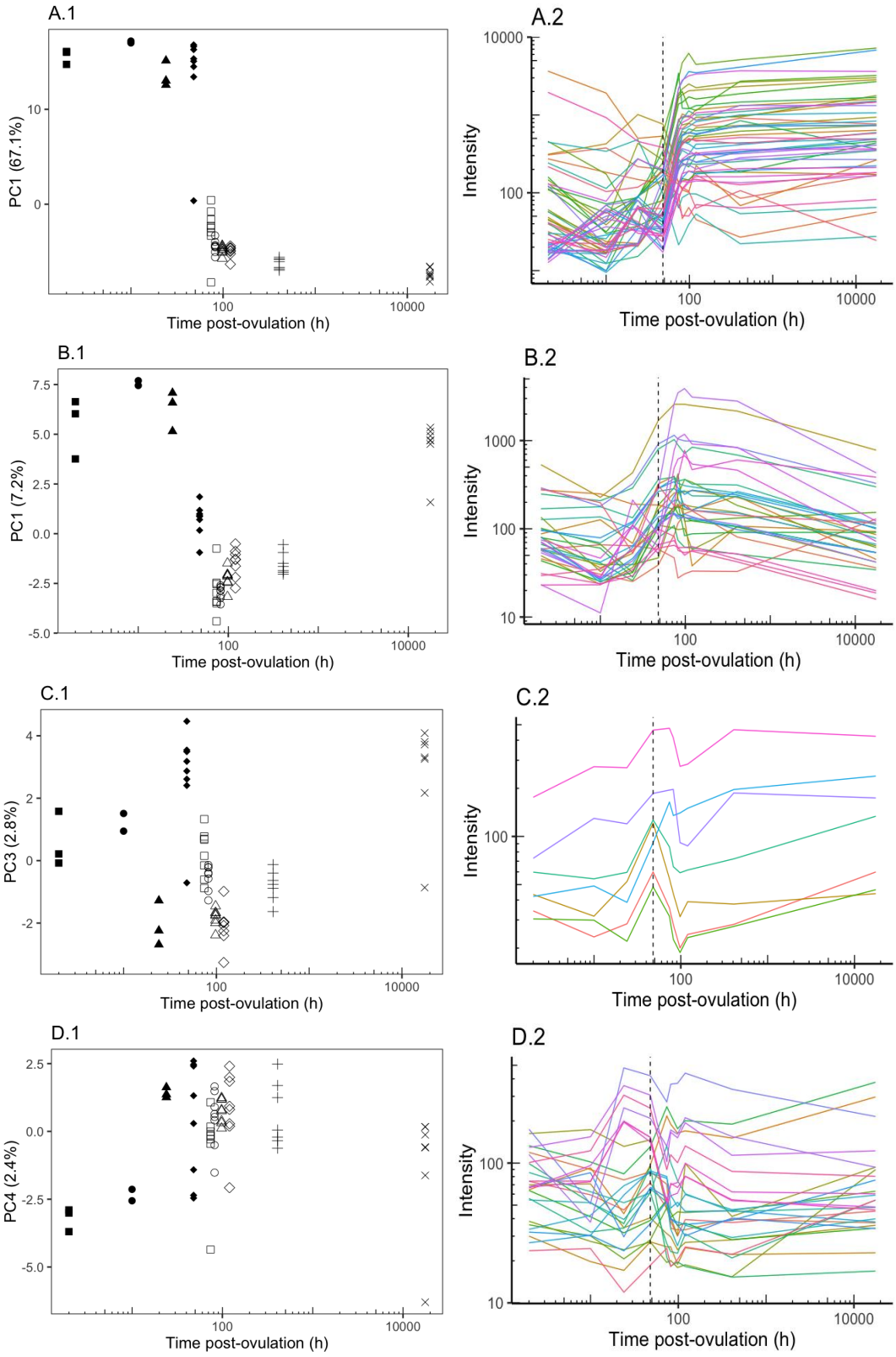


Figure 6-7 Scores plots (1) and mean log-normalised intensity of peaks with non-zero loadings for associated component (2) from sparse principal component analysis of log-normalised peak intensity from negative ion mode analysis of polar extracts of sexual *Daphnia magna* embryos collected between 2 h and 2 y (17592 h) post-ovulation. Sparse principal components 1 (A) through to 4 (D) are shown. Vertical dashed line indicates approximate onset of developmental arrest (48 h). Coloured lines indicate mean intensity of individual peaks.

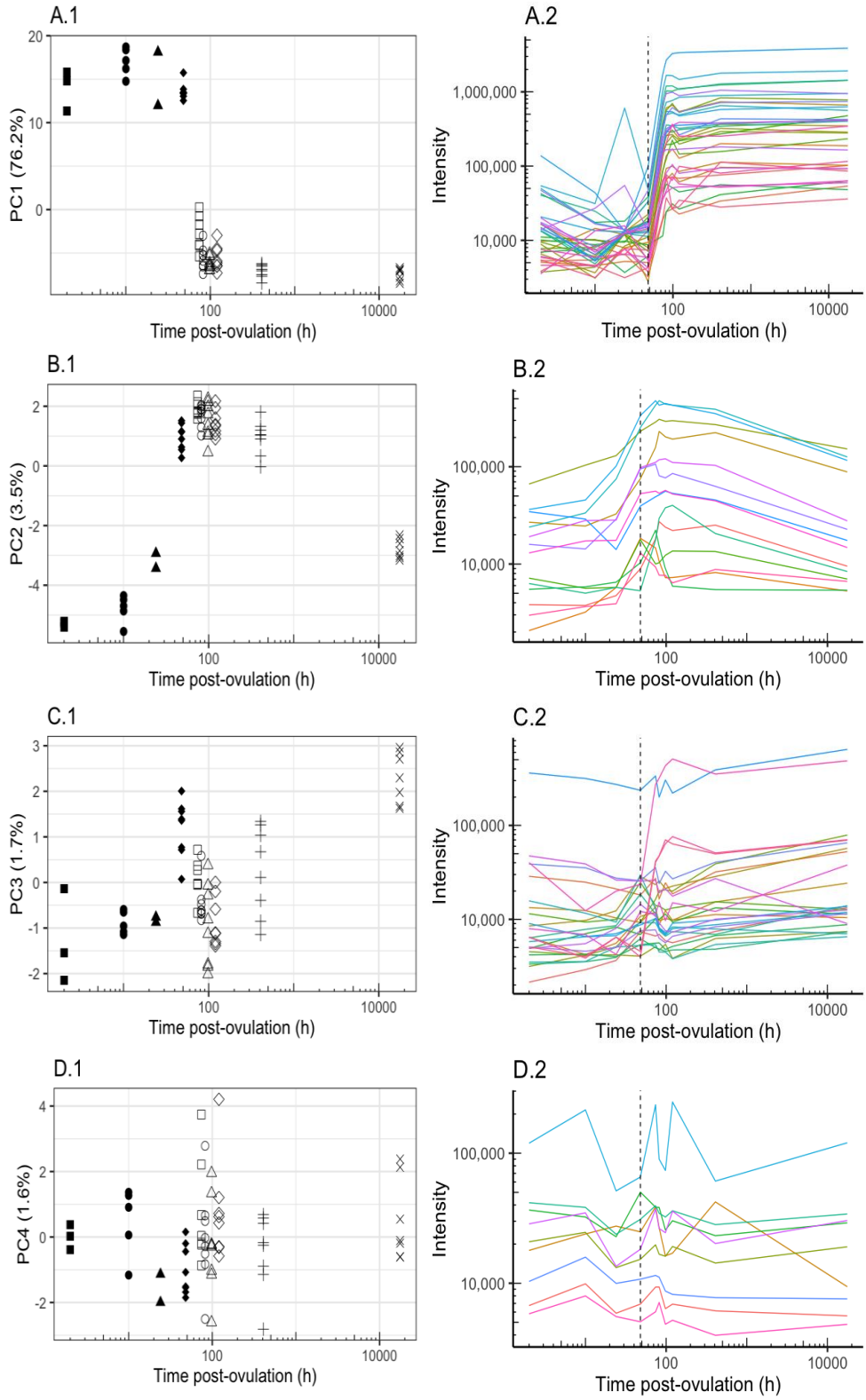


Figure 6-8 Scores plots (1) and mean log-normalised intensity of peaks with non-zero loadings for associated component (2) from sparse principal component analysis of log-normalised peak intensity from positive ion mode analysis of polar extracts of sexual *Daphnia magna* embryos collected between 2 h and 2 y (17592 h) post-ovulation. Sparse principal components 1 (A) through to 4 (D) are shown. Vertical dashed line indicates approximate onset of developmental arrest (48 h). Coloured lines indicate mean intensity of individual peaks.

In both negative ion (Figure 6-7;A.2) and positive ion (Figure 6-8;A.2) polar analyses, the peaks assigned to non-zero loadings in the first sPC demonstrated predominately the large positive fold changes in peak intensity between 48 h and 74 h, as previously identified and described in Chapter 5. In the negative ion dataset, of the 49 peaks assigned non-zero loadings in sPC1 (Table 6-2), a total of 39 peaks demonstrated $\log_2$ fold changes >2 at 74 h v 48 h. Half of these peaks (24) were statistically significant following a Student's t-test and FDR correction (Table 6-4). Within the positive ion dataset, sPC1 also largely depicts the large fold changes discussed in the previous chapter; of the 36 peaks with non-zero loadings for this component, a total of 34 displayed a large (> 2), positive $\log_2$ fold change in intensity at 74 h v 48 h, of which 27 were significant (FDR-adjusted *p*-value less than 0.01; Table 6-4). The two peaks previously discussed and putatively annotated as trehalose (positive ion; Table 11-4) and UDP-glucose (negative ion; Table 11-3) were also present within these peak lists.

Table 6-4 Table showing the number of peaks detected in polar metabolomic analysis with non-zero loadings for the first and second sparse principal components (sPC) that demonstrated significant differences in intensity between time points. Significance was determined by a Student's t-test and false discovery rate-adjusted *p*-value of less than 0.01.

			Time points compared (h)		
	sPC	**Total**	**48 v 24**	**74 v 48**	**2 w v 2 y**
Negative ion	**1**	49	0	24	4
	2	32	15	6	22
Positive ion	**1**	36	4	27	3
	2	14	6	5	9

A single peak (*m/*z = 370.1933764) displayed a high mean intensity in the positive ion dataset at 24 h. However, upon further inspection (Figure 11-3), this may be due

to a disparity between the biological replicates at this time point and was thus not considered significant. There were additionally some peaks identified in negative ion that had negative fold changes between 48 h and 74 h that were assigned non-zero sPC loadings for sPC1. Upon investigation of the intensity of these peaks throughout the duration of diapause sampled, all eight peaks displayed a general trend of a reduction in intensity with time post-ovulation (Figure 6-9). Particularly, a steep reduction in intensity between samples collected at 48 h and those at 74 h, although only two of these fold changes were statistically significant (Table 6-4).

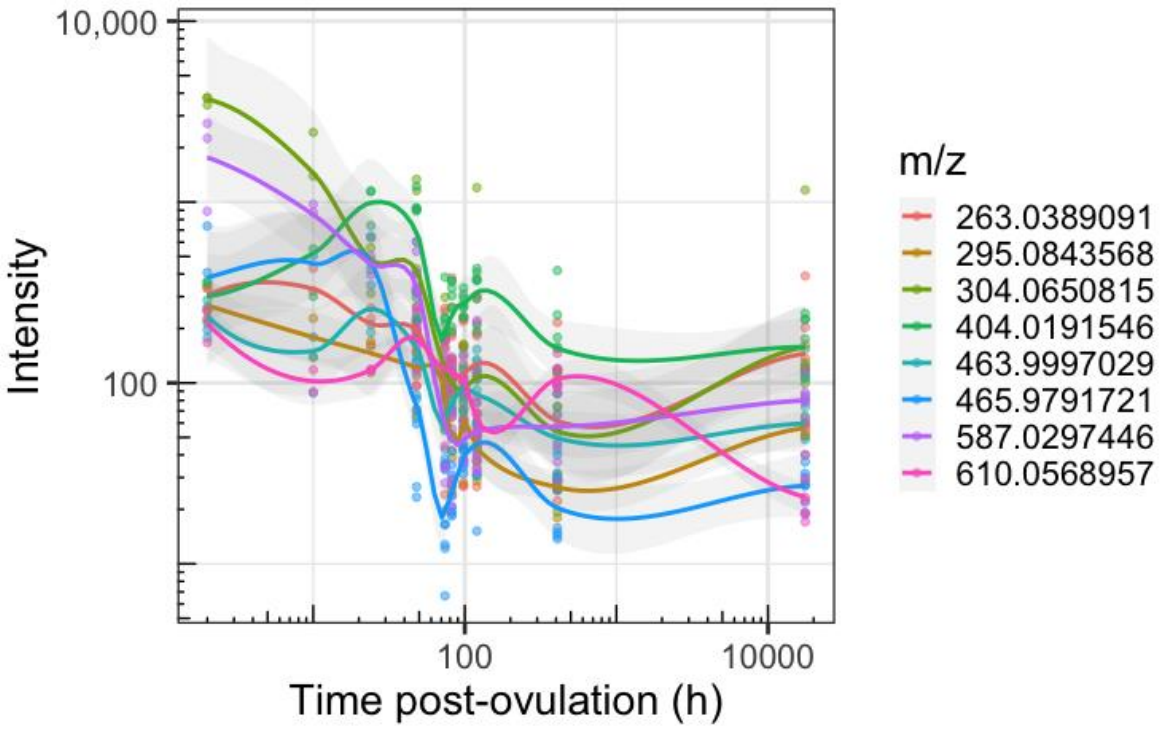


Figure 6-9 Intensity of peaks with a non-zero loadings in the first sparse principal component identified through sparse principal component analysis of log-normalised peak intensity matrix obtaining follow the negative ion analysis of polar *Daphnia magna* embryo extracts collected between 2 h and 2 y (17592 h) (post-ovulation that displayed a general decrease in intensity. Individual points represent biological replicates. Lines are fitted with locally weighted scatterplot smoothing (LOESS; span = 0.3). Grey shading indicates the modelled 95% confidence intervals.

Investigation of sPC2 in both polar analyses (B.1; Figure 6-7 (negative ion) and Figure 6-8 (positive ion)) revealed three distinct clusters of samples: sampled collected during pre-diapause development (2–24 h), embryos in early developmental arrest (from 48–408 h), and embryos in late diapause (2 y). Visualisation of the mean intensity of peaks with non-zero loadings for sPC2 (B.2; Figure 6-7 (negative ion) and Figure 6-8 (positive ion)) reveals an increase in intensity in relation to the initiation of diapause (occurring at 48 h). Crucially, the behaviour of these peaks is different to that of the peaks in sPC1 in two ways. First, this intensity increase begins earlier with just under half of the peaks with non-zero loadings in this component in both negative and positive ion analysis having significant positive fold changes between 24 h and 48 h (Table 6-4). Second, although the majority of peaks reached an intensity maximum at 74 h and maintained a similar intensity in samples collected in early diapause (74–408 h post-ovulation), their intensity is markedly lower in embryos collected after 2y in diapause. The putatively annotated glutathione (*m/z* = 328.05840) and oxidised glutathione (*m/z* = 611.14466) peaks discussed in Chapter 4, were also amongst the peaks with non-zero loadings for the negative ion dataset (Table 11-3). There were also several peaks with particularly high intensities at the initiation of diapause, which are plotted in Figure 6-10. The two peaks from negative ion analysis (Figure 6-10; A) were unfortunately not annotated but both demonstrated large negative $\log_2$ fold changes at 17592 h v 408 h (m/z = 283.08827, 2.8 fold; m/z = 473.59224, 6.5-fold). In the positive ion dataset, one of the peaks (*m/z* = 284.09943) was assigned 5 possible molecular formulae. One of these is associated with the putative compound identity of either guanosine or 8-hydroxy-deoxyguanosine ([M+H$^+$] adduct). The intensity of

this peak throughout the time course of diapause investigated is plotted in Figure 6-10; B. The intensity of an additional peak that is also putatively annotated as a guanosine or 8-hydroxy-deoxyguanosine metabolite but different adduct ([M+K^+]) was also assigned a non-zero loading to sPC2 and displayed a similar intensity trajectory through the course of diapause sampled. Both putative guanosine peaks were also detected in the dataset containing subitaneous embryos. As no later points were collected for subitaneous embryos, I was unable to compare the full profiles of this peak between the two embryo types. However, there were no significant differences (t (10) = -0.978, p = 0.351) in intensity of the peak m/z = 284.09943 in pre-diapause embryos at 24 h (87962; SD = 16361) versus at 10 h in subitaneous embryos (96015; SD = 13084), representing embryos from the two developmental trajectories at a comparable number of cells ca. 1000 cells. Suggesting therefore, this pattern of guanosine metabolite levels may not be specific to diapause development.

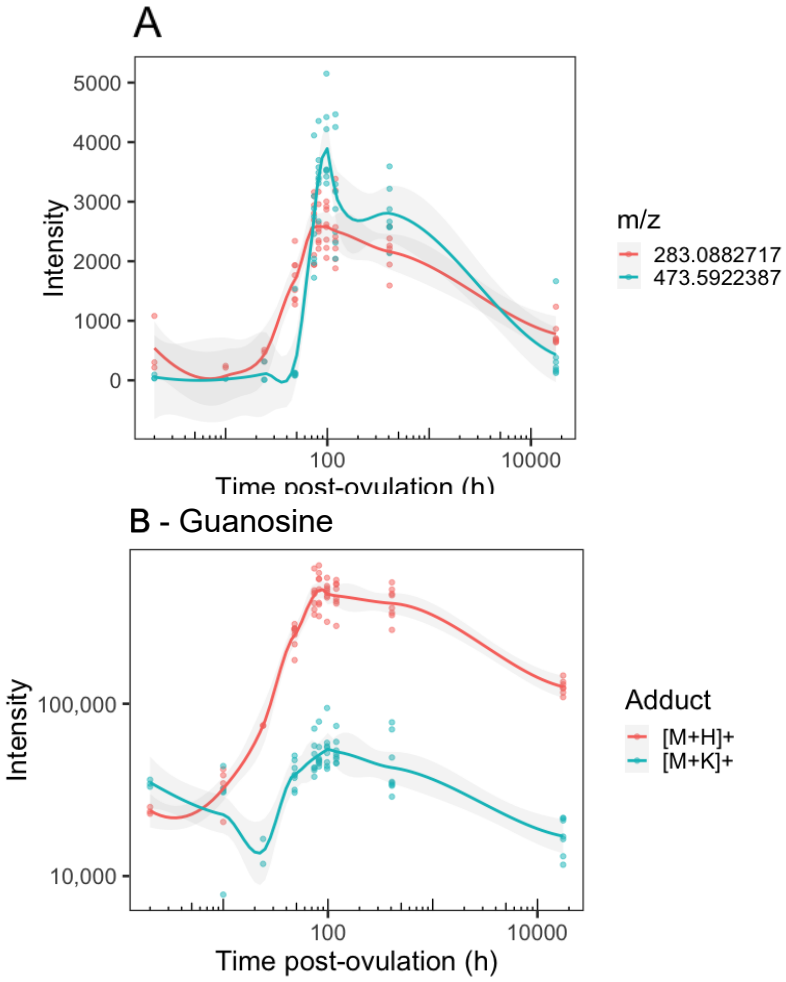


Figure 6-10 Intensity of peaks with a non-zero loadings in the second sparse principal component that demonstrated a prominent spike in intensity at 74 h post-ovulation. These are identified through sparse principal component analysis of log-normalised peak intensity matrix obtaining from the negative ion (A) and positive ion (B) metabolomics analysis of polar *Daphnia magna* embryo extracts collected between 2 h and 2 y (17592 h) post-ovulation. Additional titles indicate putative peak annotations. Individual points represent biological replicates. Lines are fitted with locally weighted scatterplot smoothing (LOESS; span = 0.3). Grey shading indicates the modelled 95% confidence intervals.

The third and fourth sPCs in both polar analyses described only a small percentage of the variance within each dataset, measured to be less than 3% (Table 6-2). No specific separation of sample classes along sPC4 in relation to time sampled in diapause development was observed (D.1; Figure 6-7 (negative ion) and Figure 6-8 (positive ion)). These were therefore not investigated further. However, in both analyses, sPC3 gave rise to the separation of particular time points from the rest of

the embryos sampled (C.1; Figure 6-7 (negative ion) and Figure 6-8 (positive ion)). Thus, peaks assigned non-zero loadings to these components tended to demonstrate specific increases or decreases in intensity at the aforementioned time points. For example, 7 peaks were assigned non-zero loadings to the third sPC in the analysis of negative ion data (Table 6-2). Embryos collected after 2 y separated out into a distinct cluster away from other samples along this component (Figure 6-7; C.1). Additionally, the 48 h time point also separated out from the other samples. A distribution of samples along sPC3 is also observed for the positive ion dataset shown in Figure 6-8; C.1. Whilst no peaks with non-zero loadings were successfully annotated from the negative ion analysis, two peaks from the positive ion dataset were assigned putative compound identities associated with postulated formulae. Both peaks were, amongst other possibilities, assigned formulae and corresponding putative dipeptide identities (Table 11-4). The first peak (*m/z* = 269.16144) was identified as a dipeptide formed of histidine and leucine/isoleucine. The second peak (*m/z* = 276.15594) was identified as a glutamate and lysine dipeptide, although for both peaks, several structural configurations are possible that will require further investigation to elucidate. Both peaks display an interesting profile (Figure 6-11) that spikes at 48 h following the initiation of diapause, dropping over the immediate subsequent time points and increasing steadily again at 2 w and 2 y during late diapause.

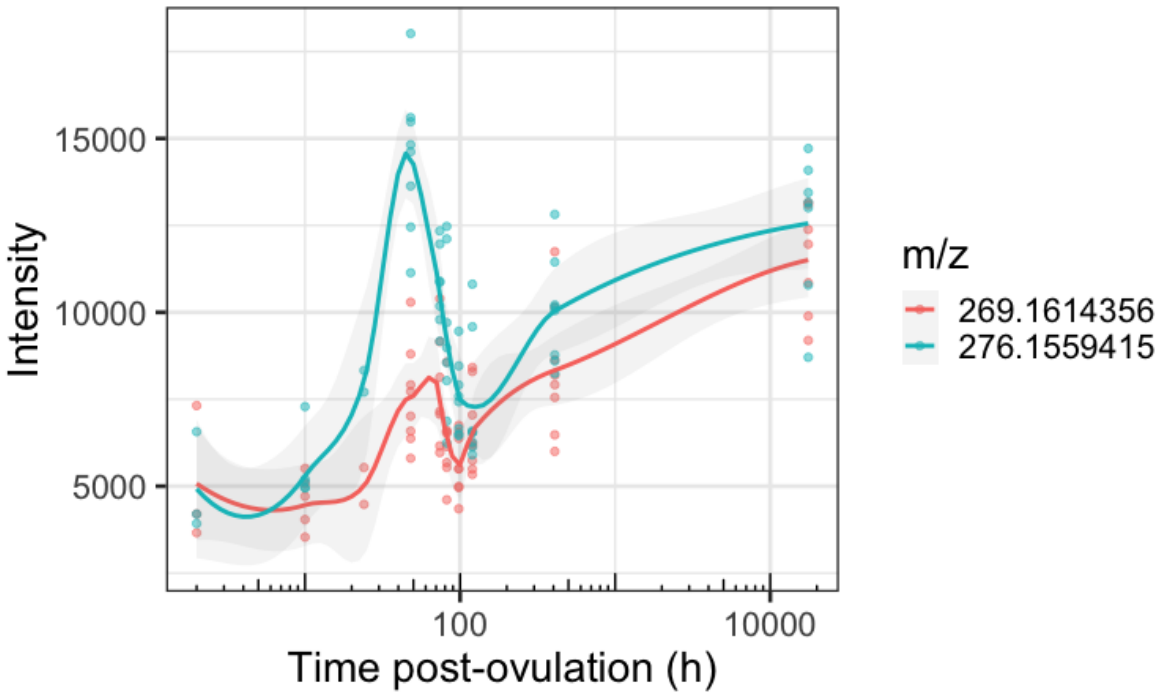


Figure 6-11 Intensity of putatively annotated dipeptide peaks with a non-zero loadings in the third sparse principal component that demonstrated a prominent spike in intensity at 48 h post-ovulation. These are identified through sparse principal component analysis of log-normalised peak intensity matrix obtaining from the positive ion analysis of polar *Daphnia magna* embryo extracts collected between 2 h and 17592 h (2 y) post-ovulation. Individual points represent biological replicates. Lines are fitted with locally weighted scatterplot smoothing (LOESS; span = 0.3). Grey shading indicates the modelled 95% confidence intervals.

Dipeptides contain, in various configurations, two amino acid molecules and are intermediaries in protein degradation (Thamotharan *et al.*, 1997). Typically, the balance between protein synthesis and proteolysis is carefully regulated to ensure that cellular integrity is maintained (Lecker, 2006). Furthermore, the protein half-life is an effective means by which important cellular processes can be regulated (e.g. cell division is controlled through the targeted degradation of cyclins that drives progression through the eukaryotic cell cycle (King *et al.*, 1996)). Therefore, a spike in the quantity of cellular dipeptides at 48 h coinciding with the initiation of diapause suggests a significant perturbation in protein turnover associated with onset of developmental arrest may occur. This potential increase in dipeptides could

represent specific temporal degradation of proteins at the onset of developmental arrest and thus accumulation of dipeptide derivatives. This may be relevant to either the termination of cellular processes, resulting mechanistically in the initiation of dormancy, or reducing the complement of active proteins to conserve resources during dormancy. For example, investigating protein synthesis in diapausing embryos of the annual killifish *Austrofundulus limnaeus*, the authors reported a reduction in protein synthesis of over 90% in comparison to pre-diapause embryos (Podrabsky and Hand, 2000). These authors suggest that inhibition of protein synthesis drastically reduces organism energetic demands facilitating the profound metabolic suppression also observed during diapause in this species.

Alternatively, during the analysis of genes differentially expressed between diapausing and non-diapausing *D. magna* embryos, several overrepresented GO-annotations for down-regulated genes in sexually-produced embryos were related to protein translation and ribosomal processes (Figure 4-4). However, expression of these genes was comparatively higher in diapausing embryos at 48 h than in pre-diapause at 2 h and 10 h, which is suggestive of on-going capacity for protein synthesis during the diapause period. Additionally, expression of these genes is maintained at comparable levels throughout the later samples taken during the diapause period (Figure 11-4). In summary, these results (whilst inconclusive) suggest that there are alterations in protein synthesis occurring in relation to diapause in *D. magna*. However, additional study of diapause at the proteomic level is required and will likely be instrumental in deciphering changes that are mechanistic in nature from the consequential.

But even without certainty about the identities of these elements of the metabolome, observing the peaks assigned non-zero loadings to the first three sPCs in the sPCA of polar datasets revealed subsets of metabolite peaks with characteristic intensity profiles. The overwhelming majority of peaks assigned non-zero loadings to the first sPC in both polar analyses showed dramatic intensity changes in relation to the onset of diapause, then afterward maintained a consistent level of intensity for the duration of the remaining time points. Thus, the onset of diapause is associated with major changes in metabolite composition of the embryos at the time scales that I investigated. When considering other sPCs additional dynamic metabolite behaviours become apparent. Peaks assigned non-zero loadings to sPC2 tend to rise in intensity, peaking shortly after the onset of diapause (74 h), reduce in intensity through the later stages of diapause, particularly demonstrating significant fold changes at 2 y versus 408 h, indicating metabolic changes that are occurring in relation to extended dormancy. By contrast, peaks assigned non-zero loadings to sPC3 were specific to individual time points and tended to show specific spikes in intensity at the aforementioned time point, indicating that there are also metabolic changes localised to specific time intervals.

6.2.2. The molecular phenotype of diapause

The sPCA investigations revealed significant molecular changes associated with diapause, not only in relation to the onset of developmental arrest but also throughout the sampled time course. By observing the peaks assigned non-zero loadings, their behaviour builds a picture of complex molecular events and revealed sub-sets of features that increase in intensity or expression at only specific time

points. These time periods, where molecular changes are localised, likely represent formative phases relevant to the diapause phenotype. However, the nature of sPCA is that, by sparsifying the data, I am unable to determine without further interrogating the data how many other features demonstrate comparable behaviour that are not included in the previous plots as a result of a pre-specified sparsity level.

To quantify the extent and magnitude of metabolic and transcriptional changes occurring in diapausing embryos at specific points (T) within the period sampled (T = 2 h through to T = 2 y post-ovulation), multiple pairwise comparisons between successive time points were performed (as described previously in Section 2.15) by extending the analysis performed in Section 5 (which focussed specifically on the initiation of diapause) to now through the whole diapause trajectory. Beginning with the earliest sexually-produced embryos collected at 2 h post-ovulation and working successively, metabolite fold changes were calculated based on mean peak intensity (n = 2–8; Section 9.3.2) at the later time point (T_n) relative to the former time point (T_{n-1}) (e.g. T_n / T_{n-1}). Statistical significance was determined based on a two-tailed Student's t-test at $p < 0.01$ after Benjamini-Hochberg false discovery rate (FDR) correction (Benjamini and Hochberg, 1995). Separate differential gene expression analyses were performed comparing embryo samples at analogous time points i.e. 10 h v 2 h using *DESeq2* (Love, Huber and Anders, 2014), with the same *p*-adjust < 0.01 threshold. Additionally, the features that exhibited a significant change in level within a given sampling interval were assessed to determine whether they displayed a significant change at any other time interval or uniquely within the aforementioned developmental period (Figure 6-14). To further visualise the magnitude of the molecular changes occurring, volcano plots as $\log_2$ fold changes

against -$\log_{10}$ *p*-value (FDR-adjusted) were plotted in Figure 6-15. The results of both analyses showing the number of significantly changing features between time points (gene expression or metabolite intensity) are plotted in Figure 6-12 and are herein discussed in relation to the phases of diapause covered.

Finally, a quantitative measurement of observed metabolic features at each of the time points was made (Supplemental Figures 8.5–8.7) to complement my findings of expression level changes of features that are shared among time points thereby comprehensively assessing the overall molecular attributes of diapause. I begin by highlighting overall events that are observed at each stage.

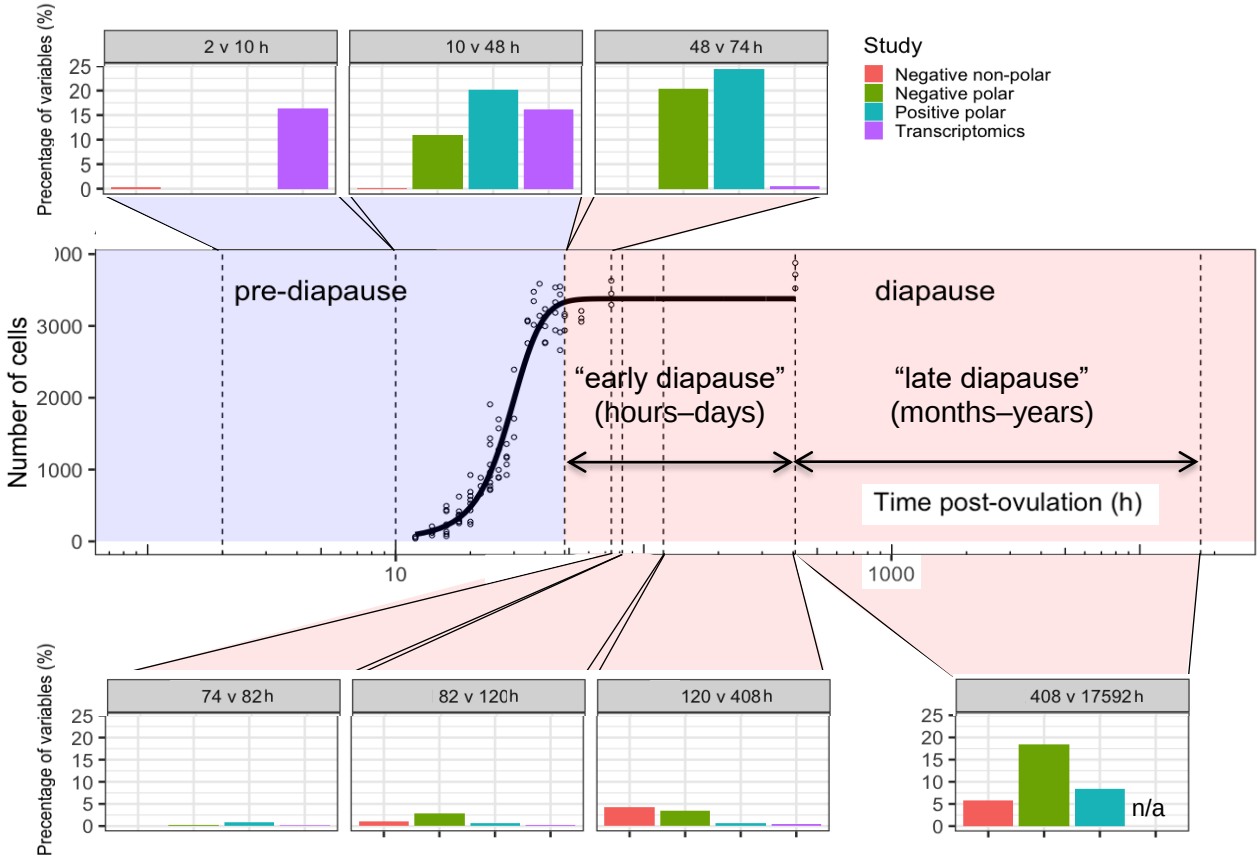


Figure 6-12 The molecular phenotype of diapause anchored by the rate of cell division through to developmental arrest. Main panel shows the stages of diapause in *Daphnia magna* as defined within this thesis based on the arrest of active development determined from the counting of DAPI-stained nuclei of sexually produced embryos at various time points post-ovulation. This encompasses pre-diapause (blue shading) wherein direct development proceeds, the initiation of diapause at the onset of developmental arrest (ca. 48 h) followed by a period of diapause (red shading). The duration of diapause investigated was further divided into "early" diapause covering the immediate days following initiation and "late" covering a longer period of 2 y, indicated by horizontal black arrows. Sampling time points are indicated in the grey annotation panel for each chart and by vertical dashed lines on the main growth curve panel. Bar charts indicate the percentage of the total number of detected features that show a significant change in magnitude between the time points indicated. Significantly changing features were determined by $p < 0.01$ (after Benjamini-Hochberg false discovery rate correction). Actual counts are given in Table 6.5.

Table 6-5 Summary of the number of features (i.e. genes or metabolites) with significant changes in expression (genes) or intensity (metabolites) from successive pairwise comparisons between *Daphnia magna* embryos sampled during diapause development. Significantly changing features were determined by $p < 0.01$ (after Benjamini-Hochberg false discovery rate correction) and are also plotted in Figure 6-12. Elements that were significant only within one time interval (plotted in Figure 6-14) and with large (> 2-fold) log_2 fold changes (FC; plotted in Figure 6-15) are also reported.

			Significant ($p < 0.01$)		Unique to time point		Large log_2 FC			
Study	Time point compared (h)	Total	No.	% of total	No.	% of significant	No.	% of significant	Up	Down
Transcriptomics	10 v 2	15393	2525	16.40	1894	75.01	1738	68.83	1217	521
	48 v 10	16790	2701	16.09	2055	76.08	2294	84.93	1215	1079
	74 v 48	15630	67	0.43	30	44.78	67	100.00	16	51
	82 v 74	14827	22	0.15	11	50.00	22	100.00	19	3
	120 v 82	15876	33	0.21	12	36.36	33	100.00	10	23
	408 v 120	15476	74	0.48	36	48.65	74	100.00	37	37
Negative ion non-polar metabolomics	10 v 2	805	2	0.25	1	50.00	0	0.00	0	0
	48 v 10	805	1	0.12	1	100.00	0	0.00	0	0
	74 v 48	805	0	0.00	0	0.00	0	0.00	0	0
	82 v 74	805	0	0.00	0	0.00	0	0.00	0	0
	120 v 82	805	9	1.12	2	22.22	0	0.00	0	0
	408 v 120	805	34	4.22	21	61.76	0	0.00	0	0
	17592 v 408	805	46	5.71	37	80.43	0	0.00	0	0
Negative ion polar metabolomic	10 v 2	353	0	0.00	0	0.00	0	0.00	0	0
	48 v 10	353	39	11.05	14	35.90	9	23.08	9	0
	74 v 48	353	72	20.40	54	75.00	28	38.89	27	1
	82 v 74	353	1	0.28	0	0.00	0	0.00	0	0
	120 v 82	353	10	2.83	1	10.00	1	10.00	0	1
	408 v 120	353	12	3.40	5	41.67	0	0.00	0	0
	17592 v 408	353	65	18.41	38	58.46	5	7.69	0	5
Positive ion polar metabolomic	10 v 2	341	0	0.00	0	0.00	0	0.00	0	0
	48 v 10	341	69	20.23	29	42.03	4	5.80	4	0
	74 v 48	341	83	24.34	47	56.63	37	44.58	37	0
	82 v 74	341	3	0.88	0	0.00	0	0.00	0	0
	120 v 82	341	2	0.59	0	0.00	0	0.00	0	0
	408 v 120	341	2	0.59	0	0.00	0	0.00	0	0
	17592 v 408	341	29	8.50	10	34.48	0	0.00	0	0

6.2.2.1. *Pre-diapause*

Embryonic development is characterized by an initial burst of transcriptional changes involving 16% of the total number of genes within *D. magna*'s genome (Table 6.5), which was subsequently followed by a more gradual increase of both metabolic activity and metabolite diversity (Figure 6.13). For instance, over 2,500 genes were significantly changed in expression at each of the first two sampled intervals, of which over 75% of these genes were uniquely changing at a specific interval of pre-diapause (Table 6.5). By contrast, only two metabolic peaks significantly differed in intensity during the first interval (from 1,499 recorded peaks during across the three experiments; Table 6.5), and 109 peaks showed significant changes in intensity during the second interval–where 40% were uniquely changing (Figure 6.14). This pattern strongly suggests that the transcriptomes are embryonic in origin, being largely (2/3) represented by the strong (> 2-fold) up-regulation of genes at increasing time points (Table 6-5; Figure 6.14). This pattern also strongly suggests that the metabolomes are reflective of the downstream actions of these genes thereby offering a unique molecular "cause-and-effect" perspective of the embryonic development towards diapause. The differential expressed (DE) transcriptome between 10 h and 48 h showed a more balanced number of strongly up and down regulated genes compared to the first interval between 2 h and 10 h. The fact that the DE transcriptomes of both intervals were unique between the sampled time-points suggests that the large fraction of initially up-regulated genes remained expressed throughout this pre-diapause stage. Indeed, only two metabolic peaks were significantly down regulated during this stage, suggesting an intensification of

metabolic activity until the initiation of diapause (Supplemental Figures 9.6-9.7). Overall, 25% of the total transcriptome (5,422 of a total 21,399 genes detected) and 7% (111 of a total 1,499 peaks detected) of the metabolome displayed significant changes during the pre-diapause period investigated, of these significantly changing elements 98% of genes and 41% of metabolites demonstrated changes uniquely within the pre-diapause period.

6.2.2.2. *Initiation of diapause*

The interval between 10 h and 48 h post-ovulation marks the transition between active development and the onset of developmental arrest. By 48 h, the sexually-produced embryos reached a stable ~3,500 cells and have thus entered diapause. The changes associated with this period and the following 74 h v 48 h period, identified by pairwise comparisons conducted to assess the molecular consequences of the initiation of diapause were previously reported in Chapter 5. Yet relative to the earlier stage, the entrance into diapause is characterized as an abrupt molecular point of departure, where the level of significantly changing gene expression rose to over 2,700 then dropped by 97.5%. Only 67 genes were differentially expressed between 48 h and 74 h. The strongly DE genes also reversed course from the earlier trend, by being significantly down regulated. By contrast to the earlier stage, only 45% of genes showing significant changes during this point of departure were uniquely DE at this transition period of diapause (Table 6.5). Yet while genome-wide transcription stalled and the number of detected non-polar metabolites dropped by 50% when in diapause (Supplemental Figure 9-6), polar metabolite production

increased overall during this transition period by 60%. Production of over 41% of these metabolites was greater than 2-fold (Figure 6.14).

Overall, the entrance into diapause is marked by an abrupt halt in changes to the transcriptome, accompanied by a sharp increase in metabolism producing a greater diversity (Figure 6-14) and amount (Figure 6-15) of metabolites.

6.2.2.3. *Early diapause*

Time points from 48 h post-ovulation onwards fall within the diapause phase where the bulk of sampling was performed. These samples cover the immediate hours and days preceding the initiation of diapause up to two weeks (408 h post-ovulation), henceforth referred to as “early diapause”. An additional time point in “late diapause” was collected 2 y after the onset of diapause for metabolomics studies only.

The first sampled time point during “early diapause” was 74 h. Thus, the 48 h to 74 h interval covers the molecular changes occurring within the first day of developmental arrest. This time point was also more extensively investigated in Section 5, yet relative to the earlier pre-diapause stage, only 67 genes (0.43% of those expressed) were differentially expressed followed by relative transcriptional inactivity over the three additional time intervals (Figure 6.12; Table 6.5). Indeed, less than 1% (175) of genes were differentially expressed at any individual time interval during the diapause phase. Of these, the majority exhibited a significant change in expression during one time point only (Figure 6-13) and were localised to either the immediate interval after the onset of diapause (78 v 48) or the longest interval sampled (408 v 120). Concurrently, early diapause over the three sampled time intervals post 74 h is also characterized by very few significant metabolite intensity changes; never

surpassing 4% of peaks detected (Figure 6.14). However, I did observe an increase in the number of significantly changing peaks in the lipidome, localised to the later time point in early diapause (408 v 120), with 34 peaks (4% of total) demonstrating a significant change in intensity (Table 6-5). Of the significant changes that did occur, all were characterised by small changes with no log_2 fold changes greater than 2-fold (Figure 6-15).

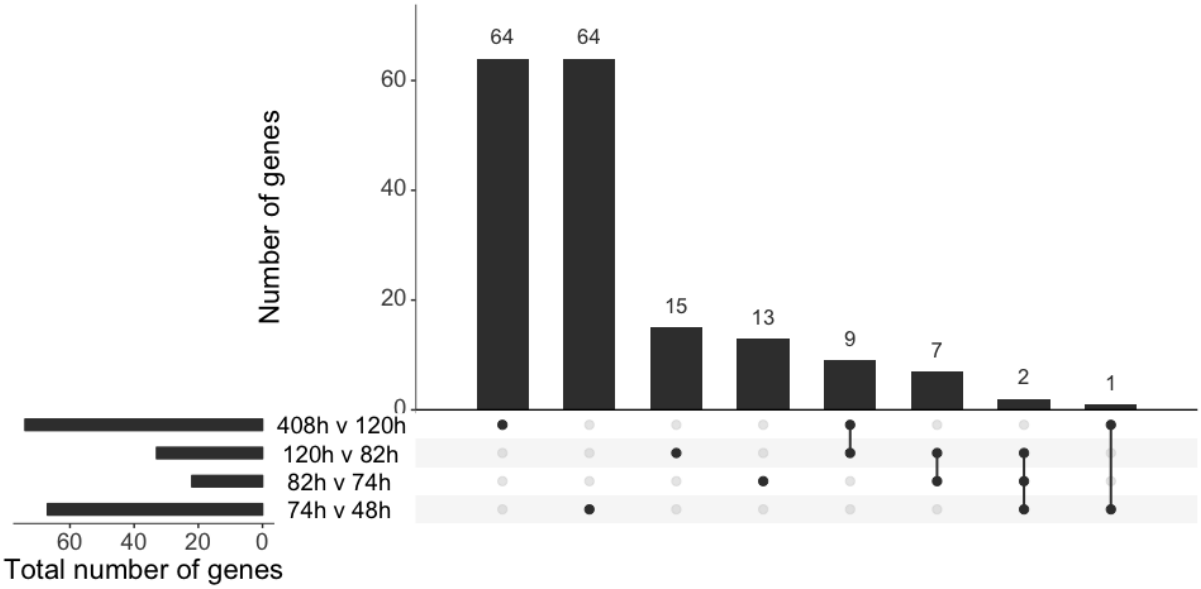


Figure 6-13 Plot of shared and unique differentially expressed genes identified in pairwise comparisons of successive sampling time points taken during early diapause (48 h–408 h). 'Total' genes refers to the total number within each time point that displayed a significant change in expression determined by *p* < 0.01 (after Benjamini-Hochberg false discovery rate correction). Empty intersections were no genes were shared between respective time points are not shown.

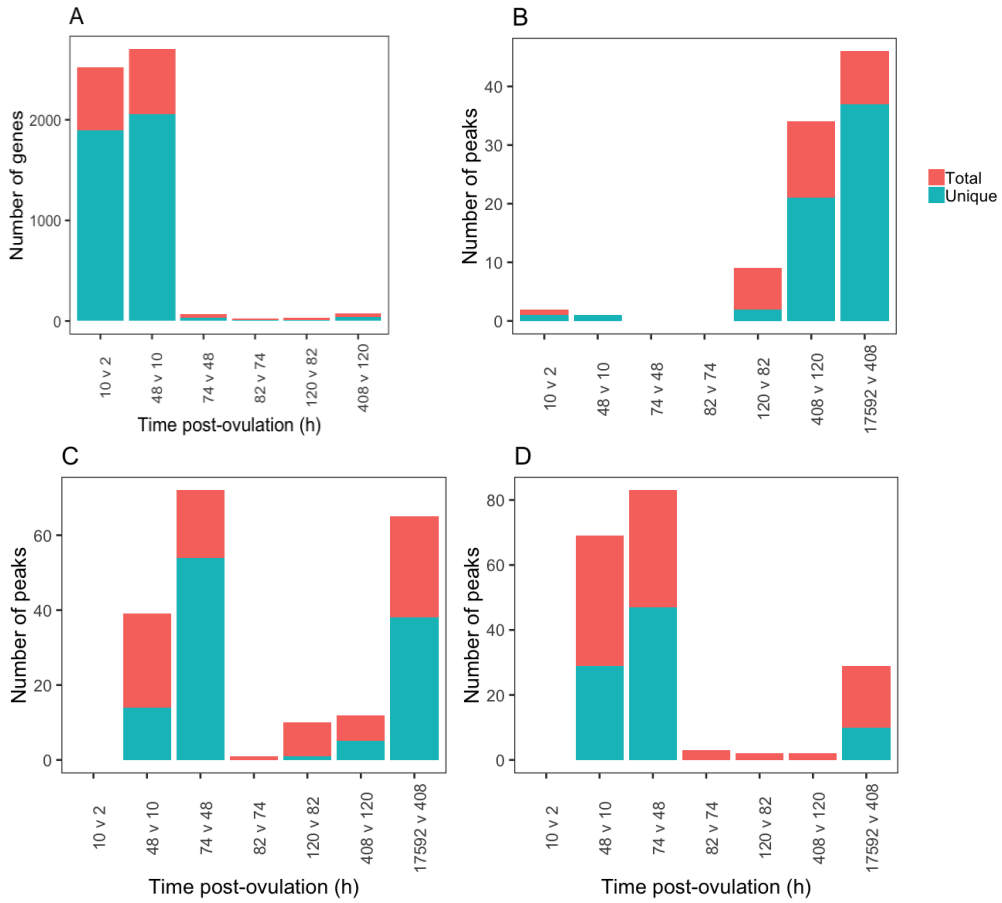
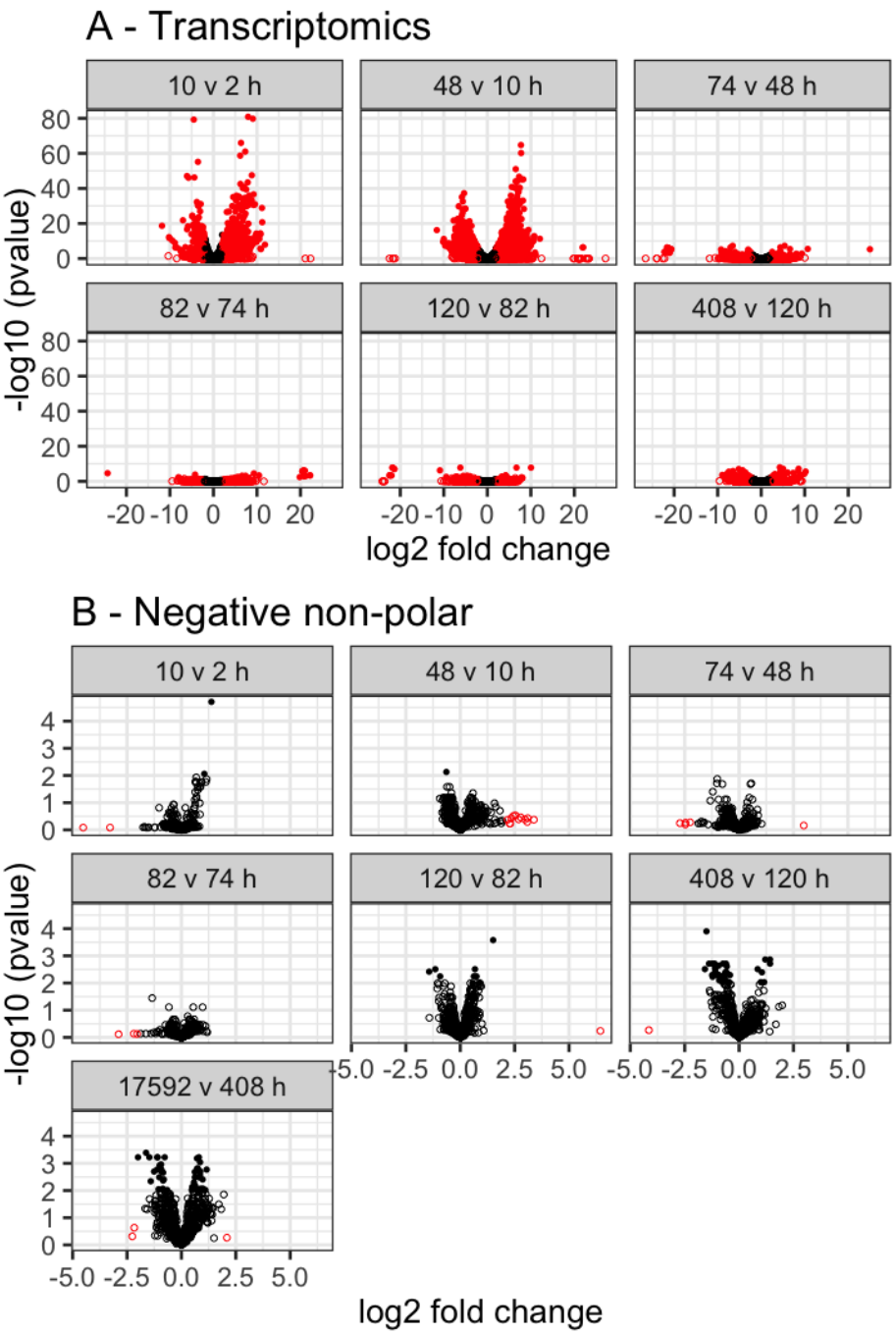


Figure 6-14 Number differentially expressed genes (A) and metabolites (B-D) with significant intensity fold changes between the indicated time points. Blue fractions of bars indicate the number of features that demonstrate a significant change only within the given developmental time interval (i.e. unique) or within at least one additional time interval (red fraction). Significantly changing features were determined by $p < 0.01$ (after Benjamini-Hochberg false discovery rate correction).


A - Transcriptomics
10 v 2 h
48 v 10 h
74 v 48 h
82 v 74 h
120 v 82 h
408 v 120 h
-log10 (pvalue)
log2 fold change
B - Negative non-polar
10 v 2 h
48 v 10 h
74 v 48 h
82 v 74 h
120 v 82 h
408 v 120 h
17592 v 408 h
-log10 (pvalue)
log2 fold change

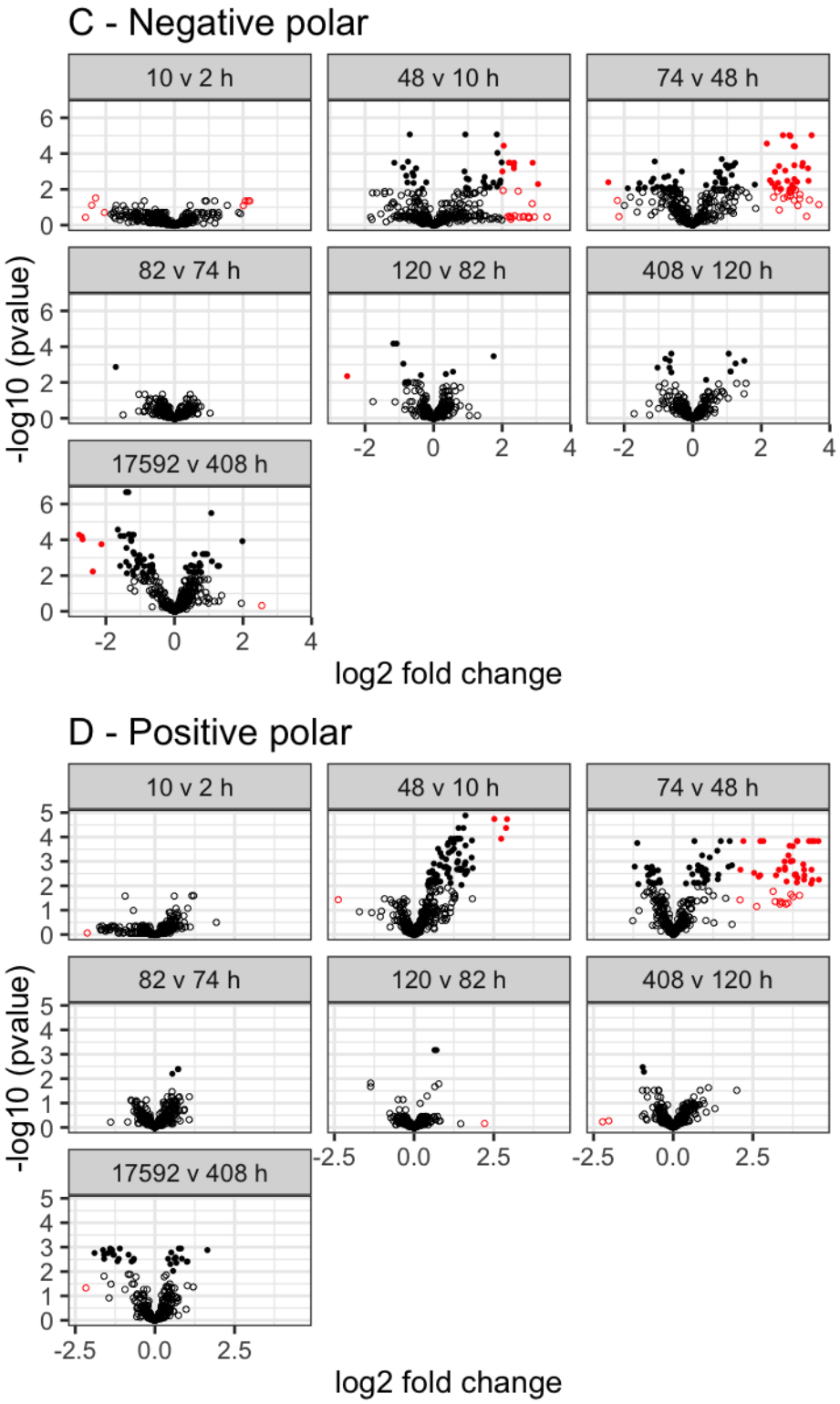


Figure 6-15 Volcano plots of differentially expressed genes (A) and metabolites (B-D) with significant fold changes between *Daphnia magna* embryos sampled at successive time points during diapause development. Metabolite analysis implemented nESI-DIMS metabolomics analysis of non-polar extracts (B; negative ion mode) and polar extracts analysed in negative ion mode (C) and positive ion mode (D). Filled points represent $p < 0.01$ (after Benjamini-Hochberg false discovery rate correction), red points indicate $\log_2$ fold change ratio > 2.

6.2.2.4. *Late diapause*

The longest time period covered within this thesis was collected for the metabolomics datasets only but consisted of embryos collected after a period in diapause of 2 years (17592 h). This already represents a substantial extension of lifespan for this comparatively short-lived species (Section 2.7). Pairwise comparisons of peak intensities at 17592 h v 408 h revealed an appreciable level of changes within the polar datasets with 65 (18% of total) and 29 (9% of total; Table 6-5) peaks showing significant changes in intensity in negative and positive ion datasets respectively. However, only 5 of these peaks also demonstrated large fold changes, all of which were in the negative ion dataset and were decreases in intensity (Figure 6-15). This finding suggests that larger and more numerous changes in detected polar metabolite peaks occur in the immediate time points covering the initiation of diapause in comparison to those occurring over long term (ca. 2 y) diapause periods. By contrast, 46 peaks changed significantly in intensity between 408 h v 17592 h in the analysis of non-polar extracts (Table 6-5). Although representing just 6% of the peaks within the dataset, this is the largest number of significantly changing peaks between any of the time intervals investigated in the non-polar dataset (Figure 6-12). Furthermore, the peak composition of the non-polar extracts are unique to this 2 year sample time, sharing only about 50% of the peaks with all of the other sample times (Figure 9-6). By contrast, in the lipidomic analysis of the pupal diapause of *Pieris napi* authors report no major changes in lipid classes and/or species in pupae throughout the course of diapause (6 d to 144 d) (Lehmann *et al.*, 2018). Although the study also included post-diapause and non-diapause samples in the analysis which represented

the major differences identified (Lehmann *et al.*, 2018), and thus could be masking more subtle, small scale changes such that I identify here.

6.3. Conclusions

The aim of the work presented in this chapter was to characterize the temporal dynamics of the transcriptomes and metabolomes of *D. magna* throughout the trajectory of diapause – thereby providing a *molecular phenotype* describing the likely processes that produce an embryo that is capable of suspended animation for decades and even centuries – by covering the transition from pre-diapause to the initiation of and early stages of diapause. Additionally, for metabolomics data, I investigated differences in early to late diapause. Unsupervised sPCA revealed distinct separation of sexually-produced *D. magna* embryos in relation to their diapause stage. There was a distinct clustering of pre-diapause and diapausing embryos in both the transcriptomic and polar metabolomic datasets, although lipidomic samples did not demonstrate this distinct separation. Thus, *D. magna* embryos display unique molecular profiles of gene expression and metabolism through the diapause trajectory. Furthermore, observing metabolite peaks and expressed genes with non-zero sPC loadings, and conducting sequential pairwise analyses of sampling time points revealed distinct localisations of molecular events and patterns of dynamic behaviour which are summarised in Figure 6-16 in relation to the respective diapause phases.

6.3.1. The molecular phenotypes of diapause

The first molecular phenotype observed consisted of elements that had elevated levels of expression during pre-diapause development, with strong down-regulation

coinciding with the initiation of diapause, followed by a consistent low level during the diapause period. This phenotype was largely restricted to the transcriptome (Figure 6-2; A.2). However, a handful of peaks detected in the lipidome (Figure 6-5; B) and negative ion analysis of polar metabolites (Figure 6-9) also displayed a decrease in intensity, although more gradual throughout pre-diapause with constant levels during the diapause period. In the transcriptome, these genes were largely annotated with functions consistent with the progression of embryonic development (e.g. mitosis, DNA replication, and the cell cycle)(Figure 5-6). Thus, is it logical that these elements drop out after 48 h when mitotic activity is supressed and diapause has initiated, and I suggest that their behaviour can be conceptually represented by the “drivers of direct development” trajectory presented in Figure 6-16. I anticipate that metabolites that also display decreases in intensity during the pre-diapause phase are likely to be substrates/products of pre-diapause metabolic activity also involved in the progress of pre-diapause development.

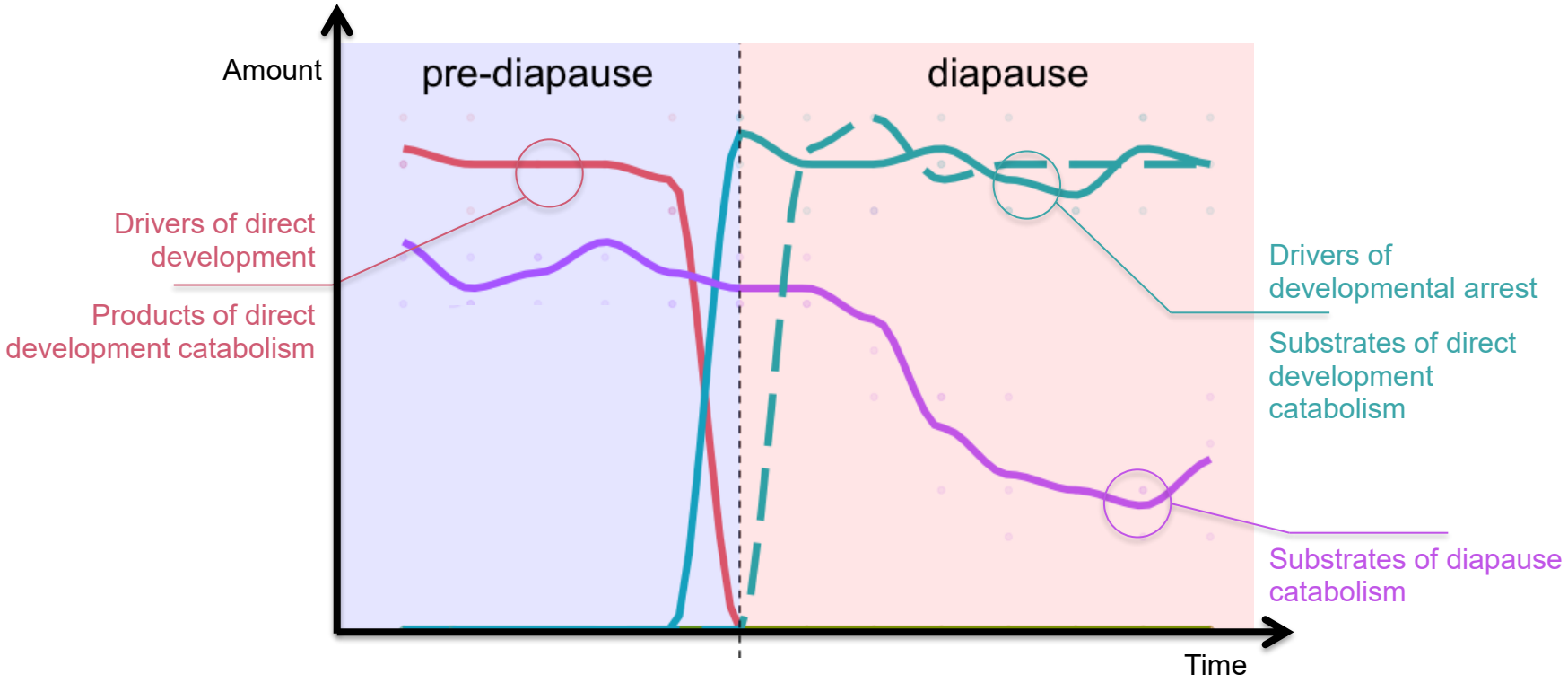


Figure 6-16 Conceptual figure describing the molecular phenotypes of embryonic diapause as described within this thesis. Coloured lines represent distinct phenotype trajectories characterised by element amount (i.e. gene expression or metabolite intensity) with time in diapause development. Shaded plot areas indicated pre-diapause (blue) and diapause (red) phases, where vertical hashed black line indicates the initiation of diapause.

A second phenotype observed covers elements demonstrating the opposite behaviour: these elements have negligible levels during pre-diapause with a strong up-regulation upon the initiation of diapause. Whilst the majority of transcriptional up-regulation occurred directly with the initiation of diapause (i.e. between 48 h and 10 h), metabolic changes occur slightly later (i.e. between 74 h and 48 h). It is currently not known whether *D. magna* diapause is accompanied by metabolic suppression as observed in other species such as *S. crassipalpis* (Ragland, Denlinger and Hahn, 2010) and *A. franciscana* (Podrabsky and Hand, 2015). However, the large step changes described in this thesis (Figure 6-7 and Figure 6-8; A.2) in relation to the initiation of diapause are indicative of major alterations in polar metabolism. I suggest that peaks with positive step changes (blue trajectory; Figure 6-16) may be metabolic substrates utilised during direct development that increase in relative intensity after the initiation of diapause as they are no longer being utilised. Alternatively, as with the putative trehalose metabolite identified, peaks with increased intensity during diapause could also be physiological relevant, potentially providing cryo-protection (Vesala *et al.*, 2012) or energy reserves stock-piled prior to the initiation of developmental arrest for use during the diapause period or for post-diapause development (Hahn and Denlinger, 2007). Indeed, putatively annotated up-regulated genes at the initiation of diapause included superoxide dismutases (Figure 5-5) and glutathione peroxidases and transferases (Figure 4-8), suggesting that elements with increased levels during the diapause period have permissive roles, (e.g. the mitigation of oxidative stress) and are thus ensuring the successful progression of diapause development.

A single “late” diapause sample was collected for metabolomics datasets and in contrast to early diapause samples a comparatively large number of peaks demonstrated significant fold changes between 2 w (408 h) and 2 y in diapause (Figure 6-6). Most specifically, during this interval observable differences in non-polar metabolites were detectable that until now appeared to be somewhat refractory to diapause associated perturbations. I postulate that these changes represent the utilisation of energy reserves to drive core molecular processes occurring during the extended diapause period i.e. “substrates of diapause catabolism” (Figure 6-16; purple trajectory). These features are of particular interest, in that a full elucidation of the availability and adequate management of energetic resources during *D. magna* diapause as well as the other molecular changes occurring will likely be relevant to understanding how this comparably short lived species can survive the dramatically extended periods of dormancy that have been documented.

Ultimately the results in this chapter describe very complex molecular behaviour throughout the diapause developmental trajectory that will require further investigation and validation. Results reiterate the major molecular changes observed in relation to the initiation of diapause but highlight that there are observable molecular changes occurring after the onset of developmental arrest into long-term dormancy.

7. Conclusions and future perspectives

The overall aim of this thesis was to develop *Daphnia magna* as a model system for the molecular investigation of embryonic diapause.

In Section 3 the laboratory induction of diapause and biomass required for DIMS-based metabolomics was investigated. In Section 4 the development of subitaneous and diapause-destined embryos was assessed both on a cellular and molecular level through the collection of three large omics datasets, covering the transcriptome, polar, and non-polar metabolome and the fluorescent staining of embryonic nuclei. Having identified the point of diapause initiation in *D. magna,* the specific molecular changes associated with the onset of diapause were investigated in Section 5. Finally in Section 6, the molecular changes occurring throughout the phases of pre-diapause and diapause were investigated, including the alterations to the metabolome occurring after 2 years in diapause.

In Section 3, I successfully established procedures for the stable generation of sexually derived diapausing embryos under controlled laboratory conditions. A combination of reduced food (0.03 mgC *Daphnia*$^{-1}$), and increased population density (62 *D. magna* L^{-1}) using a specific *D. magna* isolate (L7) and a male:female ratio of 1:4, demonstrated a consistently high and reproducible yield of diapausing offspring. Furthermore, these offspring were viable after up to 2 y in diapause as determined by embryo hatching and survival to reproduction and thus demonstrate an increased lifespan relative to subitaneous offspring. The experimental system I optimised and utilise throughout this thesis is particularly unique in that subitaneous and diapause-destined offspring were produced under the same, stable environmental conditions: a

compounding factor that has largely been ignored in molecular investigations of diapause performed to date (Kankare *et al.*, 2016).

However, the work presented in this section is by no means an exhaustive investigation of the induction of diapause in *D. magna.* A much greater body of work is required, including elucidation of the molecular events between maternal perception of diapause-inducing stimuli and the physiological switch from asexual to sexual reproduction. An initial starting point would be to focus on differences in maternal whole animal and/or ovarian transcription between the two forms of reproduction. However, this would be further complicated by the diversity of environmental cues known to induce diapause in *Daphnia sp.* For example, it is unknown to what degree the signal transduction pathways that perceive photoperiodic, predatory, nutritional and thermal cues overlap and furthermore whether there are phenotypic differences between the offspring produced under different inductive stimuli (Stross and Hill, 1965; Brunner and Halcrow, 1977; Alekseev and Lampert, 2001; Ślusarczyk, Dawidowicz and Rygielska, 2005; Koch, von Elert and Straile, 2009).

In Section 4, fluorescent staining of embryonic nuclei revealed that subitaneous and diapause-destined *D. magna* embryos exhibit unique profiles of cell division from the outset of embryonic development. Diapause-destined embryos are slower to divide and display a decrease in the rate of cell division after approximately 30 h, likely indicating the progression to developmental arrest. I was able to successfully ascertain that sexually produced *D. magna* embryos enter developmental arrest at ~3,500 cells, which occurs before 48 h post-ovulation when embryos are maintained at 20°C. Critically, this allowed me to define and frame observations of molecular

data within the phases of pre-diapause (<48 h post-ovulation) and diapause (>48 h post-ovulation) development. A logical extension of this work will be to extend observations to multiple isolates of *D. magna*, to ascertain whether my observations regarding the stage (i.e. number of cells) and profile of cell division prior to developmental arrest are applicable to the species as a whole. Additionally, whole mount preparations of both subitaneous and diapause-destined embryos throughout the same developmental period will complement my observations made from squashed embryos. Fluorescence microscopy of appropriately stained (e.g. DAPI, SYBR® Green) whole mounts will preserve embryo architecture and thus provide evidence of the morphological stage of diapause initiation (i.e. point of developmental arrest) in *D. magna*. Observations should also be extended to other *Daphnia* species to assess whether the stage of developmental arrest is conserved.

I then assessed the specific molecular changes occurring between subitaneous and diapause-destined embryos across molecular data shared between the two embryo types: 2–48 h post-ovulation. A subset of genes and metabolites were found to vary between embryo types throughout the time course investigated including, genes relating to the cell cycle and putatively annotated glutathione metabolites, indicating that unique processes are occurring in the alternative developmental trajectories. Furthermore, inherent differences in the metabolome or transcriptome between diapause–destined and subitaneous embryos were also described including down-regulation in diapause-destined embryos of genes associated with protein translation and ribosomal processes. Several putatively annotated triacylglycerides were also significantly higher in intensity in diapause-destined embryos versus subitaneous embryos which is in line with previous reports that the two embryo types differ in fatty

acid composition (Abrusan, Fink and Lampert, 2007). The work in this chapter emphasised that beyond the morphological differences, sexually– and asexually–derived *D. magna* are different on a molecular level from the outset of development. This further emphasises the knowledge gap regarding the processes occurring during the maternal induction of diapause, both in signal transduction and the physiological preparations that will require robust investigation to fully understand the diapause process. Furthermore, it also highlights the futility of the convention thus far in studies that attempt to understand diapause whilst relying on observations made solely by comparing directly developing to diapause life stages.

Having identified when developmental arrest occurs, I described in Section 5 profound transcriptomic and metabolomic changes occurring in relation to the onset of diapause. This was achieved through pairwise comparisons of the diapause trajectory alone considering time points pre– (10 h) and post–diapause (48 h and 74 h). 13% (2,231) of genes were significantly differentially expressed at 48 h versus 10 h with roughly half up- and down-regulated. These included down-regulated genes relating to DNA replication and the cell cycle, that have previously been identified in other studies investigating the mechanisms of diapause including *pcna* (Tammariello and Denlinger, 1998; Hayward *et al.*, 2005; Koštál *et al.*, 2009; Hao *et al.*, 2016; Ziv *et al.*, 2017) and cyclin genes (Zhao *et al.*, 2012; Hao *et al.*, 2016; Shimizu, Mukai and Goto, 2016),

I hypothesise that genes down-regulated at the initiation of diapause represent likely drivers of active development and genes up-regulated are potential effectors of developmental arrest. And whilst a down-regulation of genes associated with cell division offers a plausible mechanism by which diapause is initiated, application of

further data analysis to these gene lists should include upstream regulator analysis such that available in Ingenuity® Pathway Analysis (QIAGEN Bioinformatics; Krämer *et al.*, 2014) which may shed light on the molecules of interest (e.g. transcription factors) involved in the upstream regulation of diapause-associated developmental arrest.

In contrast, the largest perturbations to the polar metabolome occurred later than gene expression at 74 h v 48 h, with 8% and 11% of peaks in negative and positive ion mode respectively displaying significant, large changes in intensity. Interestingly the majority of peaks displayed an increase in intensity upon the initiation of diapause including a peak assigned putative disaccharide identity including among the 26 possible isomers trehalose. Identifying whether this peak does indeed correspond to trehalose will require the use of prior chromatographic separation and/or fragmentation (tandem MS) with comparison to spectra from authentic trehalose standards (Sumner *et al.*, 2007). Trehalose may be important during *D. magna* diapause as it has been implicated in anhydrobiosis, desiccation resistance (Hengherr *et al.*, 2011; MacRae, 2016) and cryoprotection (Su et al., 2017) during diapause in other species. Spectrophotometric assays have previously been used to investigate the activity of enzymes involved in trehalose synthesis (Trehalose-6-phosphate synthase and trehalose-6-phosphatase) in addition to quantifying the levels of trehalose in whole organisms (pupae; Guo *et al.*, 2015). Confirming that functional trehalose synthesis enzymes are present in *D. magna* and correlate with transcriptional levels would be a first step towards more robustly investigating the relationship between trehalose and diapause in *D. magna.* Following this, RNAi-mediated suppression of genes relating to trehalose and assessment of any

correlated effects on the diapause phenotype (e.g. changes to the incidence of and/or duration of diapause, in addition to possible alterations in desiccation and/or cold tolerance).

Whilst both datasets in isolation inform us that there are significant molecular changes associated with the onset of developmental arrest, when considered together and with multiple sampling stages they highlight the dynamic nature of diapause in that these changes occur along a different timeline with metabolomics occurring later than gene expression. A suppression of metabolism is considered a hallmark of diapause (Hand and Podrabsky, 2000). Yet there is no direct evidence to suggest that the same extreme suppression of metabolism is present in *Daphnia sp.* as has been observed in other species. However, I believe that the large fold changes may represent the accumulation of intermediates due the reduction in metabolic process and suggest that investigation of active metabolism in *Daphnia sp.* will prove useful in the interpretation of molecular data. Specifically, comparing the oxygen consumption (i.e. through respirometry measurements) of actively developing asexual and pre-diapause sexual embryos at comparable cell numbers will give an indication of whether the two embryo types demonstrate alternate metabolic rates. Furthermore, investigation of whether the oxygen consumption of sexual embryos drops upon the initiation of diapause will confirm the magnitude of metabolic depression during diapause in *D. magna* (Reynolds and Hand, 2009).

Analysis performed in Section 6 revealed that *D. magna* embryos display unique molecular profiles of gene expression and metabolism through the phases of diapause investigated, covering the transition from pre-diapause to the initiation of and early stages of diapause and for metabolomics only, extending into late diapause

(2 y). Largely, these results reiterated those presented in previous analyses, particularly emphasising the drastic nature of the molecular changes associated with diapause initiation. Most specifically, during the interval between 2 w (408 h) and 2 y in diapause, differences in non-polar metabolites were detectable. I postulate that these changes may represent the utilisation of energy reserves to drive core molecular processes occurring during the extended diapause period. Extending investigations of gene expression into the later stages of dormancy and the collection additional time points for metabolite investigation will be needed to understand how this comparably short lived species can survive the dramatically extended periods of dormancy that have been documented (Orsini, Spanier and De Meester, 2012; Frisch *et al.*, 2014). This will involve building a more complete picture of how energy reserves are utilised in addition to identifying the core molecular processes occurring during the extended diapause period. These may be relevant not just to diapause but in understanding the processes of extended lifespan and ageing.

The immediate future work associated with this project will be the integration of all three ‘omics’ datasets using advanced statistical techniques that were beyond the scope of this thesis. Analyses aimed at identifying correlated features (i.e. genes and metabolites) from multiple datasets such as regularized canonical correlation analysis (González *et al.*, 2008) as implemented in mixOmics (Lê Cao, González and Déjean, 2009) should be applied to identify key genes and metabolites that are associated with the phases of diapause development (i.e. pre-diapause, initiation and maintenance). This will ultimately allow me to further prioritise features of interest for further investigation including the aforementioned examples of more in-depth genetic functional studies and more thorough metabolite compound identification.

In conclusion, the work presented in this thesis successfully presents *Daphnia magna* as a model system for the investigation of the molecular processes occurring during embryonic diapause. Outlining its suitability through a robust experimental design that has allowed me to define the previously unknown stages of diapause in this species. In addition to highlighting the scale and temporal changes in both gene expression and metabolism associated with the alternative developmental trajectory of diapause.

8. References

Abdi, H. and Williams, L. J. (2010) 'Principal component analysis', *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(4), pp. 433–459. doi: 10.1002/wics.101.

Abrusan, G., Fink, P. and Lampert, W. (2007) 'Biochemical limitation of resting egg production in Daphnia', *Limnology and Oceanography*, 52(4), pp. 1724–1728. doi: 10.4319/lo.2007.52.4.1724.

Aebersold, R. and Mann, M. (2003) 'Mass spectrometry-based proteomics', *Nature*, 422(6928), pp. 198–207. doi: 10.1038/nature01511.

Alekseev, V. and Lampert, W. (2001) 'Maternal control of resting-egg production in Daphnia.', *Nature*, 414(6866), pp. 899–901. doi: 10.1038/414899a.

Alekseev, V. R. and Lajus, D. L. (2009) 'Maternal and direct effects of natural-like changes of photoperiod and food condition manipulation on life history parameters in daphnia', *Aquatic Ecology*, 43(2), pp. 415–421. doi: 10.1007/s10452-008-9168-9.

Amsalem, E., Galbraith, D. A., Cnaani, J., Teal, P. E. A. and Grozinger, C. M. (2015) 'Conservation and modification of genetic and physiological toolkits underpinning diapause in bumble bee queens', *Molecular Ecology*, 24(22), pp. 5596–5615. doi: 10.1111/mec.13410.

Arbačiauskas, K. (2004) 'Life-history characteristics and fitness in descendents of parthenogenetic and ex-ephippio females of Daphnia magna', *Hydrobiologia*, 526(1 SPEC. ISS.), pp. 211–218. doi: 10.1023/B:HYDR.0000041602.62586.47.

Arbačiauskas, K. and Lampert, W. (2003) 'Seasonal adaptation of ex-ephippio and parthenogenetic offspring of Daphnia magna: Differences in life history and physiology', *Functional Ecology*, 17(4), pp. 431–437. doi: 10.1046/j.1365-2435.2003.00748.x.

Arendt, D. and Nübler-Jung, K. (1997) 'Dorsal or ventral: Similarities in fate maps and gastrulation patterns in annelids, arthropods and chordates', *Mechanisms of Development*, 61(1–2), pp. 7–21. doi: 10.1016/S0925-4773(96)00620-X.

Arrese, E. L. and Soulages, J. L. (2010) 'Insect Fat Body: Energy, Metabolism, and Regulation', *Annual Review of Entomology*, 55(1), pp. 207–225. doi: 10.1146/annurev-ento-112408-085356.

Baer, K. N. and Goulden, C. E. (1998) 'Evaluation of a high-hardness COMBO medium and frozen algae for Daphnia magna.', *Ecotoxicology and environmental safety*, 39(3), pp. 201–6. doi: 10.1006/eesa.1997.1627.

Bailly, V., Lamb, J., Sung, P., Prakash, S. and Prakash, L. (1994) 'Specific complex formation between yeast RAD6 and RAD18 proteins: A potential mechanism for targeting RAD6 ubiquitin-conjugating activity to DNA damage sites', *Genes and Development*, 8(7), pp. 811–820. doi: 10.1101/gad.8.7.811.

Bale, J. S. (1996) 'Insect cold hardiness: A matter of life and death', *European Journal of Entomology*, 93, pp. 369–382.

Bao, B. and Xu, W.-H. (2011) 'Identification of gene expression changes associated with the initiation of diapause in the brain of the cotton bollworm, Helicoverpa armigera', *BMC Genomics*, 12(1), p. 224. doi: 10.1186/1471-2164-12-224.

Bashan, M. and Cakmak, O. (2005) 'Changes in Composition of Phospholipid and

Triacylglycerol Fatty Acids Prepared from Prediapausing and Diapausing Individuals of Dolycoris baccarum and Piezodorus lituratus (Heteroptera: Pentatomidae)', *Annals of the Entomological Society of America*, 98(4), pp. 575–579. doi: 10.1603/0013-8746(2005)098[0575:CICOPA]2.0.CO;2.
Baumgartner, M. F. and Tarrant, A. M. (2017) 'The Physiology and Ecology of Diapause in Marine Copepods', *Annual Review of Marine Science*, 9(1), pp. 387–411. doi: 10.1146/annurev-marine-010816-060505.
Baumgartner, M. F. and Tarrant, A. M. (2017) 'The Physiology and Ecology of Diapause in Marine Copepods', *Annual Review of Marine Science*, 9(1), pp. 387–411. doi: 10.1146/annurev-marine-010816-060505.
Becker, D., Brinkmann, B. F., Zeis, B. and Paul, R. J. (2011) 'Acute changes in temperature or oxygen availability induce ROS fluctuations in Daphnia magna linked with fluctuations of reduced and oxidized glutathione, catalase activity and gene (haemoglobin) expression.', *Biology of the cell / under the auspices of the European Cell Biology Organization*, 103(8), pp. 351–63. doi: 10.1042/BC20100145.
Belmonte, G. and Pati, A. C. (2007) 'Hatching rate and diapause duration in eggs of Paracartia latisetosa (Copepoda: Calanoida)', *Journal of Plankton Research*, 29(SUPPL. 1), pp. 39–47. doi: 10.1093/plankt/fbl064.
Benjamini, Y. and Hochberg, Y. (1995) 'Controlling the False Discovery Rate : A Practical and Powerful Approach to Multiple Testing Author (s): Yoav Benjamini and Yosef Hochberg Source : Journal of the Royal Statistical Society . Series B (Methodological), Vol . 57 , No . 1 Published by ':, *J. R. Statist. Soc.*, 57(1), pp. 289–300.
Bersanelli, M., Mosca, E., Remondini, D., Giampieri, E., Sala, C., Castellani, G. and Milanesi, L. (2016) 'Methods for the integration of multi-omics data: Mathematical aspects', *BMC Bioinformatics*, 17(2). doi: 10.1186/s12859-015-0857-9.
Bligh, E.G. and Dyer, W. J. (1959) 'Canadian Journal of Biochemistry and Physiology', *Canadian Journal of Biochemistry and Physiology*, 37(8).
Boersma, M., Boriss, H. and Mitchell, S. E. (2000) 'Maternal effects after sexual reproduction in Daphnia magna', 22(2), pp. 279–285.
Boersma, M. and Kreutzer, C. (2002) 'Life at the Edge : Is Food Quality Really of Minor Importance at Low Quantities ?', *Ecology*, 83(9), pp. 2552–2561.
Bold, H. C. (1949) 'The Morphology of Chlamydomonas chlamydogama, Sp. Nov.', *Bulletin of the Torrey Botanical Club*, 26, pp. 101–108.
Bommer, U. A., Lutsch, G., Stahl, J. and Bielka, H. (1991) 'Eukaryotic initiation factors eIF-2 and eIF-3 : interactions , structure and localization in ribosomal initiation complexes', pp. 1007–1019.
Bonner, A. J. and Beyene, J. (2017) 'Evaluating the performance of sparse principal component analysis methods in high-dimensional data scenarios', *Communications in Statistics: Simulation and Computation*. Taylor & Francis, 46(5), pp. 3794–3811. doi: 10.1080/03610918.2015.1004268.
Booksmythe, I., Gerber, N., Ebert, D. and Kokko, H. (2018) 'Daphnia females adjust sex allocation in response to current sex ratio and density', *Ecology Letters*, pp. 1–9. doi: 10.1111/ele.12929.
Bradshaw, A. W. E., Armbruster, P. A. and Holzapfel, C. M. (2009) 'Fitness Consequences of Hibernal Diapause in the Pitcher-Plant Mosquito, Wyeomyia smithii', *Ecological Society of America*, 79(4), pp. 1458–1462.
Brendonck, L. and Meester, L. De (2003) 'Egg banks in freshwater zooplankton :

evolutionary and ecological archives in the sediment', pp. 65–84.
Bro, R. and Smilde, A. K. (2014) 'Principal component analysis', *Analytical Methods*, 6(9), pp. 2812–2831. doi: 10.1039/c3ay41907j.
Brunner, H. C. and Halcrow, K. (1977) 'Experimental induction of the production of ephippia by Daphnia magna straus (Cladocera)', *Crusaceana*.
Burns, C. W. (2000) 'Crowding-induced changes in growth, reproduction and morphology of Daphnia', *Freshwater Biology*, 43(1), pp. 19–29. doi: 10.1046/j.1365-2427.2000.00510.x.
Cáceres, C. E. and Tessier, A. J. (2003) 'How long to rest: The ecology of optimal dormancy and environmental constraint', *Ecology*, 84(5), pp. 1189–1198. doi: 10.1890/0012-9658(2003)084[1189:HLTRTE]2.0.CO;2.
Cáceres, C. E. and Tessier, A. J. (2004) 'Incidence of diapause varies among populations of Daphnia pulicaria', *Oecologia*, 141(3), pp. 425–431. doi: 10.1007/s00442-004-1657-5.
Cattie, D. J., Richardson, C. E., Reddy, K. C., Ness-Cohn, E. M., Droste, R., Thompson, M. K., Gilbert, W. V. and Kim, D. H. (2016) 'Mutations in Nonessential eIF3k and eIF3l Genes Confer Lifespan Extension and Enhanced Resistance to ER Stress in Caenorhabditis elegans', *PLoS Genetics*, 12(9), pp. 1–20. doi: 10.1371/journal.pgen.1006326.
Chandramouli, K. and Qian, P.-Y. (2009) 'Proteomics: Challenges, Techniques and Possibilities to Overcome Biological Sample Complexity', *Human Genomics and Proteomics*, 2009, pp. 1–22. doi: 10.4061/2009/239204.
Chang, A. C., Cha, J., Koentgen, F. and Reddel, R. R. (2005) 'The Murine Stanniocalcin 1 Gene Is Not Essential for Growth and Development', *Society*, 25(23), pp. 10604–10610. doi: 10.1128/MCB.25.23.10604.
Chang, A. C. M., Hook, J., Lemckert, F. A., McDonald, M. M., Nguyen, M. A. T., Hardeman, E. C., Little, D. G., Gunning, P. W. and Reddel, R. R. (2008) 'The murine stanniocalcin 2 gene is a negative regulator of postnatal growth', *Endocrinology*, 149(5), pp. 2403–2410. doi: 10.1210/en.2007-1219.
Chang, J. and Ramasamy, S. (2014) 'Identification and Expression Analysis of Diapause Hormone and Pheromone Biosynthesis Activating Neuropeptide (DH-PBAN) in the Legume Pod Borer , Maruca vitrata Fabricius', 9(1), pp. 1–11. doi: 10.1371/journal.pone.0084916.
Chen, D.-F., Lin, C., Wang, H.-L., Zhang, L., Dai, L., Jia, S.-N., Zhou, R., Li, R., Yang, J.-S., Yang, F., Clegg, J. S., Nagasawa, H. and Yang, W.-J. (2016) 'An La-related protein controls cell cycle arrest by nuclear retrograde transport of tRNAs during diapause formation in Artemia', *BMC Biology*. BMC Biology, 14(1), p. 16. doi: 10.1186/s12915-016-0239-4.
Chen, S., Chen, D.-F., Yang, F., Nagasawa, H. and Yang, W.-J. (2011) 'Characterization and processing of superoxide dismutase-fused vitellogenin in the diapause embryo formation: a special developmental pathway in the brine shrimp, Artemia parthenogenetica.', *Biology of reproduction*, 85(1), pp. 31–41. doi: 10.1095/biolreprod.110.090340.
Chen, Y., Jiang, T., Zhu, J., Xie, Y., Tan, Z., Chen, Y., Tang, S., Hao, B., Wang, S., Huang, J. and Shen, X. (2017) 'Transcriptome sequencing reveals potential mechanisms of diapause preparation in bivoltine silkworm Bombyx mori (Lepidoptera : Bombycidae)', *Comparative Biochemistry and Physiology - Part D*. Elsevier, 24(July), pp. 68–78. doi: 10.1016/j.cbd.2017.07.003.

Chen, Y. rong, Jiang, T., Zhu, J., Xie, Y. chen, Tan, Z. cheng, Chen, Y. hua, Tang, S. ming, Hao, B. fang, Wang, S. peng, Huang, J. shan and Shen, X. jia (2017) 'Transcriptome sequencing reveals potential mechanisms of diapause preparation in bivoltine silkworm Bombyx mori (Lepidoptera: Bombycidae)', *Comparative Biochemistry and Physiology - Part D: Genomics and Proteomics*. Elsevier, 24(July), pp. 68–78. doi: 10.1016/j.cbd.2017.07.003.
Chong, J., Soufan, O., Li, C., Caraus, I., Li, S., Bourque, G., Wishart, D. S. and Xia, J. (2018) 'MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis', *Nucleic Acids Research*. Oxford University Press, 46(W1), pp. W486–W494. doi: 10.1093/nar/gky310.
Choudhuri, A., Maitra, U. and Evans, T. (2013) 'Translation initiation factor eIF3h targets specific transcripts to polysomes during embryogenesis', *Proceedings of the National Academy of Sciences*, 110(24), pp. 9818–9823. doi: 10.1073/pnas.1302934110.
Churchill, G. A. (2002) 'Fundamentals of experimental design for cdna microarrays', *Nature Genetics*, 32(4S), pp. 490–495. doi: 10.1038/ng1031.
Clegg, J. S. (1997) 'EMBRYOS OF ARTEMIA FRANCISCANA SURVIVE FOUR YEARS OF CONTINUOUS ANOXIA : THE CASE FOR COMPLETE METABOLIC RATE DEPRESSION', 475, pp. 467–475.
Clegg, J. S. (2001) 'Cryptobiosis - a peculiar state of biological', *Comparative Biochemistry and Physiology Part B*, 128, pp. 613–624. doi: 10.1016/S1096-4959(01)00300-1.
Colbourne, J. K., Pfrender, M. E., Gilbert, D., Thomas, W. K., Tucker, A., Oakley, T. H., Tokishita, S., Aerts, A., Arnold, G. J., Basu, M. K., Bauer, D. J., Cáceres, C. E., Carmel, L., Casola, C., Choi, J.-H., Detter, J. C., Dong, Q., Dusheyko, S., Eads, B. D., Fröhlich, T., Geiler-Samerotte, K. a, Gerlach, D., Hatcher, P., Jogdeo, S., Krijgsveld, J., Kriventseva, E. V, Kültz, D., Laforsch, C., Lindquist, E., Lopez, J., Manak, J. R., Muller, J., Pangilinan, J., Patwardhan, R. P., Pitluck, S., Pritham, E. J., Rechtsteiner, A., Rho, M., Rogozin, I. B., Sakarya, O., Salamov, A., Schaack, S., Shapiro, H., Shiga, Y., Skalitzky, C., Smith, Z., Souvorov, A., Sung, W., Tang, Z., Tsuchiya, D., Tu, H., Vos, H., Wang, M., Wolf, Y. I., Yamagata, H., Yamada, T., Ye, Y., Shaw, J. R., Andrews, J., Crease, T. J., Tang, H., Lucas, S. M., Robertson, H. M., Bork, P., Koonin, E. V, Zdobnov, E. M., Grigoriev, I. V, Lynch, M. and Boore, J. L. (2011) 'The ecoresponsive genome of Daphnia pulex.', *Science (New York, N.Y.)*, 331(6017), pp. 555–61. doi: 10.1126/science.1197761.
Colinet, H., Renault, D., Charoy-Guével, B. and Com, E. (2012) 'Metabolic and proteomic profiling of diapause in the aphid parasitoid praon volucre', *PLoS ONE*, 7(2). doi: 10.1371/journal.pone.0032606.
Conde-Porcuna, J. M., Ramos-Rodríguez, E. and Pérez-Martínez, C. (2014) 'In situ production of empty ephippia and resting eggs by an obligate parthenogenetic Daphnia population', *Journal of Plankton Research*, 36(1), pp. 157–169. doi: 10.1093/plankt/fbt072.
Cuenca Cambronero, M. and Orsini, L. (2018) 'Resurrection of Dormant Daphnia magna: Protocol and Applications', *Journal of Visualized Experiments*, (131), pp. 1–13. doi: 10.3791/56637.
Dai, D., Wang, Q., Li, X., Liu, J., Ma, X. and Xu, W. (2016) 'Klotho inhibits human follicular thyroid cancer cell growth and promotes apoptosis through regulation of the expression of stanniocalcin-1', *Oncology Reports*, 35(1), pp. 552–558. doi:

10.3892/or.2015.4358.
Davison, J. (1969) ‘Activation of the ephippial egg of Daphnia pulex.’, *The Journal of general physiology*, 53(5), pp. 562–575. doi: 10.1085/jgp.53.5.562.
Deng, H.-W. (1996) ‘Environmental and genetic control of sexual reproduction in Daphnia’, *Heredity*, 76(5), pp. 449–458. doi: 10.1038/hdy.1996.67.
Denlinger, D. L. (1986) ‘Dormancy in tropical insects’, *Annual Review of Entomology*, 31, pp. 239–264. doi: 10.1146/annurev.ento.31.1.239.
Denlinger, D. L. (2002) ‘Regulation of Diapause’, *Annual Review of Entomology*, 47, pp. 93–122.
Denlinger, D. L. and Armbruster, P. A. (2014) ‘Mosquito Diapause’, *Annual Review of Entomology*, 59(1), pp. 73–93. doi: 10.1146/annurev-ento-011613-162023.
Denlinger, D. L. and Armbruster, P. a (2014) ‘Mosquito diapause.’, *Annual review of entomology*, 59, pp. 73–93. doi: 10.1146/annurev-ento-011613-162023.
Denlinger, D. L., Yocum, G. D. and Rinehart, J. P. (2012) *10. Hormonal Control of Diapause*, *Insect Endocrinology*. Elsevier. doi: 10.1016/B978-0-12-384749-2.10010-X.
Dieterle, F., Ross, A., Schlotterbeck, G. and Senn, H. (2006) ‘Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in1H NMR metabonomics’, *Analytical Chemistry*, 78(13), pp. 4281–4290. doi: 10.1021/ac051632c.
Dmochowska-Ślęzak, K., Fliszkiewicz, M., Giejdasz, K. and Krystyna, Ż. (2015) ‘The antioxidant system in diapausing and active red mason bee Osmia bicornis’, *Entomology, Physiological*, pp. 82–89. doi: 10.1111/phen.12090.
Dong, Y., Desneux, N., Lei, C. and Niu, C. (2014) ‘Transcriptome characterization analysis of Bactrocera minax and new insights into its pupal diapause development with gene expression analysis.’, *International journal of biological sciences*, 10(9), pp. 1051–63. doi: 10.7150/ijbs.9438.
Draper, J., Lloyd, A. J., Goodacre, R. and Beckmann, M. (2013) ‘Flow infusion electrospray ionisation mass spectrometry for high throughput, non-targeted metabolite fingerprinting: A review’, *Metabolomics*, 9(SUPPL.1), pp. 4–29. doi: 10.1007/s11306-012-0449-x.
Drillet, G., Hansen, B. W. and Kiørboe, T. (2011) ‘Resting egg production induced by food limitation in the calanoid copepod Acartia tonsa’, *Limnology and Oceanography*, 56(6), pp. 2064–2070. doi: 10.4319/lo.2011.56.6.2064.
Duerr, J. M. and Podrabsky, J. E. (2010) ‘Mitochondrial physiology of diapausing and developing embryos of the annual killifish Austrofundulus limnaeus: Implications for extreme anoxia tolerance’, *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology*, 180(7), pp. 991–1003. doi: 10.1007/s00360-010-0478-6.
Dunn, W. B., Bailey, N. J. C. and Johnson, H. E. (2005) ‘Measuring the metabolome: Current analytical technologies’, *Analyst*, 130(5), pp. 606–625. doi: 10.1039/b418288j.
Dunn, W. B., Broadhurst, D., Begley, P., Zelena, E., Francis-Mcintyre, S., Anderson, N., Brown, M., Knowles, J. D., Halsall, A., Haselden, J. N., Nicholls, A. W., Wilson, I. D., Kell, D. B. and Goodacre, R. (2011) ‘Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry’, *Nature Protocols*, 6(7), pp. 1060–1083. doi: 10.1038/nprot.2011.335.

Durando, M., Tateishi, S. and Vaziri, C. (2013) 'A non-catalytic role of DNA polymerase g in recruiting Rad18 and promoting PCNA monoubiquitination at stalled replication forks', *Nucleic Acids Research*, 41(5), pp. 3079–3093. doi: 10.1093/nar/gkt016.
Durant, D. R., Berens, A. J., Toth, A. L. and Rehan, S. M. (2016) 'Transcriptional profiling of overwintering gene expression in the small carpenter bee, Ceratina calcarata', *Apidologie*, 47(4), pp. 572–582. doi: 10.1007/s13592-015-0402-x.
Ebert, D. (1991) 'The effect of size at birth, maturation threshold and genetic differences on the life-history of Daphnia magna', *Oecologia*, 86, pp. 243–250.
Ellers, J. and Van Alphen, J. J. M. (2002) 'A trade-off between diapause duration and fitness in female parasitoids', *Ecological Entomology*, 27(3), pp. 279–284. doi: 10.1046/j.1365-2311.2002.00421.x.
Erichson, N. B., Zeng, P., Manohar, K., Brunton, S. L., Kutz, J. N. and Aravkin, A. Y. (2018) 'Sparse Principal Component Analysis via Variable Projection', pp. 1–12. Available at: http://arxiv.org/abs/1804.00341.
Fan, L., Lin, J., Zhong, Y. and Liu, J. (2013a) 'Shotgun Proteomic Analysis on the Diapause and Non- Diapause Eggs of Domesticated Silkworm Bombyx mori', 8(4). doi: 10.1371/journal.pone.0060386.
Fan, L., Lin, J., Zhong, Y. and Liu, J. (2013b) 'Shotgun Proteomic Analysis on the Diapause and Non-Diapause Eggs of Domesticated Silkworm Bombyx mori', *PLoS ONE*, 8(4). doi: 10.1371/journal.pone.0060386.
Farkas, T., Nemecz, G. and Csengeri, I. (1984) 'Differential response of lipid metabolism and membrane physical state by an actively and passively overwintering planktonic crustacean', *Lipids*, 19(6), pp. 436–442. doi: 10.1007/BF02537405.
Feder, M. E. and Mitchell-Olds, T. (2003) 'Evolutionary and ecological functional genomics', 4(8), pp. 651–657. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12897776.
Feder, M. E. and Walser, J. C. (2005) 'The biological limitations of transcriptomics in elucidating stress and stress responses', *Journal of Evolutionary Biology*, 18(4), pp. 901–910. doi: 10.1111/j.1420-9101.2005.00921.x.
Feist, P. and Hummon, A. B. (2015) 'Proteomic challenges: Sample preparation techniques for Microgram-Quantity protein analysis from biological samples', *International Journal of Molecular Sciences*, 16(2), pp. 3537–3563. doi: 10.3390/ijms16023537.
Fenelon, J. C., Banerjee, A. and Murphy, B. D. (2014) 'Embryonic diapause: Development on hold', *International Journal of Developmental Biology*, 58(2–4), pp. 163–174. doi: 10.1387/ijdb.140074bm.
Fenelon, J. C., Lefèvre, P. L., Banerjee, A. and Murphy, B. D. (2017) 'Regulation of diapause in carnivores', *Reproduction in Domestic Animals*, 52, pp. 12–17. doi: 10.1111/rda.12835.
Ferrari, D. C. and Hebert, P. D. N. (1982) 'The induction of sexual reproduction in Daphnia magna: Genetic differences between arctic and temperate populations', *Canadian Journal of Zoology*, 60, pp. 2143–2148. doi: 10.1139/z82-274.
Fiehn, O. (2002) 'Metabolomics--the link between genotypes and phenotypes.', *Plant molecular biology*, 48(1–2), pp. 155–71. doi: 082/11.
Fielenbach, N. and Antebi, A. (2008) 'C. elegans dauer formation and the molecular basis of plasticity', *Genes & development*, 22(16), pp. 2149–65. doi:

10.1101/gad.1701508.
Fitzsimmons, J. M. and Innes, D. J. (2006) ‘Inter-genotype variation in reproductive response to crowding among Daphnia pulex’, *Hydrobiologia*, 568(1), pp. 187–205. doi: 10.1007/s10750-006-0104-5.
Flatt, T., Tu, M.-P. and Tatar, M. (2005) ‘Hormonal pleiotropy and the juvenile hormone regulation ofDrosophila development and life history’, *BioEssays*, 27(10), pp. 999–1010. doi: 10.1002/bies.20290.
Freudenthal, B. D., Gakhar, L., Ramaswamy, S. and Washington, T. S. (2010) ‘Structure of Monoubiquitinated PCNA and Implications for Translesion Synthesis and DNA Polymerase Exchange’, *Nat Struct Mol Biol.*, 17(4), pp. 479–484. doi: 10.1038/nsmb.1776.Structure.
Frisch, D., Morton, P. K., Chowdhury, P. R., Culver, B. W., Colbourne, J. K., Weider, L. J. and Jeyasingh, P. D. (2014) ‘A millennial-scale chronicle of evolutionary responses to cultural eutrophication in *Daphnia*’, *Ecology Letters*, 17(3), pp. 360–368. doi: 10.1111/ele.12237.
Fritsch, M. and Richter, S. (2015) ‘How the cladoceran heterogonic life cycle evolved-insights from gamogenetic reproduction and direct development in Cyclestherida’, *Evolution & Development*, 17(6), pp. 356–366. doi: 10.1111/ede.12163.
Gagliardi, A. D., Kuo, E. Y. W., Raulic, S., Wagner, G. F. and DiMattia, G. E. (2005) ‘Human stanniocalcin-2 exhibits potent growth-suppressive properties in transgenic mice independently of growth hormone and IGFs’, *American Journal of Physiology-Endocrinology and Metabolism*, 288(1), pp. E92–E105. doi: 10.1152/ajpendo.00268.2004.
Gerrish, G. a. and Cáceres, C. E. (2003) ‘Genetic versus environmental influence on pigment variation in the ephippia of Daphnia pulicaria’, *Freshwater Biology*, 48(11), pp. 1971–1982. doi: 10.1046/j.1365-2427.2003.01136.x.
Gertsman, I. and Barshop, B. A. (2018) ‘Promises and pitfalls of untargeted metabolomics’, *Journal of Inherited Metabolic Disease*. Journal of Inherited Metabolic Disease, 41(3), pp. 355–366. doi: 10.1007/s10545-017-0130-7.
Giansanti, P., Tsiatsiani, L., Low, T. Y. and Heck, A. J. R. (2016) ‘Six alternative proteases for mass spectrometry-based proteomics beyond trypsin’, *Nature Protocols*. Nature Publishing Group, 11(5), pp. 993–1006. doi: 10.1038/nprot.2016.057.
Giebelhausen, B. and Lampert, W. (2001) ‘Temperature reaction norms of Daphnia magna : the effect of food concentration’, *Freshwater Biology*, 46, pp. 281–289. doi: 10.1046/j.1365-2427.2001.00630.x.
Gilbert, J. J. (2003) ‘Environmental and endogenous control of sexuality in a rotifer life cycle: Developmental and population biology’, *Evolution and Development*, 5(1), pp. 19–24. doi: 10.1046/j.1525-142X.2003.03004.x.
Gilbert, J. J. (2004) ‘Females from resting eggs and parthenogenetic eggs in the rotifer Brachionus calyciflorus: Lipid droplets, starvation resistance and reproduction’, *Freshwater Biology*, 49(11), pp. 1505–1515. doi: 10.1111/j.1365-2427.2004.01282.x.
Glazier, D. S. (1992) ‘Effects of Food , Genotype , and Maternal Size and Age on Offspring Investment in Daphnia Magna Published by : Ecological Society of America Stable URL : http://www.jstor.org/stable/1940168 REFERENCES Linked references are available on JSTOR for this arti’, *Ecology*, 73(3), pp. 910–926.
Gliwicz, M. and Lampert, W. (1990) ‘Food Thresholds in Daphnia Species in the Absence and Presence of Blue-Green Filaments’, *Ecological Society of America*,

71(2), pp. 691–702.
Graveley, B. R., Brooks, A. N., Carlson, J. W., Duff, M. O., Landolin, J. M., Yang, L., Artieri, C. G., Van Baren, M. J., Boley, N., Booth, B. W., Brown, J. B., Cherbas, L., Davis, C. A., Dobin, A., Li, R., Lin, W., Malone, J. H., Mattiuzzo, N. R., Miller, D., Sturgill, D., Tuch, B. B., Zaleski, C., Zhang, D., Blanchette, M., Dudoit, S., Eads, B., Green, R. E., Hammonds, A., Jiang, L., Kapranov, P., Langton, L., Perrimon, N., Sandler, J. E., Wan, K. H., Willingham, A., Zhang, Y., Zou, Y., Andrews, J., Bickel, P. J., Brenner, S. E., Brent, M. R., Cherbas, P., Gingeras, T. R., Hoskins, R. A., Kaufman, T. C., Oliver, B. and Celniker, S. E. (2011) 'The developmental transcriptome of Drosophila melanogaster', *Nature*. Nature Publishing Group, 471(7339), pp. 473–479. doi: 10.1038/nature09715.
Guidetti, R., Altiero, T. and Rebecchi, L. (2011) 'On dormancy strategies in tardigrades', *Journal of Insect Physiology*. Elsevier Ltd, 57(5), pp. 567–576. doi: 10.1016/j.jinsphys.2011.03.003.
Gulbrandsen, J. and Johnsen, G. H. (1990) 'Temperature-dependent development of parthenogenetic embryos in Daphnia pulex de Geer', 12(3), pp. 443–453.
Guo, Q., Hao, Y.-J., Li, Y., Zhang, Y.-J., Ren, S., Si, F.-L. and Chen, B. (2015) 'Gene cloning, characterization and expression and enzymatic activities related to trehalose metabolism during diapause of the onion maggot Delia antiqua (Diptera: Anthomyiidae).', *Gene*. Elsevier B.V., 565(1), pp. 106–15. doi: 10.1016/j.gene.2015.03.070.
Guo, R., Gu, J., Zong, S., Wu, M. and Yang, M. (2018) 'Structure and mechanism of mitochondrial electron transport chain', *Biomedical Journal*. Elsevier Ltd, 41(1), pp. 9–20. doi: 10.1016/j.bj.2017.12.001.
Gyllström, M. and Hansson, L.-A. (2004) 'Dormancy in freshwater zooplankton: Induction, termination and the importance of benthic-pelagic coupling', *Aquatic Sciences*, 66(3), pp. 274–295. doi: 10.1007/s00027-004-0712-y.
Hahn, D. A. and Denlinger, D. L. (2007) 'Meeting the energetic demands of insect diapause: Nutrient storage and utilization', *Journal of Insect Physiology*, 53(8), pp. 760–773. doi: 10.1016/j.jinsphys.2007.03.018.
Hahn, D. A. and Denlinger, D. L. (2011) 'Energetics of Insect Diapause', *Annual Review of Entomology*, 56(1), pp. 103–121. doi: 10.1146/annurev-ento-112408-085436.
Hairston, N. G. J., Van Brunt, R. a, Kearns, C. M. and Engstron, D. R. (1995) 'Egg and survivorship of diapausing eggs in a sediment egg bank', *Ecology*, 76(6), pp. 1706–1711. doi: Doi 10.2307/1940704.
Han, J., Danell, R. M., Patel, J. R., Gumerov, D. R., Scarlett, C. O., Speir, J. P., Parker, C. E., Rusyn, I., Zeisel, S. and Borchers, C. H. (2008) 'Towards high-throughput metabolomics using ultrahigh-field Fourier transform ion cyclotron resonance mass spectrometry.', *Metabolomics : Official journal of the Metabolomic Society*, 4(2), pp. 128–140. doi: 10.1007/s11306-008-0104-8.
Han, X. and Gross, R. W. (2005) 'Shotgun lipidomics: Electrospray ionization mass spectrometric analysis and quantitation of cellular lipidomes directly from crude extracts of biological samples', *Mass Spectrometry Reviews*, 24(3), pp. 367–412. doi: 10.1002/mas.20023.
Hand, S. C., Denlinger, D. L., Podrabsky, J. E. and Roy, R. (2016) 'Mechanisms of animal diapause: recent developments from nematodes, crustaceans, insects, and fish', *American Journal of Physiology - Regulatory, Integrative and Comparative*

Physiology, 310(11), pp. R1193–R1211. doi: 10.1152/ajpregu.00250.2015.
Hand, S. C. and Podrabsky, J. E. (2000) 'Bioenergetics of diapause and quiescence in aquatic animals', 349.
Hansen, J. M. and Harris, C. (2015) 'Glutathione during embryonic development', *Biochimica et Biophysica Acta - General Subjects*. Elsevier B.V., 1850(8), pp. 1527–1542. doi: 10.1016/j.bbagen.2014.12.001.
Hanski, I. (1988) 'Four kinds of extra long diapause in insects: A review of theory and observations', *Ann. Zool. Fennici*, pp. 37–53. doi: 10.1016/j.ajog.2016.11.297.
Hao, Y. J., Zhang, Y. J., Si, F. L., Fu, D. Y., He, Z. B. and Chen, B. (2016) 'Insight into the possible mechanism of the summer diapause of Delia antiqua (Diptera: Anthomyiidae) through digital gene expression analysis', *Insect Science*, 23(3), pp. 438–451. doi: 10.1111/1744-7917.12323.
Hardie, D. G., Carling, D. and Halford, N. (1994) 'Roles of the Snf1/Rkin1/AMP-activated protein kinase family in the response to environmental and nutritional stress', *seminars in Cell Biology*, 5, pp. 409–416. Available at: https://www.sciencedirect.com/science/article/pii/S0166445X18305186.
Hasin, Y., Seldin, M. and Lusis, A. (2017) 'Multi-omics approaches to disease', *Genome Biology*. Genome Biology, 18(1), pp. 1–15. doi: 10.1186/s13059-017-1215-1.
Hayward, S. A. L. (2014) 'Application of functional "Omics" in environmental stress physiology: Insights, limitations, and future challenges', *Current Opinion in Insect Science*. Elsevier Inc, 4(1), pp. 35–41. doi: 10.1016/j.cois.2014.08.005.
Hayward, S. a L., Pavlides, S. C., Tammariello, S. P., Rinehart, J. P. and Denlinger, D. L. (2005) 'Temporal expression patterns of diapause-associated genes in flesh fly pupae from the onset of diapause through post-diapause quiescence', *Journal of Insect Physiology*, 51(6 SPEC. ISS.), pp. 631–640. doi: 10.1016/j.jinsphys.2004.11.009.
Hemmati, C., Moharramipour, S. and Talebi, A. A. (2017) 'Diapause Induced by Temperature and Photoperiod Affects Fatty Acid Compositions and Cold Tolerance of Phthorimaea Operculella (Lepidoptera: Gelechiidae)', *Environmental Entomology*, 46(6), pp. 1456–1463. doi: 10.1093/ee/nvx167.
Hengherr, S., Heyer, A. G., Brümmer, F. and Schill, R. O. (2011) 'Trehalose and Vitreous States: Desiccation Tolerance of Dormant Stages of the Crustaceans Triops and Daphnia', *Physiological and Biochemical Zoology*, 84(2), pp. 147–153. doi: 10.1086/658499.
Herman, W. S. (1973) 'The endocrine basis of reproductive inactivity in monarch butterflies overwintering in cenral California', *J. Insect Physiol.*, 19, pp. 1883–1887.
Herman, W. S. (1975) 'Endocrine regulation of posteclosion enlargement of the male and female reproductive gland in monarch butterflies.', *Gen. Comp. Endocrionl*, 26, pp. 534–540.
Hiruta, C. and Tochinai, S. (2014) 'Formation and structure of the ephippium (resting egg case) in relation to molting and egg laying in the water flea Daphnia pulex De Geer (Cladocera: Daphniidae)', *Journal of Morphology*, 275(7), pp. 760–767. doi: 10.1002/jmor.20255.
Hobaek, A. and Larsson, P. (1990) 'Sex determination in Daphnia magna.', *Ecology*, 71(6), pp. 2255–2268.
Hopper, A. K., Pai, D. A. and Engelke, D. R. (2010) 'Cellular dynamics of tRNAs and their genes', *FEBS Letters*. Federation of European Biochemical Societies, 584(2),

pp. 310–317. doi: 10.1016/j.febslet.2009.11.053.
Houthoofd, K., Braeckman, B. P., Lenaerts, I., Brys, K., Vreese, A. De, Eygen, S. Van and Vanfleteren, J. R. (2002) 'Ageing is reversed , and metabolism is reset to young levels in recovering dauer larvae of C . elegans', 37, pp. 1015–1021.
Ignace, D. D., Dodson, S. I. and Kashian, D. R. (2011) 'Identification of the critical timing of sex determination in Daphnia magna (Crustacea, Branchiopoda) for use in toxicological studies', *Hydrobiologia*, 668(1), pp. 117–123. doi: 10.1007/s10750-010-0534-y.
Innes, D. J., Fox, C. J. and Winsor, G. L. (2000) 'Avoiding the cost of males in obligately asexual Daphnia pulex (Leydig)', *Proceedings of the Royal Society B: Biological Sciences*, 267(1447), pp. 991–997. doi: 10.1098/rspb.2000.1101.
Jackson, R. J., Hellen, C. U. T. and Pestova, T. V. (2010) 'The Mechanism Of Eucaryotic Translation Initiation And Principles Of Its Regulation', *Nat Rev Mol Cell Biol.*, 11(2), pp. 113–127. doi: 10.1038/nrm2838.THE.
Jemec, A., Drobne, D., Tišler, T., Trebše, P., Roš, M. and Sepčić, K. (2007) 'The applicability of acetylcholinesterase and glutathione S-transferase in Daphnia magna toxicity test', *Comparative Biochemistry and Physiology - C Toxicology and Pharmacology*, 144(4), pp. 303–309. doi: 10.1016/j.cbpc.2006.10.002.
Jena, K., Kar, P. K., Babu, C. S., Giri, S., Singh, S. S. and Prasad, B. C. (2013) 'Comparative Study of Total Hydroperoxides and Antioxidant Defense System in the Indian Tropical Tasar Silkworm, *Antheraea mylitta* , in Diapausing and Non-Diapausing Generations', *Journal of Insect Science*, 13(123), pp. 1–11. doi: 10.1673/031.013.12301.
Jolliffe, I. T. (2002) *Principal Component Analysis*. doi: 10.2307/1270093.
Jordão, R., Casas, J., Fabrias, G., Campos, B., Piña, B., Lemos, M. F. L., Soares, A. M. V. M., Tauler, R. and Barata, C. (2015) 'Obesogens beyond vertebrates: Lipid perturbation by tributyltin in the crustacean Daphnia magna', *Environmental Health Perspectives*, 123(8), pp. 813–819. doi: 10.1289/ehp.1409163.
Jovanović-Galović, A., Blagojević, D. P., Grubor-Lajšić, G., Worland, R. and Spasić, M. B. (2004) 'Role of antioxidant defense during different stages of preadult life cycle in European corn borer (Ostrinia nubilalis, Hubn.): Diapause and metamorphosis', *Archives of Insect Biochemistry and Physiology*, 55(2), pp. 79–89. doi: 10.1002/arch.10126.
Kang, D. S., Cotten, M. A., Denlinger, D. L. and Sim, C. (2016) 'Comparative transcriptomics reveals key gene expression differences between diapausing and Non-Diapausing Adults of Culex pipiens', *PLoS ONE*, 11(4), pp. 1–14. doi: 10.1371/journal.pone.0154892.
Kankare, M., Parker, D. J., Merisalo, M., Salminen, T. S. and Hoikkala, A. (2016) 'Transcriptional differences between diapausing and non-diapausing D. Montana females reared under the same photoperiod and temperature', *PLoS ONE*, 11(8), pp. 1–18. doi: 10.1371/journal.pone.0161852.
Kato, Y., Shiga, Y., Kobayashi, K., Tokishita, S., Yamagata, H., Iguchi, T. and Watanabe, H. (2011) 'Development of an RNA interference method in the cladoceran crustacean Daphnia magna.', *Development genes and evolution*, 220(11–12), pp. 337–45. doi: 10.1007/s00427-011-0353-9.
Kato, Y., Tokishita, S., Ohta, T. and Yamagata, H. (2004) 'A vitellogenin chain containing a superoxide dismutase-like domain is the major component of yolk proteins in cladoceran crustacean Daphnia magna.', *Gene*, 334, pp. 157–65. doi:

10.1016/j.gene.2004.03.030.
Kaupinis, A., Aitmanaitė, L., Strepetkaitė, D., Valius, M., Lazutka, J. R. and Arbačiauskas, K. (2016) 'Proteomic and gene expression differences between post-diapause and subitaneous offspring phenotypes in the cyclic parthenogen Daphnia pulex', *Hydrobiologia*, pp. 1–17. doi: 10.1007/s10750-016-3057-3.
Kaupinis, A., Aitmanaitė, L., Strepetkaitė, D., Valius, M., Lazutka, J. R. and Arbačiauskas, K. (2017) 'Proteomic and gene expression differences between post-diapause and subitaneous offspring phenotypes in the cyclic parthenogen Daphnia pulex', *Hydrobiologia*, 798(1), pp. 87–103. doi: 10.1007/s10750-016-3057-3.
Kawasaki, T., Sakata, M. and Namiki, H. (2004) 'Elemental characterization of Daphnia resting eggs by X-ray analytical microscopy.', *Zoological science*, 21(10), pp. 1019–1023. doi: 10.2108/zsj.21.1019.
Keilin, D. (1959) 'The Leeuwenhoek Lecture - The problem of anabiosis or latent life: history and current concept', *Proceedings of the Royal Society of London. Series B - Biological Sciences*, 150(939), pp. 149–191. doi: 10.1098/rspb.1959.0013.
Kim, H.-J., Suga, K., Kim, B.-M., Rhee, J.-S., Lee, J.-S. and Hagiwara, A. (2015) 'Light-dependent transcriptional events during resting egg hatching of the rotifer Brachionus manjavacas', *Marine Genomics*. Elsevier B.V., 20, pp. 25–31. doi: 10.1016/j.margen.2015.02.003.
King, A. M., Toxopeus, J. and Macrae, T. H. (2014) 'Artemin, a diapause-specific chaperone, contributes to the stress tolerance of Artemia cysts and influences their release from females', *The Journal of Experimental Biology*, pp. 1719–1724. doi: 10.1242/jeb.100081.
King, R. W., Deshaies, R. J., Peters, J. and Kirschner, M. W. (1996) 'How Proteolysis Drive the Cell Cycle', *Science*, 274(5293), pp. 1652–1659.
Klüttgen, B., Dülmer, U., Engels, M. and Ratte, H. T. (1994) 'ADaM, an artificial freshwater for the culture of zooplankton', *Water Research*, 28(3), pp. 743–746. doi: 10.1016/0043-1354(94)90157-0.
Koch, U., von Elert, E. and Straile, D. (2009) 'Food quality triggers the reproductive mode in the cyclical parthenogen Daphnia (Cladocera)', *Oecologia*, 159(2), pp. 317–324. doi: 10.1007/s00442-008-1216-6.
Koštál, V. (2006) 'Eco-physiological phases of insect diapause', *Journal of Insect Physiology*, 52(2), pp. 113–127. doi: 10.1016/j.jinsphys.2005.09.008.
Koštál, V., Šimůnková, P., Kobelková, A. and Shimada, K. (2009) 'Cell cycle arrest as a hallmark of insect diapause: Changes in gene transcription during diapause induction in the drosophilid fly, Chymomyza costata', *Insect Biochemistry and Molecular Biology*, 39(12), pp. 875–883. doi: 10.1016/j.ibmb.2009.10.004.
Koštál, V., Štětina, T., Poupardin, R., Korbelová, J. and Bruce, A. W. (2017) 'Conceptual framework of the eco-physiological phases of insect diapause development justified by transcriptomic profiling', *Proceedings of the National Academy of Sciences*, 114(32), p. 201707281. doi: 10.1073/pnas.1707281114.
Kotov, A. A. and Boikova, O. S. (2001) 'Study of the late embryogenesis of Daphnia (Anomopoda , " Cladocera ", Branchiopoda) and a comparison of development in Anomopoda and Ctenopoda', (1925), pp. 127–143.
Kubrak, O. I., Kučerová, L., Theopold, U. and Nässel, D. R. (2014) 'The Sleeping Beauty: How Reproductive Diapause Affects Hormone Signaling, Metabolism, Immune Response and Somatic Maintenance in Drosophila melanogaster.', *PloS one*, 9(11), p. e113051. doi: 10.1371/journal.pone.0113051.

Kučerová, L., Kubrak, O. I., Bengtsson, J. M., Strnad, H., Nylin, S., Theopold, U. and Nässel, D. R. (2016) 'Slowed aging during reproductive dormancy is reflected in genome-wide transcriptome changes in Drosophila melanogaster', *BMC Genomics*, 17(1), p. 50. doi: 10.1186/s12864-016-2383-1.
Kučerová, L., Kubrak, O. I., Bengtsson, J. M., Strnad, H., Nylin, S., Theopold, U. and Nässel, D. R. (2016) 'Slowed aging during reproductive dormancy is reflected in genome-wide transcriptome changes in Drosophila melanogaster', *BMC Genomics*, 17(1), pp. 1–25. doi: 10.1186/s12864-016-2383-1.
Kuthiala, A. and Chippendale, G. M. (1989) 'Relationship between the fatty acids of fat body triacylglycerol and lipophorin diacylglycerol of non-diapause and diapause larvae of the southwestern corn borer, Diatraea grandiosella', *Archives of Insect Biochemistry and Physiology*, 12(2), pp. 123–131. doi: 10.1002/arch.940120206.
Lampert, W., Lampert, K. P. and Larsson, P. (2012) 'Induction of male production in clones of Daphnia pulex by the juvenoid hormone methyl farnesoate under short photoperiod', *Comparative Biochemistry and Physiology - C Toxicology and Pharmacology*. Elsevier Inc., 156(2), pp. 130–133. doi: 10.1016/j.cbpc.2012.05.004.
Latta IV, L. C., Frederick, S. and Pfrender, M. E. (2011) 'Diet Restriction and Life-History Trade-Offs in Short- and Long- Lived Species of Daphnia', *J Exp Zool A Ecol Genet Physiol.*, 315(A)(10), pp. 610–617. doi: 10.1037/a0017530.A.
LeBlanc, G. A. and Cochrane, B. J. (1987) 'Identification of multiple glutathione S-transferases from Daphnia magna', *Comparative Biochemistry and Physiology -- Part B: Biochemistry and*, 88(1), pp. 39–45. doi: 10.1016/0305-0491(87)90076-9.
Lecker, S. H. (2006) 'Protein Degradation by the Ubiquitin-Proteasome Pathway in Normal and Disease States', *Journal of the American Society of Nephrology*, 17(7), pp. 1807–1819. doi: 10.1093/nq/s9-VIII.206.466.
Lee, A. S. Y., Kranzusch, P. J. and Cate, J. H. D. (2015) 'EIF3 targets cell-proliferation messenger RNAs for translational activation or repression', *Nature*, 522(7554), pp. 111–114. doi: 10.1038/nature14267.
Lehmann, P., Pruisscher, P., Agren, R., Va, L., Wiklund, C., Wheat, C. W. and Gotthard, K. (2018) 'Metabolome dynamics of diapause in the butterfly Pieris napi : distinguishing maintenance , termination and post-diapause phases'. doi: 10.1242/jeb.169508.
Lehmann, P., Pruisscher, P., Posledovich, D., Carlsson, M., Käkelä, R., Tang, P., Nylin, S., Wheat, C. W., Wiklund, C. and Gotthard, K. (2016) 'Energy and lipid metabolism during direct and diapause development in a pierid butterfly', *The Journal of Experimental Biology*, 219(19), pp. 3049–3060. doi: 10.1242/jeb.142687.
Lei, Z., Huhman, D. V and Sumner, L. W. (2011) 'Mass spectrometry strategies in metabolomics.', *The Journal of biological chemistry*, 286(29), pp. 25435–42. doi: 10.1074/jbc.R111.238691.
Lessman, C. A. and Herman, W. S. (1983) 'Seasonal variation in hmolymph juvenile hormone of adult monarchs (Danaus p.plexippus: Lepidoptera)', *Canadian journal of Zoology*, 61, pp. 88–94.
Lessman, C. A., Rollins, L. and Herman, W. S. (1982) 'Effects of juvenile hormones I,II and III on reproductive tract growth in male and female monarch butterflies.', *Comp. Biochem. Physiol.*, A 71, pp. 141–144.
Li, H., Yuan, S., Su, G., Li, M., Wang, Q. and Zhu, G. (2017) 'Whole-Life-Stage Characterization in the Basic Biology of Daphnia magna and E ff ects of TDCIPP on Growth, Reproduction, Survival, and Transcription of Genes'. doi:

10.1021/acs.est.7b04569.
Li, N. and Oberley, T. D. (1998) 'Modulation of antioxidant enzymes, reactive oxygen species, and glutathione levels in manganese superoxide dismutase- overexpressing NIH/3T3 fibroblasts during the cell cycle', *Journal of Cellular Physiology*, 177(1), pp. 148–160.
Li, Y., Zhang, L., Chen, H., Kostal, V., Simek, P., Moos, M. and Denlinger, D. L. (2015) 'Shifts in metabolomic profiles of the parasitoid Nasonia vitripennis associated with elevated cold tolerance induced by the parasitoid's diapause, host diapause and host diet augmented with proline', *Insect Biochemistry and Molecular Biology*, 63(May), pp. 1–13. doi: 10.1016/j.ibmb.2015.05.012.
Liu, W., Tan, Q., Zhu, L., Li, Y., Zhu, F., Lei, C. and Wang, X. (2017) 'Absence of juvenile hormone signalling regulates the dynamic expression profiles of nutritional metabolism genes during diapause preparation in the cabbage beetle Colaphellus bowringi', 26, pp. 530–542. doi: 10.1111/imb.12316.
Liu, Y. Q., Zheng, X. X., Ma, H. F., Xia, R. X., Li, Y. P. and Zhang, Q. R. (2016) 'Supercooling Capacity and Cold Tolerance of the Wild Silkworm, Antheraea pernyi (Lepidoptera: Saturniidae)', *Journal of Economic Entomology*, 109(4), pp. 1619–1627. doi: 10.1093/jee/tow121.
Liu, Z., Cai, M., Yu, P., Chen, M., Wu, D., Zhang, M. and Zhao, Y. (2018) 'Age-dependent survival, stress defense, and AMPK in Daphnia pulex after short-term exposure to a polystyrene nanoplastic', *Aquatic Toxicology*. Elsevier, 204(August), pp. 1–8. doi: 10.1016/J.AQUATOX.2018.08.017.
Love, M. I., Huber, W. and Anders, S. (2014) 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biology*, 15(12), p. 550. doi: 10.1186/s13059-014-0550-8.
Lu, Y. and Xu, W. (2010) 'Proteomic and Phosphoproteomic Analysis at Diapause Initiation in the Cotton Bollworm , Helicoverpa armigera research articles', *Journal of proteome research*, pp. 5053–5064.
Lu, Y., Zhang, Q. and Xu, W. (2014) 'Global Metabolomic Analyses of the Hemolymph and Brain during the Initiation , Maintenance , and Termination of Pupal Diapause in the Cotton Bollworm , Helicoverpa armigera', 9(6). doi: 10.1371/journal.pone.0099948.
Lürling, M., Roozen, F., Van Donk, E. and Goser, B. (2003) 'Response of Daphnia to substances released from crowded congeners and conspecifics', *Journal of Plankton Research*, 25(8), pp. 967–978. doi: 10.1093/plankt/25.8.967.
Lyu, K., Zhu, X., Chen, R., Chen, Y. and Yang, Z. (2014) 'Molecular cloning of manganese superoxide dismutase gene in the cladoceran Daphnia magna: Effects of microcystin, nitrite, and cadmium on gene expression profiles', *Aquatic Toxicology*. Elsevier B.V., 148, pp. 55–64. doi: 10.1016/j.aquatox.2013.12.031.
Lyu, K., Zhu, X., Wang, Q., Chen, Y. and Yang, Z. (2013) 'Copper/Zinc Superoxide Dismutase from the Cladoceran *Daphnia magna* : Molecular Cloning and Expression in Response to Different Acute Environmental Stressors', *Environmental Science & Technology*, p. 130710143609005. doi: 10.1021/es4015212.
Macel, M., van dam, N. M. and Keurentjes, J. J. B. (2010) 'Metabolomics: The chemistry between ecology and genetics', *Molecular Ecology Resources*, 10(4), pp. 583–593. doi: 10.1111/j.1755-0998.2010.02854.x.
MacRae, T. H. (2003) 'Molecular chaperones, stress resistance and development in Artemia franciscana', *Seminars in Cell and Developmental Biology*, 14(5), pp. 251–

258. doi: 10.1016/j.semcdb.2003.09.019.
MacRae, T. H. (2010) 'Gene expression, metabolic regulation and stress tolerance during diapause', *Cellular and Molecular Life Sciences*, 67(14), pp. 2405–2424. doi: 10.1007/s00018-010-0311-0.
MacRae, T. H. (2016) 'Stress tolerance during diapause and quiescence of the brine shrimp, Artemia', *Cell Stress and Chaperones*, pp. 9–18. doi: 10.1007/s12192-015-0635-7.
Maga, G. (2003) 'Proliferating cell nuclear antigen (PCNA): a dancer with many partners', *Journal of Cell Science*, 116(15), pp. 3051–3060. doi: 10.1242/jcs.00653.
Magalhães, R. S. S., De Lima, K. C., de Almeida, D. S. G., De Mesquita, J. F. and Eleutherio, E. C. A. (2017) 'Trehalose-6-Phosphate as a Potential Lead Candidate for the Development of Tps1 Inhibitors: Insights from the Trehalose Biosynthesis Pathway in Diverse Yeast Species', *Applied Biochemistry and Biotechnology*. Applied Biochemistry and Biotechnology, 181(3), pp. 914–924. doi: 10.1007/s12010-016-2258-6.
De Meester, L. (1993) 'Inbreeding and Outbreeding Depression in Daphnia Oecologia Inbreeding and outbreeding depression in Daphnia', *Oecologia*, 96(1), pp. 80–84.
Meller, C. L., Meller, R., Simon, R. P., Culpepper, K. M. and Podrabsky, J. E. (2012a) 'Cell cycle arrest associated with anoxia-induced quiescence, anoxic preconditioning, and embryonic diapause in embryos of the annual killifish Austrofundulus limnaeus', *J Comp Physiol B.*, 182(7), pp. 909–920. doi: 10.1007/s00360-012-0672-9.Cell.
Meller, C. L., Meller, R., Simon, R. P., Culpepper, K. M. and Podrabsky, J. E. (2012b) 'Cell cycle arrest associated with anoxia-induced quiescence, anoxic preconditioning, and embryonic diapause in embryos of the annual killifish Austrofundulus limnaeus', *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology*, 182(7), pp. 909–920. doi: 10.1007/s00360-012-0672-9.
Méndez, J. and Stillman, B. (2000) 'Chromatin association of human origin recognition complex, cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis.', *Molecular and cellular biology*, 20(22), pp. 8602–12. doi: 10.1128/MCB.20.22.8602-8612.2000.Updated.
Meng, G., Yao, J., Wang, L. and Zhao, L. (2011) 'Variation in glutathione status associated with induction and initiation of diapause in eggs of the bivoltine strain of the silkworm Bombyx mori', *Physiological Entomology*, 36(2), pp. 173–179. doi: 10.1111/j.1365-3032.2011.00783.x.
Menon, S. G. and Goswami, P. C. (2007) 'A redox cycle within the cell cycle: Ring in the old with the new', *Oncogene*, 26(8), pp. 1101–1109. doi: 10.1038/sj.onc.1209895.
Mesak, F., Tatarenkov, A. and Avise, J. C. (2015) 'Transcriptomics of diapause in an isogenic self-fertilizing vertebrate', *BMC Genomics*. BMC Genomics, 16(1), p. 989. doi: 10.1186/s12864-015-2210-0.
Mi, H., Huang, X., Muruganujan, A., Tang, H., Mills, C., Kang, D. and Thomas, P. D. (2017) 'PANTHER version 11: Expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements', *Nucleic Acids Research*, 45(D1), pp. D183–D189. doi: 10.1093/nar/gkw1138.
Mi, H., Muruganujan, A., Casagrande, J. T. and Thomas, P. D. (2013) 'Large-scale gene function analysis with the panther classification system', *Nature Protocols*, 8(8), pp. 1551–1566. doi: 10.1038/nprot.2013.092.

Miao, Y., Tenor, J. L., Toffaletti, D. L., Washington, E. J., Liu, J., Shadrick, W. R., Schumacher, M. A., Lee, R. E., Perfect, J. R. and Brennan, R. G. (2016) 'Structures of trehalose-6-phosphate phosphatase from pathogenic fungi reveal the mechanisms of substrate recognition and catalysis', *Proceedings of the National Academy of Sciences*, 113(26), pp. 7148–7153. doi: 10.1073/pnas.1601774113.
Michaud, M. R. and Denlinger, D. L. (2007) 'Shifts in the carbohydrate, polyol, and amino acid pools during rapid cold-hardening and diapause-associated cold-hardening in flesh flies (Sarcophaga crassipalpis): A metabolomic comparison', *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology*, 177(7), pp. 753–763. doi: 10.1007/s00360-007-0172-5.
Mitsumasu, K., Azuma, M., Niimi, T., Yamashita, O. and Yaginuma, T. (2008) 'Changes in the expression of soluble and integral-membrane trehalases in the midgut during metamorphosis in Bombyx mori.', *Zoological science*, 25(7), pp. 693–8. doi: 10.2108/zsj.25.693.
Mittmann, B., Ungerer, P., Klann, M., Stollewerk, A. and Wolff, C. (2014) 'Development and staging of the water flea Daphnia magna (Straus, 1820; Cladocera, Daphniidae) based on morphological landmarks', *EvoDevo*. EvoDevo, 5(1), p. 12. doi: 10.1186/2041-9139-5-12.
Miyazaki, S., Kikuchi, H., Iino, I., Uehara, T., Setoguchi, T., Fujita, T., Hiramatsu, Y., Ohta, M., Kamiya, K., Kitagawa, K., Kitagawa, M., Baba, S. and Konno, H. (2014) 'Anti-VEGF antibody therapy induces tumor hypoxia and stanniocalcin 2 expression and potentiates growth of human colon cancer xenografts', *International Journal of Cancer*, 135(2), pp. 295–307. doi: 10.1002/ijc.28686.
Moribe, Y., Oka, K., Niimi, T., Yamashita, O. and Yaginuma, T. (2010) 'Expression of heat shock protein 70a mRNA in Bombyx mori diapause eggs.', *Journal of insect physiology*. Elsevier Ltd, 56(9), pp. 1246–52. doi: 10.1016/j.jinsphys.2010.03.023.
Muñoz-Mejía, G. and Martínez-Jerónimo (2007) 'Impact of algae and their concentrations on the reproduction and longevity of cladocerans', *Annales de Limnologie - International Journal of Limnology*, 43(3), pp. 167–177. doi: 10.1051/limn:2007011.
Murphy, S. M. and Lill, J. T. (2010) 'Winter predation of diapausing cocoons of slug caterpillars (Lepidoptera: Limacodidae).', *Environmental entomology*, 39(6), pp. 1893–902. doi: 10.1603/EN10094.
Nakagaki, M., Takei, R., Nagashima, E. and Yaginuma, T. (1991) 'Cell cycles in embryos of the silkworm, Bombyx mori: G2-arrest at diapause stage', *Roux's Archives of Developmental Biology*, 200(4), pp. 223–229. doi: 10.1007/BF00361341.
Nakanishi, T., Kato, Y., Matsuura, T. and Watanabe, H. (2014) 'CRISPR/Cas-mediated targeted mutagenesis in Daphnia magna', *PLoS ONE*, 9(5). doi: 10.1371/journal.pone.0098363.
Navis, S., Waterkeyn, A., Putman, A., De Meester, L., Vanermen, G. and Brendonck, L. (2015) 'Timing matters: Sensitivity of Daphnia magna dormant eggs to fenoxycarb exposure depends on embryonic developmental stage', *Aquatic Toxicology*. Elsevier B.V., 159, pp. 176–183. doi: 10.1016/j.aquatox.2014.12.016.
Navis, S., Waterkeyn, A., Voet, T., De Meester, L. and Brendonck, L. (2013) 'Pesticide exposure impacts not only hatching of dormant eggs, but also hatchling survival and performance in the water flea Daphnia magna', *Ecotoxicology*, 22(5), pp. 803–814. doi: 10.1007/s10646-013-1080-y.
Nguyen, T. V., Jung, H., Rotllant, G., Hurwood, D., Mather, P. and Ventura, T. (2018)

‘Guidelines for RNA-seq projects: applications and opportunities in non-model decapod crustacean species’, *Hydrobiologia*. Springer International Publishing, pp. 1–23. doi: 10.1007/s10750-018-3682-0.
Nie, L., Wu, G., Culley, D. E., Scholten, J. C. M. and Zhang, W. (2007) ‘Integrative analysis of transcriptomic and proteomic data: Challenges, solutions and applications’, *Critical Reviews in Biotechnology*, 27(2), pp. 63–75. doi: 10.1080/07388550701334212.
Noctor, G. and Foyer, C. H. (1998) ‘ASCORBATE AND GLUTATHIONE: Keeping Active Oxygen Under Control’, *Annual Review of Plant Physiology and Plant Molecular Biology*, 49(1), pp. 249–279. doi: 10.1146/annurev.arplant.49.1.249.
Oda, S., Tatarazako, N., Watanabe, H., Morita, M. and Iguchi, T. (2006) ‘Genetic differences in the production of male neonates in Daphnia magna exposed to juvenile hormone analogs.’, *Chemosphere*, 63(9), pp. 1477–84. doi: 10.1016/j.chemosphere.2005.09.026.
Ohtsu, T., Kimura, M. T. and Hori, S. H. (1992) ‘Energy storage during reproductive diapause in the *Drosophila melanogaster* species group’, *Journal of Comparative Physiology B*, 162(3), pp. 203–208.
Olmstead, A. W. and Leblanc, G. a (2002) ‘Juvenoid hormone methyl farnesoate is a sex determinant in the crustacean Daphnia magna.’, *The Journal of experimental zoology*, 293(7), pp. 736–9. doi: 10.1002/jez.10162.
Orr, W. C. and Sohal, R. S. (1994) ‘Extension of life-span by overexpression of superoxide dismutase and catalse in Drosophila melanogaster’, *Science*, 263(February), pp. 1128–1130.
Orsini, L., Gilbert, D., Podicheti, R., Jansen, M. and Brown, J. B. (2016) ‘Data Descriptor : Daphnia magna transcriptome by RNA-Seq across 12 environmental stressors’, (May), pp. 1–16.
Orsini, L., Spanier, K. I. and De Meester, L. (2012) ‘Genomic signature of natural and anthropogenic stress in wild populations of the waterflea Daphnia magna: Validation in space, time and experimental evolution’, *Molecular Ecology*, 21(9), pp. 2160–2175. doi: 10.1111/j.1365-294X.2011.05429.x.
Padilla, P. a. and Ladage, M. L. (2012) ‘Suspended animation, diapause and quiescence: Arresting the cell cycle in C. elegans’, *Cell Cycle*, 11(9), pp. 1672–1679. doi: 10.4161/cc.19444.
Paes, T. A. S. V., Rietzler, A. C. and Maia-Barbosa, P. M. (2016) ‘Methods for selection of Daphnia resting eggs: the influence of manual decapsulation and sodium hypoclorite solution on hatching rates’, *Brazilian Journal of Biology*, 76(4), pp. 1058–1063. doi: 10.1590/1519-6984.09415.
Paolucci, S., Van de Zande, L. and Beukeboom, L. W. (2013) ‘Adaptive latitudinal cline of photoperiodic diapause induction in the parasitoid Nasonia vitripennis in Europe’, *Journal of Evolutionary Biology*, 26(4), pp. 705–718. doi: 10.1111/jeb.12113.
Parsons, H. M., Ekman, D. R., Collette, T. W. and Viant, M. R. (2009) ‘Spectral relative standard deviation: A practical benchmark in metabolomics’, *Analyst*, 134(3), pp. 478–485. doi: 10.1039/b808986h.
Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., Kingsford, C. and Biology, C. (2017) ‘Salmon: fast and bias-aware quantification of transcript expression using dual-phase inference’, *Nature Methods*, 14(4), pp. 417–419. doi: 10.1038/nmeth.4197.Salmon.
Paul, S. (2002) ‘Effects of photoperiod, temperature, and host plant age on induction of reproductive diapause and development time in Danaus plexippus - Goehring -

2002 - Ecological Entomology - Wiley Online Library', pp. 674–685. Available at: http://0-onlinelibrary.wiley.com.libraries.colorado.edu/doi/10.1046/j.1365-2311.2002.00454.x/abstract.
Pavlides, S. C., Pavlides, S. A. and Tammariello, S. P. (2011) 'Proteomic and phosphoproteomic profiling during diapause entrance in the flesh fly, Sarcophaga crassipalpis', *Journal of Insect Physiology*. Elsevier Ltd, 57(5), pp. 635–644. doi: 10.1016/j.jinsphys.2011.03.022.
Peña, C., Hurt, E. and Panse, V. G. (2017) 'Eukaryotic ribosome assembly, transport and quality control', *Nature Structural and Molecular Biology*. Nature Publishing Group, 24(9), pp. 689–699. doi: 10.1038/nsmb.3454.
Perez-Hedo, M., Sanchez-Lopez, I. and Eizaguirre, M. (2012) 'Comparative analysis of hemolymph proteome maps in diapausing and non-diapausing larvae of sesamia nonagrioides', *Proteome Science*, 10(1), p. 58. doi: 10.1186/1477-5956-10-58.
Pietrzak, B., Bednarska, A. and Grzesiuk, M. (2010) 'Longevity of Daphnia magna males and females', *Hydrobiologia*, 643(1), pp. 71–75. doi: 10.1007/s10750-010-0138-6.
Ploomi, A., Kuusik, A., Jõgar, K., Metspalu, L., Hiiesaar, K., Karise, R., Williams, I. H., Sibul, I. and Mänd, M. (2018) 'Variability in metabolic rate and gas exchange patterns of the Colorado potato beetle of winter and prolonged diapauses', *Physiological Entomology*, 43(4), pp. 251–258. doi: 10.1111/phen.12254.
Podrabsky, J. E. and Culpepper, K. M. (2012) 'Cell cycle regulation during development and dormancy in embryos of the annual killi fi sh Austrofundulus limnaeus', *cell cycle*, 11:9, pp. 1697–1704.
Podrabsky, J. E. and Hand, S. C. (2000) 'Depression of protein synthesis during diapause in embryos of the annual killifish Austrofundulus limnaeus.', *Physiological and biochemical zoology : PBZ*, 73(6), pp. 799–808. doi: 10.1086/318106.
Podrabsky, J. E. and Hand, S. C. (2015) 'Physiological strategies during animal diapause: lessons from brine shrimp and annual killifish', *Journal of Experimental Biology*, 218(12), pp. 1897–1906. doi: 10.1242/jeb.116194.
Poelchau, M. F., Reynolds, J. A., Denlinger, D. L., Elsik, C. G. and Armbruster, P. A. (2013) 'Transcriptome sequencing as a platform to elucidate molecular components of the diapause response in the Asian tiger mosquito Aedes albopictus', *Physiological Entomology*, 38(2), pp. 173–181. doi: 10.1111/phen.12016.
Poelchau, M. F., Reynolds, J. a, Elsik, C. G., Denlinger, D. L. and Armbruster, P. a (2013a) 'Deep sequencing reveals complex mechanisms of diapause preparation in the invasive mosquito, Aedes albopictus', *Proceedings of the Royal Society B: Biological Sciences*, 280, p. 20130143. doi: http://dx.doi.org/10.1098/rspb.2013.0143.
Poelchau, M. F., Reynolds, J. a, Elsik, C. G., Denlinger, D. L. and Armbruster, P. a (2013b) 'RNA-Seq reveals early distinctions and late convergence of gene expression between diapause and quiescence in the Asian tiger mosquito, Aedes albopictus.', *The Journal of experimental biology*, 216(Pt 21), pp. 4082–90. doi: 10.1242/jeb.089508.
Poupardin, R., Schöttner, K., Korbelová, J., Provazník, J., Doležel, D., Pavlinic, D., Beneš, V. and Koštál, V. (2015) 'Early transcriptional events linked to induction of diapause revealed by RNAseq in larvae of drosophilid fly, Chymomyza costata', *BMC Genomics*, 16(1), p. 720. doi: 10.1186/s12864-015-1907-4.
Poynton, H. C., Taylor, N. S., Hicks, J., Colson, K., Chan, S., Clark, C., Scanlan, L., Loguinov, A. V, Vulpe, C. and Viant, M. R. (2011) 'Metabolomics of microliter

hemolymph samples enables an improved understanding of the combined metabolic and transcriptional responses of Daphnia magna to cadmium.', *Environmental science & technology*, 45(8), pp. 3710–7. doi: 10.1021/es1037222.
Poynton, H. C., Varshavsky, J. R., Chang, B., Cavigiolio, G., Chan, S., Holman, P. S., Loguinov, A. V, Bauer, D. J., Komachi, K., Theil, E. C., Perkins, E. J., Hughes, O. and Vulpe, C. D. (2007) 'Daphnia magna ecotoxicogenomics provides mechanistic insights into metal toxicity.', *Environmental science & technology*, 41(3), pp. 1044–50. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17328222.
Qi, X., Zhang, L., Han, Y., Ren, X., Huang, J. and Chen, H. (2015) 'De novo transcriptome sequencing and analysis of Coccinella septempunctata L. in non-diapause, diapause and diapause-terminated states to identify diapause-associated genes', *BMC Genomics*. BMC Genomics, 16(1), p. 1086. doi: 10.1186/s12864-015-2309-3.
Radzikowski, J. (2013) 'Resistance of dormant stages of planktonic invertebrates to adverse environmental conditions', *Journal of Plankton Research*, 35(4), pp. 707–723. doi: 10.1093/plankt/fbt032.
Ragland, G. J., Denlinger, D. L. and Hahn, D. A. (2010) 'Mechanisms of suspended animation are revealed by transcript profiling of diapause in the flesh fly', *Proceedings of the National Academy of Sciences*, 107(33), pp. 14909–14914. doi: 10.1073/pnas.1007075107.
Ragland, G. J. and Keep, E. (2017) 'Comparative transcriptomics support evolutionary convergence of diapause responses across Insecta', *Physiological Entomology*, 42(3), pp. 246–256. doi: 10.1111/phen.12193.
Rauwerda, H., Pagano, J. F. B., de Leeuw, W. C., Ensink, W., Nehrdich, U., de Jong, M., Jonker, M., Spaink, H. P. and Breit, T. M. (2017) 'Transcriptome dynamics in early zebrafish embryogenesis determined by high-resolution time course analysis of 180 successive, individual zebrafish embryos', *BMC Genomics*. BMC Genomics, 18(1), pp. 1–15. doi: 10.1186/s12864-017-3672-z.
Ren, X. Y., Zhang, L. S., Han, Y. H., An, T., Liu, Y., Li, Y. Y. and Chen, H. Y. (2016) 'Proteomic research on diapause-related proteins in the female ladybird, Coccinella septempunctata L.', *Bulletin of Entomological Research*, 106(02), pp. 168–174. doi: 10.1017/S0007485315000954.
Renfree, M. B. and Fenelon, J. C. (2017) 'The enigma of embryonic diapause', *Development*, 144(18), pp. 3199–3210. doi: 10.1242/dev.148213.
Reynolds, J. a and Hand, S. C. (2009) 'Decoupling development and energy flow during embryonic diapause in the cricket, Allonemobius socius.', *The Journal of experimental biology*, 212(Pt 13), pp. 2065–74. doi: 10.1242/jeb.027359.
Rinehart, J. P., Li, A., Yocum, G. D., Robich, R. M., Hayward, S. a L. and Denlinger, D. L. (2007) 'Up-regulation of heat shock proteins is essential for cold survival during insect diapause.', *Proceedings of the National Academy of Sciences of the United States of America*, 104(27), pp. 11130–11137. doi: 10.1073/pnas.0703538104.
Robbins, H. M., Van Stappen, G., Sorgeloos, P., Sung, Y. Y., MacRae, T. H. and Bossier, P. (2010) 'Diapause termination and development of encysted Artemia embryos: roles for nitric oxide and hydrogen peroxide.', *The Journal of experimental biology*, 213(Pt 9), pp. 1464–1470. doi: 10.1242/jeb.041772.
Roch, G. J. and Sherwood, N. M. (2011) 'Stanniocalcin has deep evolutionary roots in eukaryotes', *Genome Biology and Evolution*, 3(1), pp. 284–294. doi: 10.1093/gbe/evr020.

Roncalli, V., Sommer, S. A., Cieslak, M. C., Clarke, C., Hopcroft, R. R. and Lenz, P. H. (2018) 'Physiological characterization of the emergence from diapause: A transcriptomics approach', *Scientific Reports*. Springer US, 8(1), p. 12577. doi: 10.1038/s41598-018-30873-0.
Roulin, A. C., Mariadassou, M., Hall, M. D., Walser, J. C., Haag, C. and Ebert, D. (2015) 'High genetic variation in resting-stage production in a metapopulation: Is there evidence for local adaptation?', *Evolution*, 69(10), pp. 2747–2756. doi: 10.1111/evo.12770.
Roulin, A. C., Routtu, J., Hall, M. D., Janicke, T., Colson, I., Haag, C. R. and Ebert, D. (2013) 'Local adaptation of sex induction in a facultative sexual crustacean: insights from QTL mapping and natural populations of Daphnia magna.', *Molecular ecology*, pp. 3567–3579. doi: 10.1111/mec.12308.
Russell, J. J., Theriot, J. A., Sood, P., Marshall, W. F., Landweber, L. F., Fritz-Laylin, L., Polka, J. K., Oliferenko, S., Gerbich, T., Gladfelter, A., Umen, J., Bezanilla, M., Lancaster, M. A., He, S., Gibson, M. C., Goldstein, B., Tanaka, E. M., Hu, C. K. and Brunet, A. (2017) *Non-model model organisms*, *BMC Biology*. BMC Biology. doi: 10.1186/s12915-017-0391-5.
Schadt, E. E., Turner, S. and Kasarskis, A. (2010) 'A window into third-generation sequencing', *Human Molecular Genetics*, 19(R2), pp. 227–240. doi: 10.1093/hmg/ddq416.
Schein, V., Cardoso, J. C. R., Pinto, P. I. S., Anjos, L., Silva, N., Power, D. M. and Canário, A. V. M. (2012) 'Four stanniocalcin genes in teleost fish: Structure, phylogenetic analysis, tissue distribution and expression during hypercalcemic challenge', *General and Comparative Endocrinology*. Elsevier Inc., 175(2), pp. 344–356. doi: 10.1016/j.ygcen.2011.11.033.
Schwarzenberger, A., Christjani, M. and Wacker, A. (2014) 'Longevity of Daphnia and the attenuation of stress responses by melatonin', *BMC Physiology*, 14(1), pp. 1–7. doi: 10.1186/s12899-014-0008-y.
Sgolastra, F., Bosch, J., Molowny-Horas, R., Maini, S. and Kemp, W. P. (2010) 'Effect of temperature regime on diapause intensity in an adult-wintering Hymenopteran with obligate diapause', *Journal of Insect Physiology*, 56(2), pp. 185–194. doi: 10.1016/j.jinsphys.2009.10.001.
Shanley, D. P. and Kirkwood, T. B. L. (2006) 'Caloric restriction does not enhance longevity in all species and is unlikely to do so in humans', *Biogerontology*, 7(3), pp. 165–168. doi: 10.1007/s10522-006-9006-1.
Shaw, J. R., Pfrender, M. E., Eads, B. D., Klaper, R., Callaghan, A., Sibly, R. M., Colson, I., Jansen, B., Gilbert, D. and Colbourne, J. K. (2008) 'Daphnia as an emerging model for toxicological genomics', *Advances in Experimental Biology*, pp. 5–7. doi: 10.1016/S1872-2423(08)00005-7.
Shelly C. Lu, M. D. (2014) 'Glutathione Synthesis', *Biochim Biophys Acta*, 1830(5), pp. 3143–3153. doi: 10.1016/j.bbagen.2012.09.008.GLUTATHIONE.
Shimizu, Y., Mukai, A. and Goto, S. G. (2016) 'Cell cycle arrest in the jewel wasp Nasonia vitripennis in larval diapause', *Journal of Insect Physiology*. Elsevier Ltd. doi: 10.1016/j.jinsphys.2016.11.011.
Sim, C. and Denlinger, D. L. (2008) 'Insulin signaling and FOXO regulate the overwintering diapause of the mosquito Culex pipiens.', *Proceedings of the National Academy of Sciences of the United States of America*, 105(18), pp. 6777–6781. doi: 10.1073/pnas.0802067105.

Sim, C. and Denlinger, D. L. (2009) ‘A shut-down in expression of an insulin-like peptide, ILP-1, halts ovarian maturation during the overwintering diapause of the mosquito Culex pipiens’, *Insect molecular biology*, 18(3), pp. 25–32. doi: 10.1021/nl061786n.Core-Shell.
Sim, C. and Denlinger, D. L. (2009) ‘Transcription profiling and regulation of fat metabolism genes in diapausing adults of the mosquito Culex pipiens.’, *Physiological genomics*, 39(3), pp. 202–209. doi: 10.1152/physiolgenomics.00095.2009.
Sim, C. and Denlinger, D. L. (2011) ‘Catalase and superoxide dismutase-2 enhance survival and protect ovaries during overwintering diapause in the mosquito Culex pipiens’, *Journal of Insect Physiology*. Elsevier Ltd, 57(5), pp. 628–634. doi: 10.1016/j.jinsphys.2011.01.012.
Sim David L., C. D., Sim, C. and Denlinger, D. L. (2009) ‘Transcription profiling and regulation of fat metabolism genes in diapausing adults of the mosquito Culex pipiens.’, *Physiological Genomics*, 39(3), pp. 202–9. doi: 10.1152/physiolgenomics.00095.2009.
Simon, J. C., Pfrender, M. E., Tollrian, R., Tagu, D. and Colbourne, J. K. (2011) ‘Genomics of environmentally induced phenotypes in 2 extremely plastic arthropods’, *Journal of Heredity*, 102(5), pp. 512–525. doi: 10.1093/jhered/esr020.
Simon, S., Sagasser, S., Saccenti, E., Brugler, M. R., Schranz, M. E., Hadrys, H., Amato, G. and DeSalle, R. (2017) ‘Comparative transcriptomics reveal developmental turning points during embryogenesis of a hemimetabolous insect, the damselfly Ischnura elegans’, *Scientific Reports*. Springer US, 7(1), pp. 1–14. doi: 10.1038/s41598-017-13176-8.
Sinclair, B. J. and Marshall, K. E. (2018) ‘The many roles of fats in overwintering insects’, *The Journal of Experimental Biology*, 221(Suppl 1), p. jeb161836. doi: 10.1242/jeb.161836.
Ślusarczyk, M., Dawidowicz, P. and Rygielska, E. (2005) ‘Hide, rest or die: A light-mediated diapause response in Daphnia magna to the threat of fish predation’, *Freshwater Biology*, 50(1), pp. 141–146. doi: 10.1111/j.1365-2427.2004.01309.x.
Smilde, A. K., Jansen, J. J., Hoefsloot, H. C. J., Lamers, R. J. A. N., van der Greef, J. and Timmerman, M. E. (2005) ‘ANOVA-simultaneous component analysis (ASCA): A new tool for analyzing designed metabolomics data’, *Bioinformatics*, 21(13), pp. 3043–3048. doi: 10.1093/bioinformatics/bti476.
Solbreck, C. and Widenfalk, O. (2012) ‘Very long diapause and extreme resistance to population disturbance in a galling insect’, *Ecological Entomology*, 37(1), pp. 51–55. doi: 10.1111/j.1365-2311.2011.01335.x.
Solnica-Krezel, L. and Sepich, D. S. (2012) ‘Gastrulation: Making and Shaping Germ Layers’, *Annual Review of Cell and Developmental Biology*, 28(1), pp. 687–717. doi: 10.1146/annurev-cellbio-092910-154043.
Soneson, C., Love, M. I. and Robinson, M. D. (2016) ‘Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences’, *F1000Research*, 4(2), p. 1521. doi: 10.12688/f1000research.7563.2.
Southam, A. D., Payne, T. G., Cooper, H. J., Arvanitis, T. N. and Viant, M. R. (2007) ‘Dynamic range and mass accuracy of wide-scan direct infusion nanoelectrospray fourier transform ion cyclotron resonance mass spectrometry-based metabolomics increased by the spectral stitching method.’, *Analytical chemistry*, 79(12), pp. 4595–602. doi: 10.1021/ac062446p.
Southam, A. D., Payne, T. G., Cooper, H. J., Arvanitis, T. N. and Viant, M. R. (2007)

‘Dynamic range and mass accuracy of wide-scan direct infusion nanoelectrospray fourier transform ion cyclotron resonance mass spectrometry-based metabolomics increased by the spectral stitching method’, *Analytical Chemistry*, 79(12), pp. 4595–4602. doi: 10.1021/ac062446p.
Southam, A. D., Weber, R. J. M., Engel, J., Jones, M. R. and Viant, M. R. (2017) ‘A complete workflow for high-resolution spectral-stitching nanoelectrospray direct-infusion mass-spectrometry-based metabolomics and lipidomics’, *Nature Protocols*. Nature Publishing Group, 12(2), pp. 255–273. doi: 10.1038/nprot.2016.156.
Stanley, J. K., Perkins, E. J., Habib, T., Sims, J. G., Chappell, P., Escalon, B. L., Wilbanks, M. and Garcia-Reyero, N. (2013) ‘The good, the bad, and the toxic: Approaching hormesis in daphnia magna exposed to an energetic compound’, *Environmental Science and Technology*, 47(16), pp. 9424–9433. doi: 10.1021/es401115q.
Stross, R. G. and Hill, J. C. (1965) ‘Diapause induction in Daphnia requires two stimuli.’, *Science (New York, N.Y.)*, 150(3702), pp. 1462–4. doi: 10.1126/science.150.3702.1462.
Su, H., Zou, J., Zhou, Q., Yu, Q., Yang, Y. and Yang, Y. (2017) ‘Better cold tolerance of Bt-resistant Spodoptera exigua strain and the corresponding cold-tolerant mechanism’, 140, pp. 51–57. doi: 10.1016/j.pestbp.2017.06.003.
Suarez, R. K. and Moyes, C. D. (2012) ‘Metabolism in the age of “omes”’, *Journal of Experimental Biology*, 215(14), pp. 2351–2357. doi: 10.1242/jeb.059725.
Sujii, E. R., Garcia, M. A., Fontes, E. M. G. and O’neil, R. J. (2002) ‘Predation as a Mortality Factor in Populations of the Spittlebug , Deois flavopicta Stål (Homoptera : Cercopidae)’, *Neotropical Entomology*, 31, pp. 581–588.
Tammariello, S. P. and Denlinger, D. L. (1998a) ‘G0/G1 cell cycle arrest in the brain of Sarcophaga crassipalpis during pupal diapause and the expression pattern of the cell cycle regulator, proliferating cell nuclear antigen’, *Insect Biochemistry and Molecular Biology*, 28(2), pp. 83–89. doi: 10.1016/S0965-1748(97)00082-9.
Tammariello, S. P. and Denlinger, D. L. (1998b) ‘G0/G1 cell cycle arrest in the brain of Sarcophaga crassipalpis during pupal diapause and the expression pattern of the cell cycle regulator, proliferating cell nuclear antigen’, *Insect Biochemistry and Molecular Biology*, 28(2), pp. 83–89. doi: 10.1016/S0965-1748(97)00082-9.
Tang, B., Wang, S., Wang, S.-G., Wang, H.-J., Zhang, J.-Y. and Cui, S.-Y. (2018) ‘Invertebrate Trehalose-6-Phosphate Synthase Gene: Genetic Architecture, Biochemistry, Physiological Function, and Potential Applications’, *Frontiers in Physiology*, 9(January), pp. 1–13. doi: 10.3389/fphys.2018.00030.
Tatar, M. and Yin, C. M. (2001) ‘Slow aging during insect reproductive diapause: Why butterflies, grasshoppers and flies are like worms’, *Experimental Gerontology*, 36(4–6), pp. 723–738. doi: 10.1016/S0531-5565(00)00238-2.
Tatarazako, N., Oda, S., Watanabe, H., Morita, M. and Iguchi, T. (2003) ‘Juvenile hormone agonists affect the occurrence of male Daphnia.’, *Chemosphere*, 53(8), pp. 827–33. doi: 10.1016/S0045-6535(03)00761-6.
Taylor, N. S., Weber, R. J. M., Southam, A. D., Payne, T. G., Hrydziuszko, O., Arvanitis, T. N. and Viant, M. R. (2008) ‘A new approach to toxicity testing in Daphnia magna: application of high throughput FT-ICR mass spectrometry metabolomics’, *Metabolomics*, 5(1), pp. 44–58. doi: 10.1007/s11306-008-0133-3.
Taylor, N. S., Weber, R. J. M., Southam, A. D., Payne, T. G., Hrydziuszko, O., Arvanitis, T. N. and Viant, M. R. (2009) ‘A new approach to toxicity testing in Daphnia

magna: Application of high throughput FT-ICR mass spectrometry metabolomics', *Metabolomics*, 5(1), pp. 44–58. doi: 10.1007/s11306-008-0133-3.
Taylor, N. S., White, T. A. and Viant, M. R. (2017) 'Defining the baseline and oxidant perturbed lipidomic profiles of Daphnia magna', *Metabolites*, 7(1). doi: 10.3390/metabo7010011.
Tessier, A. J., Henry, L. L., Goulden, C. E., Durand, M. W. and Jul, N. (2007) 'Starvation in Daphnia : Energy Reserves and Reproductive Allocation Starvation in Daphnia : Energy reserves and reproductive allocation1', 28(4), pp. 667–676.
Thamotharan, M., Lombardo, Y. B., Bawani, S. Z. and Adibi, S. A. (1997) 'An active mechanism for completion of the final stage of protein degradation in the liver, lysosomal transport of dipeptides', *Journal of Biological Chemistry*, 272(18), pp. 11786–11790. doi: 10.1074/jbc.272.18.11786.
Thomas, P. D., Kejariwal, A., Guo, N., Mi, H., Campbell, M. J., Muruganujan, A. and Lazareva-Ulitsky, B. (2006) 'Applications for protein sequence-function evolution data: mRNA/protein expression analysis and coding SNP scoring tools', *Nucleic Acids Research*, 34(WEB. SERV. ISS.), pp. 645–650. doi: 10.1093/nar/gkl229.
Thompson, A. W. and Ortí, G. (2016) 'Annual Killifish Transcriptomics and Candidate Genes for Metazoan Diapause', *Molecular Biology and Evolution*, 33(9), pp. 2391–2395. doi: 10.1093/molbev/msw110.
Tibshirani, R., Friedman, J. H., Mayeda, T. K., Lyons, T. W., Devlin, S. J., Rajagopalan, B., Lall, U., Vajda, I., Grassberger, P., Stuetzle, W., Krzyzak, a, Linder, T., Zeger, K., Smrekar, M., Rizzo, M., Kudlicki, a, Rowicka, M., Mcknight, S. L., Pearson, R., Huttunen, H., Archive, T. B., Alva, S. and Verhoeven, M. (2011) 'The Structure of the Eukaryotic', *Science*, (December), pp. 1524–1529. doi: 10.1126/science.1212642.
DI Tommaso, P., Chatzou, M., Floden, E. W., Barja, P. P., Palumbo, E. and Notredame, C. (2017) 'Nextflow enables reproducible computational workflows', *Nature Biotechnology*, 35(4), pp. 316–319. doi: 10.1038/nbt.3820.
Toyota, K., Gavin, A., Miyagawa, S., Viant, M. R. and Iguchi, T. (2016) 'Metabolomics reveals an involvement of pantothenate for male production responding to the short-day stimulus in the water flea, Daphnia pulex', *Scientific Reports*. Nature Publishing Group, 6, pp. 1–9. doi: 10.1038/srep25125.
Toyota, K., Hiruta, C., Ogino, Y., Miyagawa, S., Okamura, T., Onishi, Y., Tatarazako, N. and Iguchi, T. (2016) 'Comparative Developmental Staging of Female and Male Water Fleas Daphnia pulex and Daphnia magna During Embryogenesis'. doi: 10.2108/zs150116.
Tu, M.-P., Yin, C.-M. and Tatar, M. (2005) 'Mutations in insulin signaling pathway alter juvenile hormone synthesis in Drosophila melanogaster', *General and Comparative Endocrinology*, 142(3), pp. 347–356. doi: 10.1016/j.ygcen.2005.02.009.
Tu, X., Wang, J., Hao, K., Whitman, D. W., Fan, Y. and Cao, G. (2015) 'Transcriptomic and proteomic analysis of pre-diapause and non- diapause eggs of migratory locust , Locusta migratoria L . (Orthoptera : Acridoidea)', *Nature Publishing Group*. Nature Publishing Group, (January), pp. 1–14. doi: 10.1038/srep11402.
Tu, X., Wang, J., Hao, K., Whitman, D. W., Fan, Y., Cao, G. and Zhang, Z. (2015) 'Transcriptomic and proteomic analysis of pre-diapause and non-diapause eggs of migratory locust, Locusta migratoria L. (Orthoptera: Acridoidea)', *Scientific Reports*. Nature Publishing Group, 5(January), p. 11402. doi: 10.1038/srep11402.

Vandekerkhove, J., Declerck, S., Brendonck, L., Conde-Porcuna, J. M., Jeppesen, E. and Meester, L. De (2005) 'Hatching of cladoceran resting eggs: temperature and photoperiod', *Freshwater Biology*, 50(1), pp. 96–104. doi: 10.1111/j.1365-2427.2004.01312.x.
Vermeulen, K., Van Bockstaele, D. R. and Berneman, Z. N. (2003) 'The cell cycle:a review of regulation,deregulation and therapeutic targets in cancer', *Cell Proliferation*, 36(3), pp. 131–149. doi: 10.1306/74D715D2-2B21-11D7-8648000102C1865D.
Vesala, L., Salminen, T. S., Ko, V., Zahradní, H. and Hoikkala, A. (2012) 'Myo-inositol as a main metabolite in overwintering flies : seasonal metabolomic profiles and cold stress tolerance in a northern drosophilid fly', 4, pp. 2891–2897. doi: 10.1242/jeb.069948.
Vesterlund, L., Jiao, H., Unneberg, P., Hovatta, O. and Kere, J. (2011) 'The zebrafish transcriptome during early development', *BMC Developmental Biology*, 11. doi: 10.1186/1471-213X-11-30.
Vis, D. J., Westerhuis, J. A., Smilde, A. K. and van der Greef, J. (2007) 'Statistical validation of megavariate effects in ASCA', *BMC Bioinformatics*, 8, pp. 1–8. doi: 10.1186/1471-2105-8-322.
Vivancos, P. D., Dong, Y., Ziegler, K., Markovic, J., Pallardó, F. V., Pellny, T. K., Verrier, P. J. and Foyer, C. H. (2010) 'Recruitment of glutathione into the nucleus during cell proliferation adjusts whole-cell redox homeostasis in Arabidopsis thaliana and lowers the oxidative defence shield', *Plant Journal*, 64(5), pp. 825–838. doi: 10.1111/j.1365-313X.2010.04371.x.
Wadsworth, C. B. and Dopman, E. B. (2015) 'Transcriptome profiling reveals mechanisms for the evolution of insect seasonality', *Journal of Experimental Biology*, 218(22), pp. 3611–3622. doi: 10.1242/jeb.126136.
Wagner, J. T. and Podrabsky, J. E. (2015) 'Extreme tolerance and developmental buffering of UV-C induced DNA damage in embryos of the annual killifish Austrofundulus limnaeus', *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology*, 323(1), pp. 10–30. doi: 10.1002/jez.1890.
Wang, z. et al (2001) 'RNA-seq: A revolutionary tool for transcriptomics', *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*, 10(1), pp. 57–63. doi: 10.1038/nrg2484.RNA-Seq.
Wang, X., Li, E. and Chen, L. (2016) 'A Review of Carbohydrate Nutrition and Metabolism in Crustaceans', *North American Journal of Aquaculture*, 78(2), pp. 178–187. doi: 10.1080/15222055.2016.1141129.
Weber, R. J. M. and Viant, M. R. (2010) 'MI-Pack: Increased confidence of metabolite identification in mass spectra by integrating accurate masses and metabolic pathways', *Chemometrics and Intelligent Laboratory Systems*, 104(1), pp. 75–82. doi: 10.1016/j.chemolab.2010.04.010.
Wilches, D. M., Laird, R. A., Floate, K. D. and Fields, P. G. (2016) 'A review of diapause and tolerance to extreme temperatures in dermestids (Coleoptera)', *Journal of Stored Products Research*. Elsevier Ltd, pp. 50–62. doi: 10.1016/j.jspr.2016.04.004.
Williams, K. D., Busto, M., Suster, M. L., So, A. K.-C., Ben-Shahar, Y., Leevers, S. J. and Sokolowski, M. B. (2006) 'Natural variation in Drosophila melanogaster diapause due to the insulin-regulated PI3-kinase', *Proceedings of the National Academy of Sciences*, 103(43), pp. 15911–15915. doi: 10.1073/pnas.0604592103.
Wu, H., Southam, A. D., Hines, A. and Viant, M. R. (2008) 'High-throughput tissue

extraction protocol for NMR- and MS-based metabolomics', *Analytical Biochemistry*, 372(2), pp. 204–212. doi: 10.1016/j.ab.2007.10.002.
Xu, J., Bao, B., Zhang, Z.-F., Yi, Y.-Z. and Xu, W.-H. (2009) 'Identification of a novel gene encoding the trehalose phosphate synthase in the cotton bollworm, Helicoverpa armigera', *Glycobiology*, 19(3), pp. 250–257. doi: cwn127 [pii] 10.1093/glycob/cwn127 ET - 2008/11/14.
Yamashita, O. (1996) 'Diapause Hormone of the Silkworm , Bombyx mori : Structure , Gene Expression and Function', 42(I), pp. 669–679.
Yoshida, M., Matsuda, H., Kubo, H. and Nishimura, T. (2016) 'Molecular characterization of Tps1 and Treh genes in Drosophila and their role in body water homeostasis', *Scientific Reports*. Nature Publishing Group, 6(July), pp. 4–5. doi: 10.1038/srep30582.
Yoshiko, Y., Aubin, J. E. and Maeda, N. (2002) 'Stanniocalcin 1 (STC1) protein and mRNA are developmentally regulated during embryonic mouse osteogenesis: The potential of STC1 as an autocrine/paracrine factor for osteoblast development and bone formation', *Journal of Histochemistry and Cytochemistry*, 50(4), pp. 483–491. doi: 10.1177/002215540205000405.
Zdobnov, E. M., Tegenfeldt, F., Kuznetsov, D., Waterhouse, R. M., Simao, F. A., Ioannidis, P., Seppey, M., Loetscher, A. and Kriventseva, E. V. (2017) 'OrthoDB v9.1: Cataloging evolutionary and functional annotations for animal, fungal, plant, archaeal, bacterial and viral orthologs', *Nucleic Acids Research*, 45(D1), pp. D744–D749. doi: 10.1093/nar/gkw1119.
Zelko, I. N., Mariani, T. J. and Folz, R. J. (2002) 'Superoxide dismutase multigene family: A comparison of the CuZn-SOD (SOD1), Mn-SOD (SOD2) and EC-SOD (SOD3) gene structures, evolution and expression', *Free Radical Biology & Medicine*, 33(3), pp. 337–349. doi: 10.1097/EDE.0000000000000438.
Zhang, L. and Baer, K. N. (2000) 'The influence of feeding, photoperiod and selected solvents on the reproductive strategies of the water flea, Daphnia magna', *Environmental Pollution*, 110(3), pp. 425–430. doi: 10.1016/S0269-7491(99)00324-3.
Zhang, Q. and Denlinger, D. L. (2012) 'Dynamics of diapause hormone and prothoracicotropic hormone transcript expression at diapause termination in pupae of the corn earworm, Helicoverpa zea.', *Peptides*. Elsevier Inc., 34(1), pp. 120–6. doi: 10.1016/j.peptides.2011.05.019.
Zhang, W., Li, F. and Nie, L. (2010) 'Integrating multiple "omics" analysis for microbial biology: Application and methodologies', *Microbiology*, 156(2), pp. 287–301. doi: 10.1099/mic.0.034793-0.
Zhao, L. C., Hou, Y. S. and Sima, Y. H. (2014) 'Changes in glutathione redox cycle during diapause determination and termination in the bivoltine silkworm, Bombyx mori', *Insect Science*, 21(1), pp. 39–46. doi: 10.1111/1744-7917.12015.
Zhao, L., Xu, X., Xu, Z., Liu, Y. and Sun, S. (2018) 'Diapause Induction , Color Change , and Cold Tolerance Physiology of the Diapausing Larvae of the Chouioia cunea (Hymenoptera : Eulophidae)', (March), pp. 1–5. doi: 10.1093/jisesa/ieu156.
Zhao, Y., Ding, X., Ye, X., Dai, Z. M., Yang, J. S. and Yang, W. J. (2012) 'Involvement of cyclin K posttranscriptional regulation in the formation of artemia diapause cysts', *PLoS ONE*, 7(2). doi: 10.1371/journal.pone.0032129.
Zhi-Hui, S., Motoko, I., Yukihiro, S., Hiroyuki, S., Kunio, I., Isobe, M. and Yamashita, O. (1994) 'Molecular characterization of ovary trehalase of the silkworm, Bombyx mori and its transcriptional activation by diapause hormone', *Biochimica et*

Biophysica Acta, 1218, pp. 366–374.
Ziv, T., Chalifa-Caspi, V., Denekamp, N., Plaschkes, I., Kierszniowska, S., Blais, I., Admon, A. and Lubzens, E. (2017) 'Dormancy in Embryos: Insight from Hydrated Encysted Embryos of an Aquatic Invertebrate', *Molecular & Cellular Proteomics*, 16(10), pp. 1746–1769. doi: 10.1074/mcp.RA117.000109.
Zou, H., Hastie, T. and Tibshirani, R. (2006) 'Sparse principal component analysis', *Journal of Computational and Graphical Statistics*, 15(2), pp. 265–286. doi: 10.1198/106186006X113430.

9. Appendix A

9.1. Summary of DAPI-stained nuclei count data describing the development of sexually and asexually produced *Daphnia magna* embryos

Table 9-1 Mean number of cells (DAPI-stained nuclei) in asexually produced subitaneous embryos of *Daphnia magna* in the hours post-ovulation

Embryo type	Time (h)	Mean number of cells	Standard deviation	Standard error of mean	Number of embryos (n)
Asexual	4	27	6.6	3.8	3
	5	60	14.0	7.0	4
	6	135	30.4	17.5	3
	7	596	206.8	119.4	3
	8	683	178.0	89.0	4
	9	1018	284.7	164.4	3
	10	1166	374.6	216.3	3
	11	1618	176.2	88.1	4
	12	2361	NA	NA	1
	13	2209	448.2	158.5	8
	14	2753	297.0	210.0	2
	15	3713	137.0	79.1	3
	16	3509	791.8	457.1	3
	18	4561	756.5	436.7	3
	25	7000	0.0	0.0	3

Table 9-2 Mean number of cells (DAPI-stained nuclei) in sexually produced diapause-destined embryos of *Daphnia magna* in the hours post-ovulation

Embryo type	Time (h)	Mean number of cells	Standard deviation	Standard error of mean	Number of embryos (n)
Sexual	12	55	14.3	8.3	3
	14	136	46.9	21.0	5
	16	264	164.8	58.3	8
	18	362	112.2	37.4	9
	20	537	224.3	79.3	8
	22	773	112.0	64.7	3
	24	1066	386.7	122.3	10
	26	1220	353.8	144.4	4
	28	1140	158.1	70.7	5
	30	1851	486.2	280.7	3
	34	2966	178.9	103.3	3
	36	3278	237.4	137.1	3
	38	3240	309.4	178.7	3
	40	2978	211.4	94.6	3
	44	3247	250.3	125.2	4
	46	3140	422.8	211.4	4
	48	3079	123.2	71.1	3
	55	3124	76.5	44.2	3
	74	3458	167.6	96.8	3
	408	3705	176.7	102.0	3

9.2. Summary of sequencing metrics, data processing and quality assessment of results from the RNA-sequencing of sexually and asexually produced *Daphnia magna* embryos

Table 9-3 Table of *Daphnia magna* embryo sample sequencing metrics, read quality and successful alignment to transcriptome

Sample ID	Embryo type	Time (h)	Replicate	Number of Raw reads	Number of raw bases	Q20 (%)	Number of read pairs before trimming	Number of read pairs after trimming	Proportion binned during trimming	Percentage trimmed reads mapped
A021	Asexual	2	1	12,803,110	1,280,311,000	94.08	6,401,555	5,558,422	0.13	99.04
A022	Asexual	2	2	13,001,624	1,300,162,400	94.48	6,500,812	5,673,379	0.13	98.81
A023	Asexual	2	3	19,026,080	1,902,608,000	94.27	9,513,040	8,338,395	0.12	98.98
A0101	Asexual	10	1	11,639,200	1,163,920,000	93.64	5,819,600	5,093,465	0.12	98.13
A0102	Asexual	10	2	18,454,426	1,845,442,600	93.63	9,227,213	8,105,375	0.12	97.88
A0103	Asexual	10	3	15,802,560	1,580,256,000	94.50	7,901,280	6,815,526	0.14	98.06
A0481	Asexual	48	1	14,630,994	1,463,099,400	93.23	7,315,497	6,386,288	0.13	96.82
A0482	Asexual	48	2	20,334,908	2,033,490,800	94.49	10,167,454	8,773,583	0.14	97.23
A0483	Asexual	48	3	13,528,460	1,352,846,000	93.87	6,764,230	5,774,582	0.15	96.28
O21	Sexual	2	1	15,902,870	1,590,287,000	94.13	7,951,435	6,917,460	0.13	99.03
O22	Sexual	2	2	12,196,826	1,219,682,600	94.57	6,098,413	5,301,951	0.13	99.05
O23	Sexual	2	3	25,948,362	2,594,836,200	94.35	12,974,181	11,302,554	0.13	98.96
O101	Sexual	10	1	24,015,302	2,401,530,200	93.70	12,007,651	10,600,747	0.12	98.12
O102	Sexual	10	2	23,008,550	2,300,855,000	94.15	11,504,275	10,211,451	0.11	97.84
O103	Sexual	10	3	20,698,932	2,069,893,200	94.27	10,349,466	9,038,219	0.13	97.71
O481	Sexual	48	1	10,260,268	1,026,026,800	94.04	5,130,134	4,463,368	0.13	97.06
O482	Sexual	48	2	19,761,796	1,976,179,600	93.76	9,880,898	8,564,512	0.13	97.55
O483	Sexual	48	3	32,318,102	3,231,810,200	92.56	16,159,051	13,965,850	0.14	96.86
E21	Sexual	74	1	11,828,782	1,182,878,200	93.11	5,914,391	5,205,894	0.12	94.16
E22	Sexual	74	2	21,989,936	2,198,993,600	93.32	10,994,968	9,561,434	0.13	96.44
E23	Sexual	74	3	11,445,090	1,144,509,000	93.06	5,722,545	4,769,227	0.17	95.52
E101	Sexual	82	1	11,283,772	1,128,377,200	93.75	5,641,886	4,830,149	0.14	97.71
E102	Sexual	82	2	12,123,374	1,212,337,400	93.19	6,061,687	5,256,934	0.13	96.66
E103	Sexual	82	3	14,029,140	1,402,914,000	93.99	7,014,570	6,080,452	0.13	96.10

E481	Sexual	120	1	15,492,426	1,549,242,600	93.40	7,746,213	6,769,232	0.13	96.39
E482	Sexual	120	2	17,056,052	1,705,605,200	94.05	8,528,026	7,393,807	0.13	96.98
E483	Sexual	120	3	29,163,540	2,916,354,000	94.36	14,581,770	12,574,240	0.14	95.94
E2W1	Sexual	408	1	13,628,066	1,362,806,600	92.25	6,814,033	6,017,122	0.12	96.05
E2W2	Sexual	408	2	16,839,186	1,683,918,600	94.08	8,419,593	7,167,447	0.15	96.55
E2W3	Sexual	408	3	13,083,886	1,308,388,600	94.39	6,541,943	5,748,456	0.12	95.81
			Mean	17,043,187	1,704,318,733	93.82	8,521,594	7,408,651	0.13	97.26
			Median	15,647,493	1,564,749,300	94.02	7,823,747	6,792,379	0.13	97.15
			Total	511,295,620	51,129,562,000	–	255,647,810	222,259,521	–	–

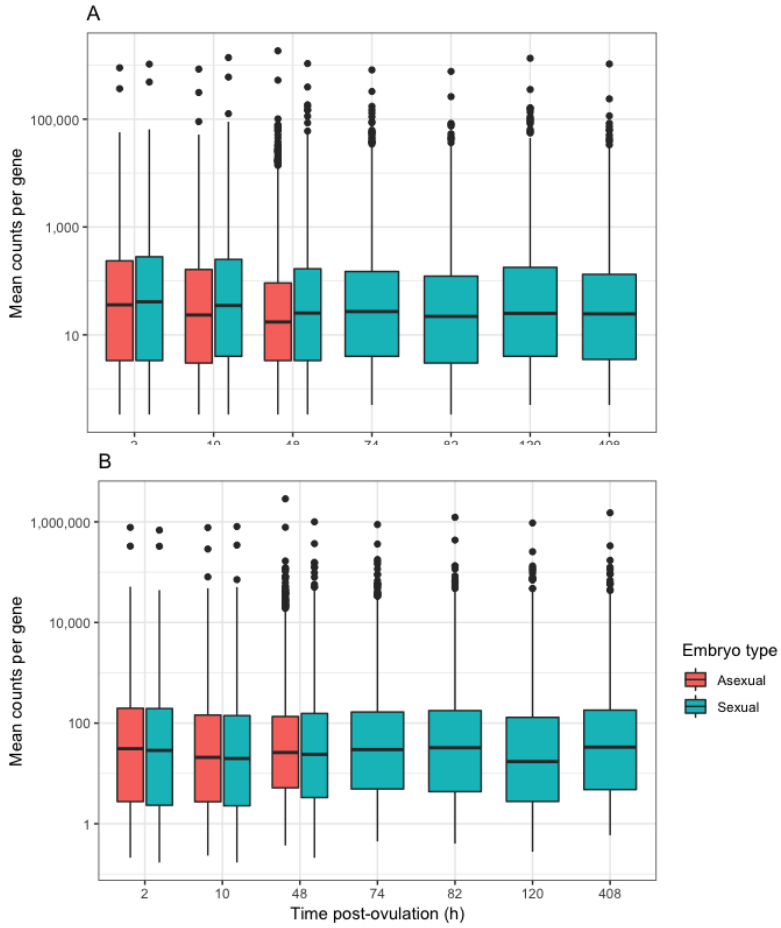


Figure 9-1 Boxplots of mean gene expression (n = 3) for pre-normalisation (A) and normalised (B) data for each time pint post-ovulation collected from asexually and sexually produced *Daphnia magna* embryos.

9.2.1. Removal of low-expression genes and data normalisation

Genes with low-counts, amounting to less than 10 counts across all biological samples (n = 30) were removed to reduce the impact of their comparatively higher biological noise on the estimation of fold changes and differential gene expression (Love, Huber and Anders, 2014). A total of 3,323 genes were removed from the dataset (Figure 9-2).

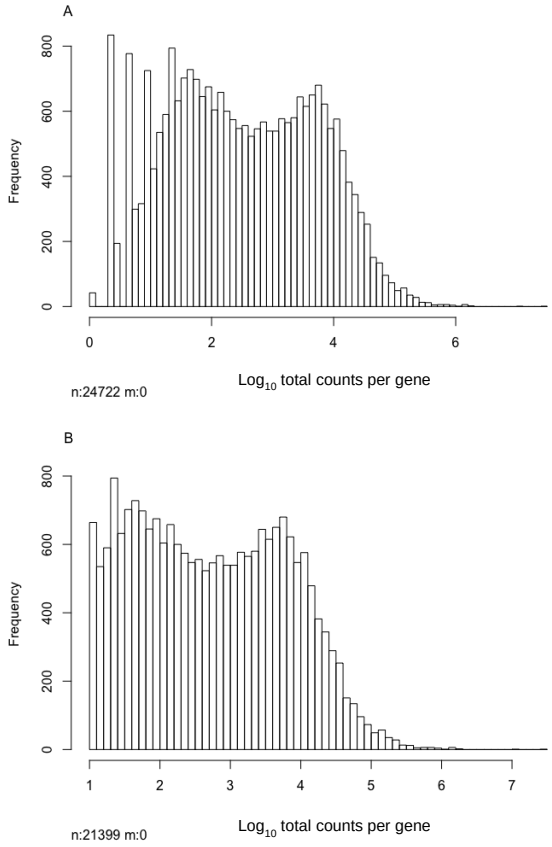


Figure 9-2 Distribution of total gene counts (log_{10}) across all biological samples (n = 30) A) before and B) after the removal of genes with <10 counts across all samples

9.2.2. Assessment and removal of outlying samples

Three samples were determined to be possible outliers based on their relative positions on PCA scores plots (Figure 9-3) and clustering based on Euclidean distance (Figure 9-4) versus the other two biological replicates of their respective sample types. These samples corresponded to one replicate each from 74 h, 120 h and 408 h post-ovulation and were removed from further analysis.

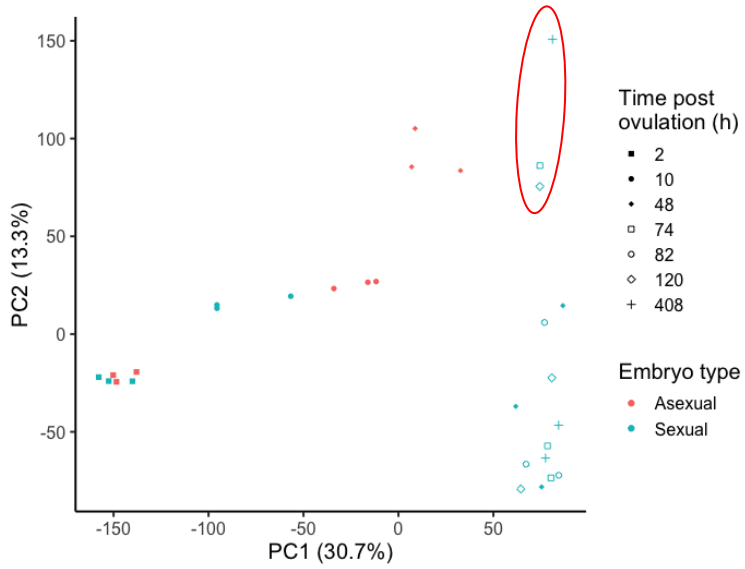


Figure 9-3 Principal component scores plots showing the first two components from PCA of log-transformed gene counts, three samples circled in red (arbitrarily drawn) were determined to be outliers based on sample scores being three or more standard deviations away from group mean score.

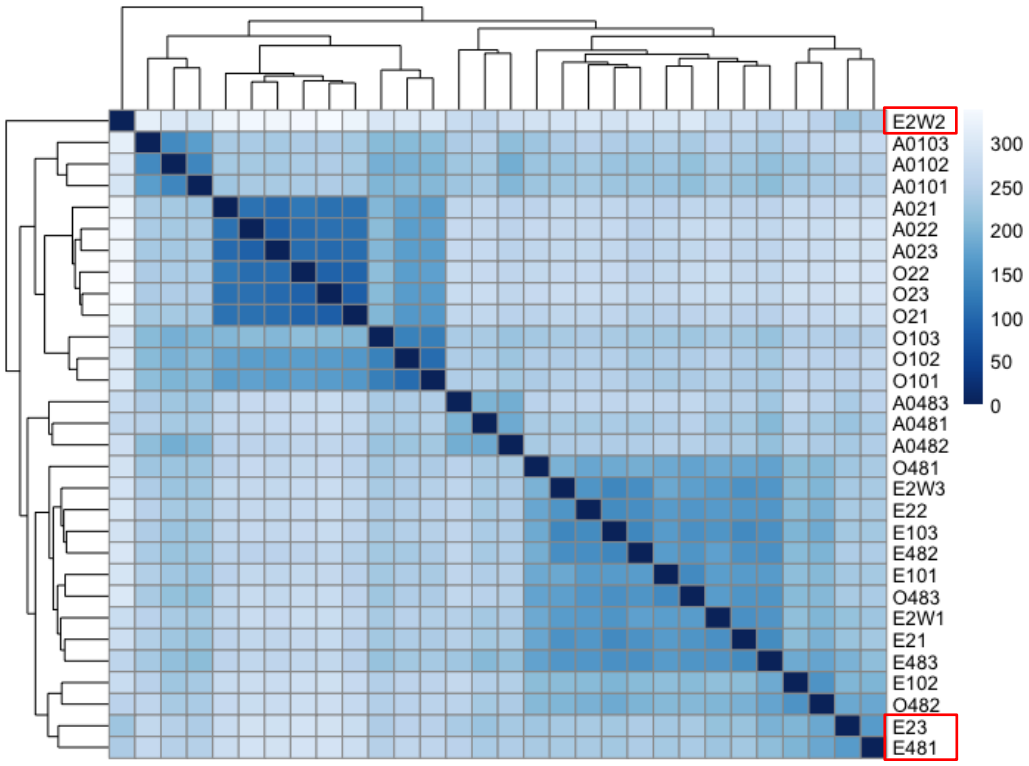


Figure 9-4 Heatmap showing the clustering of sample types and biological replicates based on Euclidean distance calculated on log-transformed gene counts, outlying samples that failed to cluster with other biological replicates of the same sample type are highlighted in red.

9.3. Summary of metabolite data processing and sample peak composition

To assess the number of shared and uniquely expressed genes and metabolites present between diapause-destined and subitaneous embryos the gene count matrix was filtered to retain only genes that were expressed in all of the biological replicates of each embryo type. For metabolomics the blank filtered matrices were filtered to remove peaks that were not present in at least 80% of biological replicates within each embryo type. This analysis reduced the number of genes and metabolites retained in the matrices–presumably due to their presence in samples only in specific time points–but revealed a high number of genes and metabolites that were unique to each embryo type (Figure 9-5).

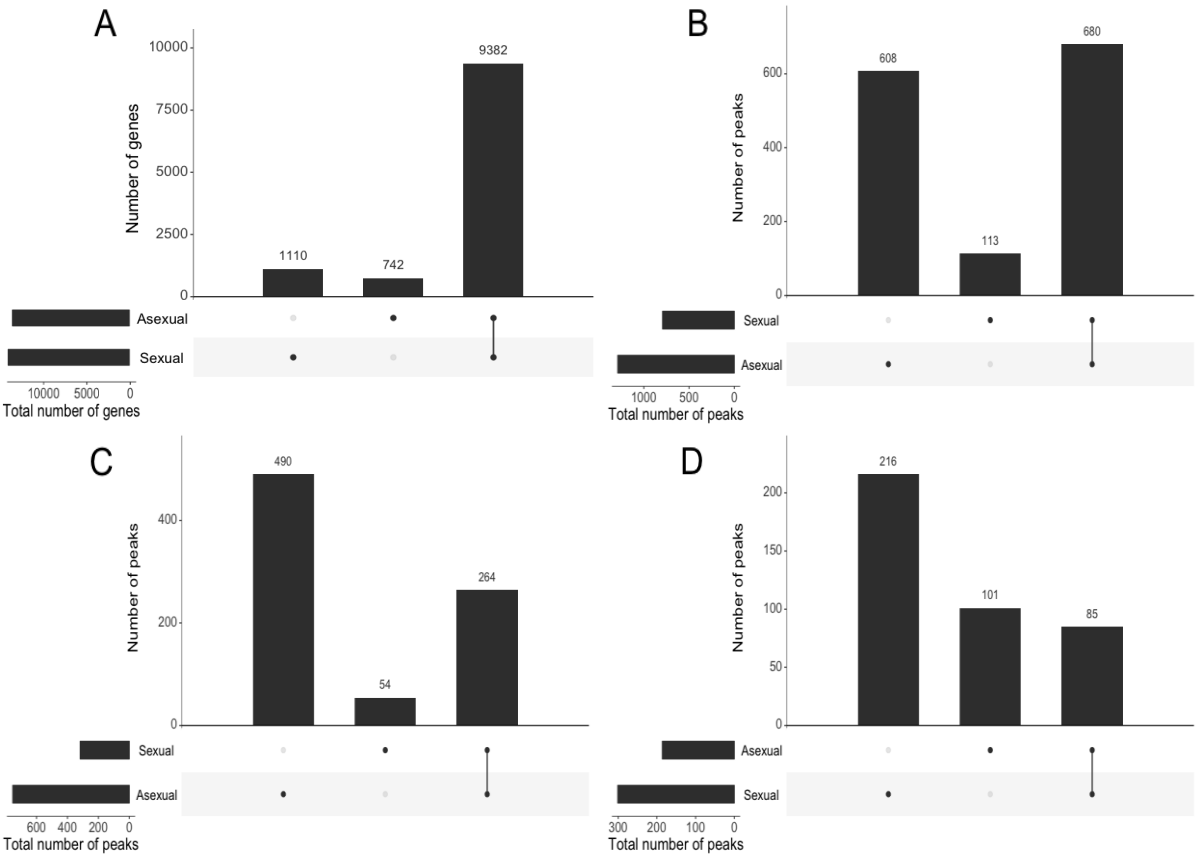


Figure 9-5 Upset plot of shared and unique genes and metabolites between diapause-destined and subitaneous embryo samples across all time points. A) Number of genes with >10 counts that occur in all biological replicates of diapause-destined and/or subitaneous samples; B-D) number of peaks that occur in at least 80% of all biological replicates of diapause-destined and/or subitaneous samples; B) Negative ion analysis of non-polar extracts; C) Negative ion analysis of polar extracts; D) Positive ion analysis of polar extracts.

Of the 11,234 genes retained in the data set, the majority (9,382; 83%) were expressed in both embryo types with only 742 (6.6%) unique to diapause-destined embryos and 1,110 (9.9%) unique to subitaneous embryos. In contrast, the analysis of the shared and unique metabolites between diapause-destined and subitaneous embryos samples across all time points revealed a larger number of features that were unique to each sample type; this was observed across all the three mass spectrometry analyses performed. Indeed, the number of unique features present in subitaneous samples was particularly high in both negative ion analyses and comparable, if not larger, than the number of features shared between the two embryo types. For example, analysis of non-polar embryo extracts in negative ion

mode resulted in 680 shared spectral features between diapause-destined and subitaneous embryos. But a further 608 features were uniquely identified in diapause-destined embryos alone. By contrast, the highest numbers of unique features were detected in diapause-destined embryos during polar ion mode and subitaneous during negative ion mode. Furthermore, the number of shared features identified during the analysis of polar embryo extracts in both negative and positive ion mode was less (264 and 85 respectively) than the number of unique features in each embryo type. Previous proteomic analysis of *Bombyx mori* eggs in diapause and non-diapause also revealed unique proteomic profiles between the two embryo types, with 253 unique proteins identified in non-diapause and 192 unique to diapause with 309 proteins shared (Fan *et al.*, 2013b).

Various filtering steps were applied throughout the analysis of the metabolomics data to retain only technically robust, biological peaks in the final results matrices (Section 2.14.1.4). Spectral noise and technically inconsistent peaks were removed, with peaks that were present only in the majority of technical replicates (2-out-of-3; Table 9-4 (Stage 1)) and that had an intensity at least 10-fold higher in biological samples versus extract blanks being retained (Table 9-4 (Stage 2)). A second phase of filtering was then performed, the first being a ‘sample filter’ to discard peaks that were not present in the majority of biological samples (Table 9-4 (Stage 4)). Followed by the removal of samples with a particularly low peak count i.e. greater than 20% missing values, and any peaks that had a relative standard deviation (RSD) in quality control (QC) samples of more than 30% (Table 9-4 (Stage 5)).

Table 9-4 Number of mass spectral peaks retained in intensity matrices throughout successive filtering stages during the processing of data obtained from the nESI-DIMS mass spectrometry metabolomics analysis of negative and positive ion analysis of *Daphnia magna* sexual and asexual embryo polar and non-polar extracts. RSD = relative standard deviation

	Stage	Negative non-polar metabolomics	Negative polar metabolomics	Positive polar metabolomics
1	**Aligned matrix**	46168	29649	29191
2	**Post-blank filtering**	22587	16789	20074
3	**80% sample filter within biological sample class**	2636	3241	3082
4	**80% sample filter across all biological samples**	825	341	286
5	**RSD <30% in QC samples**	794	272	284

Sample filtering, to retain peaks that are present in a minimum of 80% of samples **within** each biological class yielded large matrices containing in excess of 2,000 spectral features. Sample filtering **across** all biological samples (i.e. all sample classes) led to a considerable reduction in the size of the resulting intensity matrices, whilst only a small number of features were removed due to poor reproducibility in QC samples (RSD filtering). For non-polar embryo extracts analysed in negative ion mode the final intensity matrix contained a total of 794 peaks, representing a 3-fold reduction in peaks in comparison with the within sample filter matrix. In the analysis of polar extracts, in negative ion mode 272 peaks were retained and 284 peaks in positive ion mode both representing an over 10-fold reduction in the number of peaks (Table 9-4).

Following peak sample filtering, biological samples with more than 20% missing intensity values are also removed from the analysis. This is typically performed to remove samples that are particularly poor in terms of technical reproducibility (e.g.

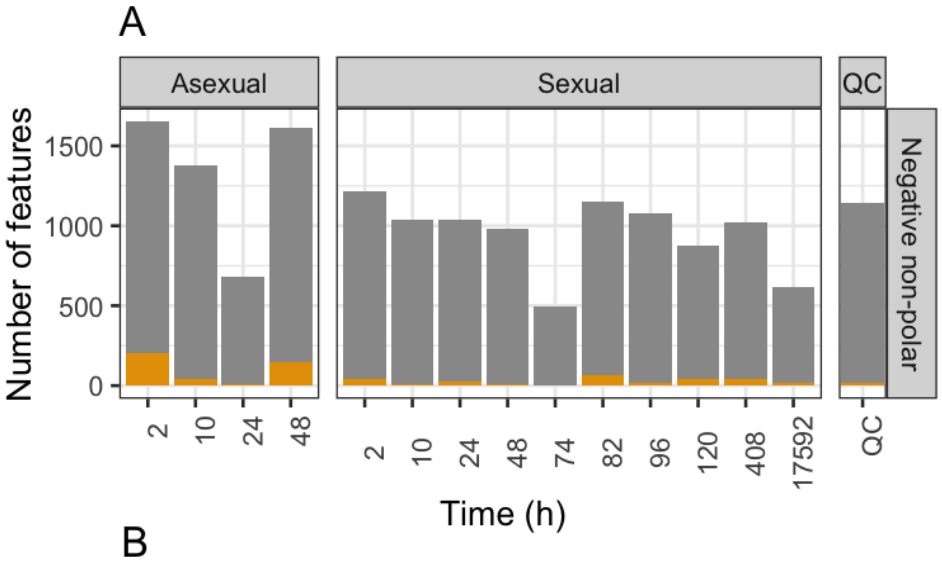
samples with a reduced number of features) due issues during the extraction or data acquisition stages that may thus, not be representative. This resulted in the removal of several biological samples and for the most part, was sporadic throughout sample classes (Table 9-5). However, there were several sample types where a particularly large number of samples were removed. In the analysis of polar extracts in negative ion mode, only 3 replicates of the sexually produced embryos 2 h and 17,592 h (2 y) after ovulation passed the threshold of < 20% missing values. Similarly, in positive ion mode, only 4 replicates of sexually produced embryos 2 h after ovulation had < 20% missing values and just a single replicate of subitaneous embryos collected 48 h after ovulation passed the missing value filter and thus, this entire group was removed from any further analysis.

Table 9-5 Table of the number of biological replicates retained in final data matrices (bold) following nESI-DIMS metabolomics analysis of polar and non-polar extracts of sexually and asexually produced *Daphnia magna* embryos collected at various time points post-ovulation. All time points contained an initial 8 biological replicates and 'processed' columns refer to the number of samples for which data was successfully acquired.

			Negative non-polar metabolomics		Negative polar metabolomics		Positive polar metabolomics	
	Time (h)	Analysed	Processed	Final matrix	Processed	Final matrix	Processed	Final matrix
Subitaneous	**2**	8	8	**8**	8	**8**	8	**8**
	10	8	8	**7**	8	**8**	8	**8**
	24	8	8	**5**	8	**8**	8	**7**
	48	8	7	**7**	7	**7**	8	**1**
Diapause-destined	**2**	8	7	**7**	8	**3**	8	**4**
	10	8	8	**8**	8	**6**	8	**5**
	24	8	7	**6**	8	**7**	8	**5**
	48	8	8	**8**	8	**8**	8	**8**
	74	8	8	**6**	8	**8**	8	**7**
	82	8	8	**8**	8	**8**	8	**8**
	96	8	8	**8**	8	**8**	8	**8**
	120	8	8	**7**	8	**8**	8	**8**
	408	8	8	**8**	7	**7**	8	**8**
	17592	8	8	**6**	8	**3**	7	**7**
	Total	112	109	**99**	110	**97**	111	**92**

9.3.1. Pairwise comparisons of subitaneous and diapause developmental *Daphnia magna* embryo samples reveals distinct peak compositions at pre-diapause and diapause

Together the highly reduced number of peaks that are present after filtering across all sample classes versus within classes and the high incidence of missing values in specific time points suggests that there are a large number of spectral features not ubiquitously detected across all samples analysed. To investigate the peak composition of samples further and try to ascertain why specific sample types were so high in missing values I investigated the total and unique numbers of peaks in each sample type across the analyses. Additionally, I assessed the percentage of peaks that were shared between sample types in pairwise comparisons. The results from these analyses are shown in Figure 9-6–Figure 9-8.



B

	Time (h)	Asexual				Sexual										
		2	10	24	48	2	10	24	48	74	82	96	120	408	17592	QC
Asexual	2	100	73	36	77	58	52	51	51	28	50	52	42	47	31	60
	10	87	100	45	87	66	62	62	61	33	59	61	49	54	37	71
	24	88	91	100	91	71	70	74	71	53	65	69	56	60	43	80
	48	79	74	39	100	57	53	55	53	29	52	53	42	49	31	62
Sexual	2	78	74	40	76	100	75	70	70	37	75	72	60	69	45	75
	10	82	82	46	82	88	100	78	79	43	82	82	68	74	53	84
	24	82	82	49	86	83	78	100	79	44	75	79	65	69	45	83
	48	86	85	50	88	87	83	83	100	46	81	83	69	76	52	88
	74	93	91	73	95	91	91	91	91	100	92	92	78	88	63	95
	82	73	70	39	73	79	74	68	70	39	100	77	65	76	48	74
	96	80	78	44	80	82	79	76	76	42	82	100	72	74	51	81
	120	78	77	44	78	84	81	77	78	44	86	88	100	80	59	82
	408	76	73	40	76	82	75	70	73	42	86	78	69	100	52	75
	17592	84	82	48	82	90	89	77	84	51	90	89	84	86	100	88
	QC	87	85	48	87	80	77	75	75	41	74	76	63	68	47	100

scale (%)
100
75
50
25
0

Figure 9-6 Peak characteristics of non-polar extracts of *Daphnia magna* embryos collected at various times post ovulation following nESI-DIMS analysis in negative ion mode and data processing to yield intensity matrices consisting of peaks present in at least 80% of biological samples (n = 8) within each biological sample class. A) Total number of peaks per class including number of unique (yellow) features. B) Percentage of shared features in pairwise comparisons between all sample types, values in cells indicate the percentage of total peaks present in sample row that are also present in given sample column.

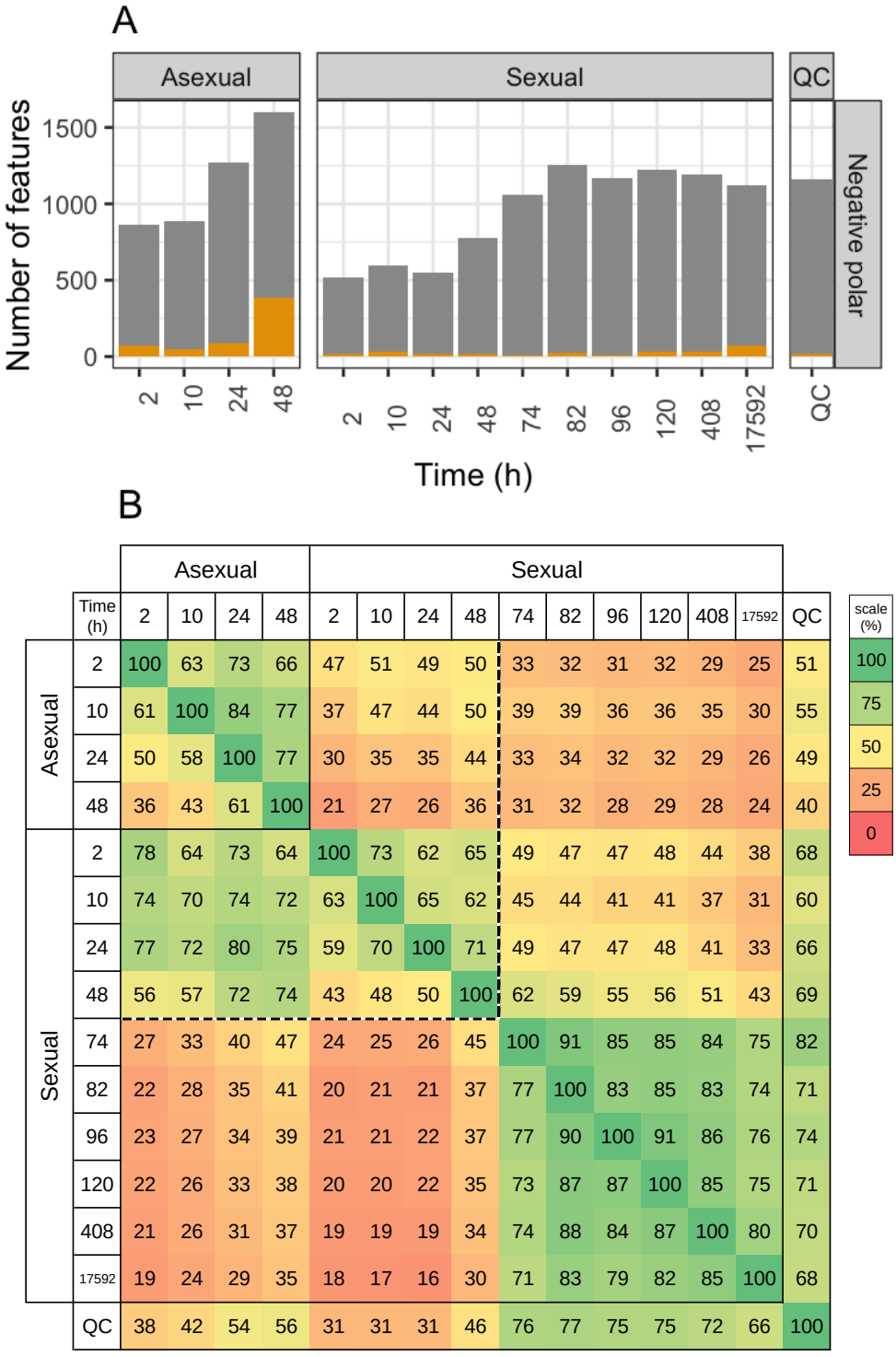


Figure 9-7 Peak characteristics of polar extracts of *Daphnia magna* embryos collected at various times post ovulation following nESI-DIMS analysis in negative ion mode and data processing to yield intensity matrices consisting of peaks present in at least 80% of biological samples (n = 8) within each biological sample class. A) Total number of peaks per class including number of unique (yellow) features. B) Percentage of shared features in pairwise comparisons between all sample types, values in cells indicate the percentage of total peaks present in sample row that are also present in given sample column.

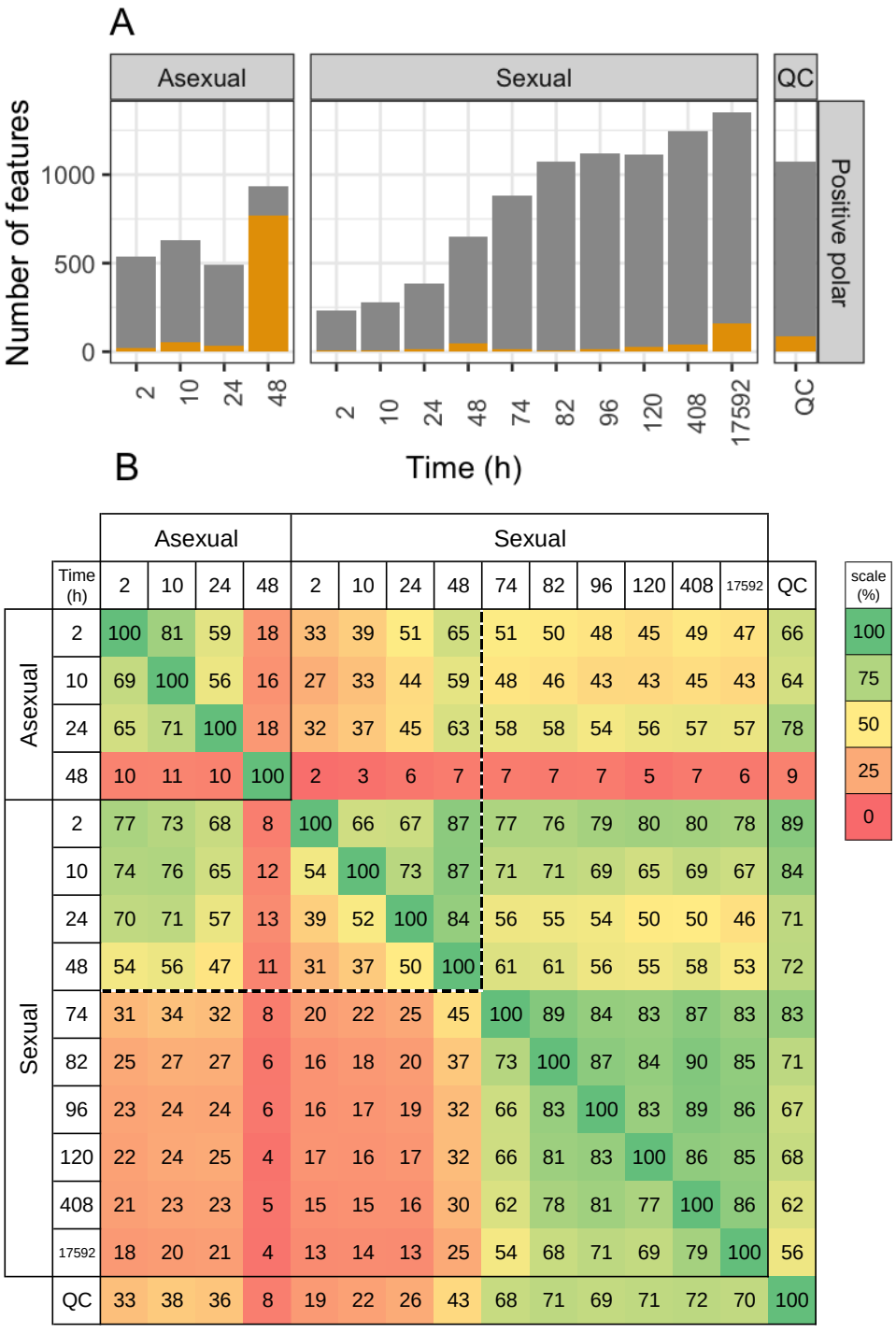


Figure 9-8 Peak characteristics of polar extracts of *Daphnia magna* embryos collected at various times post ovulation following nESI-DIMS analysis in positive ion mode and data processing to yield intensity matrices consisting of peaks present in at least 80% of biological samples (n = 8) within each biological sample class. A) Total number of peaks per class including number of unique (yellow) features. B) Percentage of shared features in pairwise comparisons between all sample types, values in cells indicate the percentage of total peaks present in sample row that are also present in given sample column.

There was a wide range in total numbers of peaks detected in each unique sample type following blank filtering (refer to methods) of the three respective analyses of *D. magna* embryo samples. The number of peaks detected in negative ion analysis of non-polar extracts shown in Figure 9-6, ranges between 493 and 1654 peaks with a mean of 1,066 (SD = 327) peaks. Negative ion analysis of polar extracts (Figure 9-7) yielded between 520 and 1,600 peaks with an average of 1,016 (SD = 307) peaks. Numbers of peaks detected were slightly lower for positive ion analysis of polar extracts (Figure 9-8) at between 230 and 1,352 peaks with an average of 799 (SD = 364) peaks.

Some individual sample types appeared to be particularly low in peak counts e.g. the lowest numbers of peaks detected in both polar analyses were in diapause-destined embryo samples collected 2 h after ovulation (Table 9-5). This potentially explains why a large number of biological samples at this time point were removed during missing value filtering. These samples may have a reduced peak count due to low intensity peaks not being distinguishable from noise features and so are removed during data analysis. However, at this stage we are unable to ascertain whether this is in relation to their developmental stage and potentially low number of cells, the amount of biomaterial or a truly less complex metabolite composition.

There appeared to be no specific relationship between sample type and peak count in the lipidomic dataset, although perhaps in general subitaneous embryo extracts appeared to contain more peaks. There were a few sporadic sample types (subitaneous embryos collected 24 h and sexually-produced embryos collected 74 h and 17,592 h (2 y) post-ovulation; Figure 9-6; A) that were specifically quite low in number of spectral features. In the analyses of polar extracts, there does appear to

be some relationship between peak number and time post-ovulation. In negative ion mode, the three sampling time points prior to the onset of developmental arrest (2–24 h) have a comparably lower peak count than embryos collected after entry into diapause. A similar pattern also appears in the positive ion analysis with a trend towards an increase in peaks detected with time post-ovulation (Figure 9-8; A) Investigation of the number of peaks that were unique to each sample type revealed that within a given analysis the number of unique peaks at specific times varied widely and is summarised in Table 9-6. The analysis of non-polar extracts was the most consistent across sample types ranging between 0.2% and 12% of the total peaks within a sample class being unique to that class with a mean of 3.6% (SD = 3.3%). Negative ion analysis of polar extracts (Figure 9-7) yielded between 0.9% and 23.9% unique peaks with an average of 4.9% (SD = 5.7%). For positive ion analysis of polar extracts (Figure 9-8) between 0.7% and 82.3% of peaks were unique to a given sample type with an average of 9.8% (SD = 20.3%) of peaks. This observation provides a possible explanation for the high number of missing values and thus removal of subitaneous samples collected 48 h after ovulation. Spectra from these samples contained a comparatively large number of peaks (934) however, 769 peaks, over 82% of those detected, where found only in this sample class. Suggesting that the metabolic composition of this sample type is quite different from the other embryo samples collected, which can be seen in Figure 9-8, sharing as little as 2% of detected peaks with other embryos samples.

Table 9-6 The proportion of metabolite peaks detected during the nESI-DIMS analysis of subitaneous and diapause embryo samples that were unique to each respective time point sampled.

		Negative non-polar metabolomics			Negative polar metabolomics			Positive polar metabolomics		
	Time (h)	Total	Unique	Unique (%)	Total	Unique	Unique (%)	Total	Unique	Unique (%)
Subitaneous	**2**	1654	204	12.3	866	74	8.5	537	19	3.5
	10	1376	41	3.0	888	49	5.5	627	55	8.8
	24	683	13	1.9	1273	85	6.7	494	36	7.3
	48	1614	150	9.3	1600	382	23.9	934	769	82.3
Diapause-destined	**2**	1220	43	3.5	520	14	2.7	230	5	2.2
	10	1038	8	0.8	599	29	4.8	278	8	2.9
	24	1035	23	2.2	550	18	3.3	388	16	4.1
	48	985	9	0.9	778	19	2.4	651	46	7.1
	74	493	1	0.2	1062	12	1.1	880	12	1.4
	82	1150	62	5.4	1253	22	1.8	1072	8	0.7
	96	1079	18	1.7	1166	10	0.9	1121	15	1.3
	120	876	43	4.9	1224	30	2.5	1111	30	2.7
	408	1024	42	4.1	1193	33	2.8	1243	42	3.4
	17592	614	17	2.8	1118	68	6.1	1352	161	11.9
	QC	1142	20	1.8	1157	13	1.1	1074	84	7.8

Additional observations regarding the number of peaks shared between any two sample types revealed clear relationships between biological samples, including large blocks of green cells indicating a high percentage of shared peaks and conversely, areas of orange/red cells representing low similarity in terms of peak composition with other samples. The influence of highly unique samples (e.g. subitaneous, 48 h embryos; Figure 9-8) and those particularly low in peak count (e.g. subitaneous 24 h and diapause-destined, 74 h and 17,592 h (2 y) post-ovulation embryos; Figure 9-6), as anticipated, lead to particularly low percentages of shared features with other sample types.

Many of the peaks in lipidomic samples collected from sexually produced embryos (Figure 9-6; B) were shared amongst time points and were also observed in subitaneous embryos. However, the distinct section of yellow cells in the upper right corner of this figure indicates that there is less consistency with subitaneous embryo peaks with an average of 53% (SD = 12%) found also within diapausing and diapause-destined embryos meaning that just under half of the peaks are unique to subitaneous embryos only.

Similar patterns were also apparent when conducting pairwise comparisons of the peak composition between polar embryo extracts. With clear blocks of high and low percentages of shared peaks dividing subitaneous, diapause-destined and diapausing embryos already in developmental arrest. Overall, samples from sexually produced embryos already in diapause i.e. 74 h onwards, exhibited a high percentage of shared peaks with other samples within the diapause period. For the negative ion analysis of polar extracts an average of 82% (SD = 5.5%) of peaks were shared between samples collected during diapause and in positive ion mode, 79% (SD = 9.1%). However, these diapause samples shared very low percentages of peaks with subitaneous and pre-diapause embryos e.g. an average of just 22% (SD = 8.8%) of peaks in pre-diapause positive ion analysis of polar extracts were also present in diapause (+74 h) samples. Pre-diapause negative polar samples also shared a high percentage of peaks with subitaneous embryos (mean = 71% (SD = 7.1%)) whilst conversely, subitaneous embryos shared on average just 39% (SD = 9.8%) of peaks with diapause-embryos over the comparable time points (2–48 h). With a similar pattern also present in samples analysed in positive ion mode but slightly masked by the highly divergent 48 h subitaneous embryos.

These results indicate that there are large numbers of peaks present in diapausing embryos absent in earlier pre-diapause developmental stages. Furthermore, they demonstrate that the peak composition of embryos in subitaneous development is quite different to that of diapause.

The observation of clear demarcations of adjacent time points that share a high number of peaks or that have a low percentage of shared peaks are largely in line with the clear separation of samples groups observed in PCA scores plots presented in Figure 4-2. Together these suggest that there are large-scale shifts in the metabolites present in sexually produced embryos associated with the onset of diapause. Additionally in Section 6 where some of these changes are investigated in terms of changes in the relative intensity of consistent peaks across the samples this analysis revealed that for the majority of peaks, large positive fold changes were observed between pre-diapause and diapause sampling points. Therefore, the general trend for a reduced number of peaks could indicate that many peaks present later in diapause development are too low abundance to be retained during the data processing in pre-diapause samples.

9.3.2. Alternative data processing reveals metabolic features unique to diapause and subitaneous development

An additional quality control sample was included during the acquisition and processing of mass spectra. This was formed from a pool of sexually produced samples only that facilitated the processing of sexual data only with the intention of hopefully retaining any diapause specific peaks that would otherwise be removed (Section 2.14.1.4). This had an effect on the size of the peak matrices with an

additional 11 peaks in the lipidomic dataset, and for polar extracts an additional 57 and 81 peaks were present in the final intensity matrices of positive and negative ion analyses respectively (Table 9-7).

Table 9-7 Table of mass spectral peaks retained in intensity matrices throughout successive filtering stages during the processing of data obtained from the nESI-DIMS mass spectrometry negative and positive ion analysis of *Daphnia magna* diapause embryo polar and non-polar extracts. RSD = relative standard deviation.

	Stage	Negative non-polar metabolomics	Negative polar metabolomics	Positive polar metabolomics
1	**Aligned matrix**	43631	26211	25797
2	**Post-blank filtering**	20074	13276	14986
3	**80% sample filter within biological sample class**	1921	1920	2074
4	**80% sample filter across all biological samples**	813	354	344
5	**RSD <30% in QC samples**	805	353	341

Number of peaks

Furthermore, the peak composition of the final data matrices processed including only diapause embryos are quite different. The number of unique and shared peaks between the final intensity matrices from data processed from all biological samples including subitaneous samples and diapause samples alone are shown in Figure 9-9.

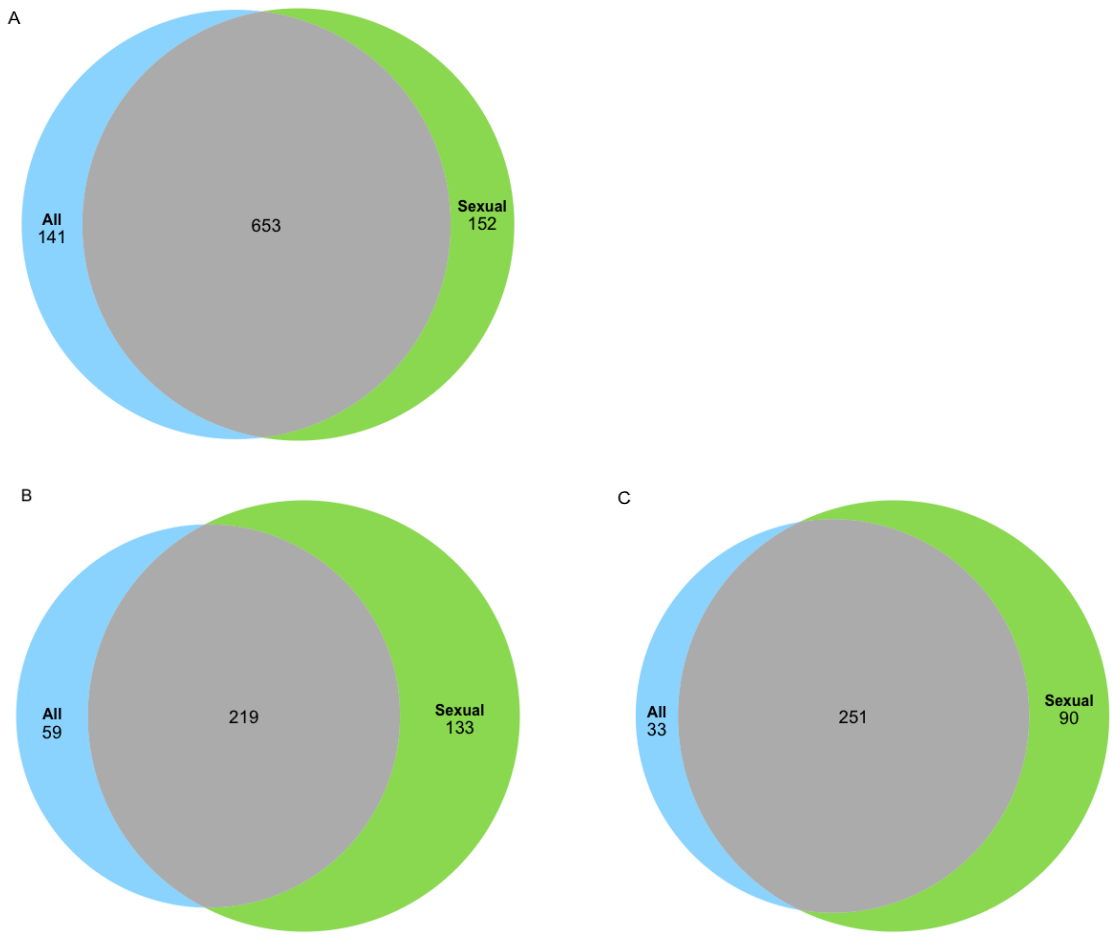


Figure 9-9 Euler diagrams indicating the number of intersecting metabolite peaks between datasets processed including both subitaneous and diapause embryo samples (All) and diapause embryo samples only (Sexual) from the nano-electrospray ionisation direct infusion mass spectrometry analysis of non-polar extracts (A; negative ion mode) and polar extracts analysed in negative ion mode (B) and positive ion mode (C).

As anticipated, additional diapause specific peaks were present in results matrices when only diapause embryos were considered. The percentage of total peaks detected in each analysis that were unique to sexually produced embryos ranged from 19% of peaks in lipidomic datasets to 26% and 37% in positive ion and negative ion analysis respectively of polar extracts.

However, the processing of samples in this manner had the unanticipated effect of the loss of some additional samples during the missing value filtering stage. Notably, only 2 replicates extracts of 10 h samples analysed in negative ion mode and 24 h samples analysed in positive ion mode had less than 20% missing values (Table 9-8)

a loss of 4 and 3 replicates respectively in comparison with the analysis containing all samples (Table 9-5). Furthermore, the number of 2 h samples retained did not increase, suggesting that the metabolic composition of early sexually produced embryos, prior to the onset of diapause are distinctly different to embryos at later stages in diapause development.

Table 9-8 Table of the number of biological replicates retained in final data matrices (bold) following nESI-DIMS metabolomics analysis of polar and non-polar extracts of sexual *Daphnia magna* embryos collected at various time points post-ovulation. All time points contained an initial 8 biological replicates and 'processed' columns refer to the number of samples for which data was successfully acquired.

		Negative non-polar metabolomics		Negative polar metabolomics		Positive polar metabolomics	
Time (h)	**Analysed**	**Processed**	**Final matrix**	**Processed**	**Final matrix**	**Processed**	**Final matrix**
2	8	7	7	8	3	8	3
10	8	8	8	8	2	8	5
24	8	7	6	8	8	8	2
48	8	8	8	8	8	8	8
74	8	8	6	8	8	8	7
82	8	8	8	8	8	8	8
96	8	8	7	8	8	8	8
120	8	8	7	8	8	8	8
408	8	8	8	7	7	8	8
17592	8	8	7	8	7	7	7
Total	80	78	**72**	79	**67**	79	**64**

9.4. Leverage/SPE plots of ANOVA-simultaneous component analysis (ASCA) variable sub-models for lipidomic and metabolomic datasets

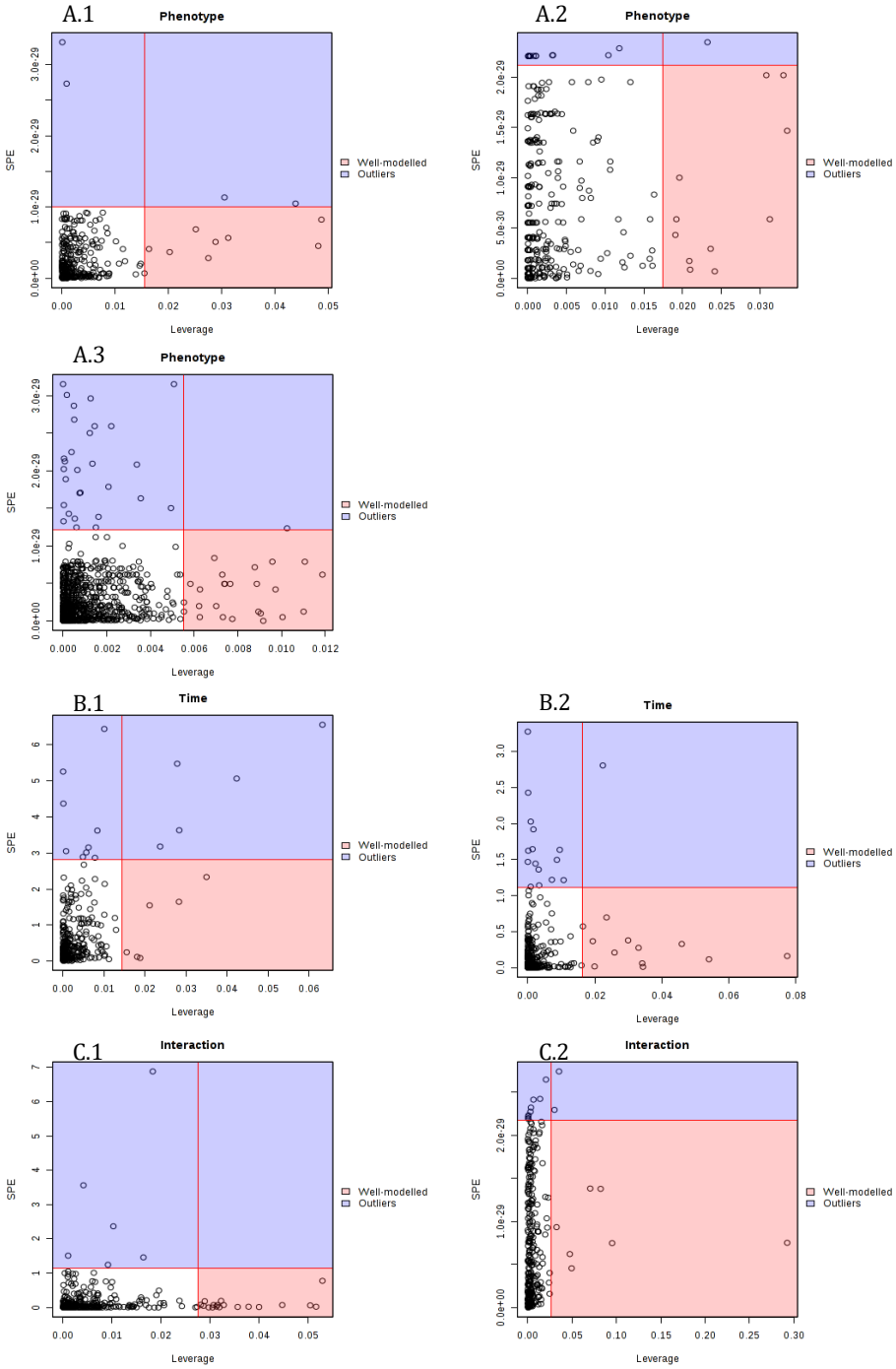


Figure 9-10 Leverage/SPE scatter plots of ANOVA-simultaneous component analysis (ASCA) variable submodels for (A) phenotype (embryo type), (B) time (h post-ovulation) and the (C) interactions thereof for analysis conducted in negative ion mode on polar extracts (.1) positive ion mode on polar extracts (.2) and negative ion analysis on non-polar extracts (.3). Permutation testing of sub-models for time and interactions were insignificant ($p > 0.05$) for analysis of non-polar extracts.

9.5. Gene ontology annotation and overrepresentation analysis of genes differentially expressed in *Daphnia magna* sexual and asexual embryos 2–48 h post-ovulation

Table 9-9 Gene ontology overrepresentation analysis of genes differentially expressed with time (h) post-ovulation after false discovery rate correction ($p < 0.1$). Background contained 19,244 genes, and original input contained 4989 genes

Domain	PANTHER GO-Slim term	Background	Input	Fold enrichment	raw *p*-value	FDR
Biological Process	DNA metabolic process (GO:0006259)	254	157	1.19	0.082	19.200
	DNA repair (GO:0006281)	128	84	1.27	0.094	11.000
	DNA replication (GO:0006260)	97	61	1.21	0.238	7.000
	lipid transport (GO:0006869)	39	27	1.33	0.244	6.380
	neuromuscular synaptic transmission (GO:0007274)	15	12	1.54	0.310	5.200
	cellular amino acid catabolic process (GO:0009063)	44	28	1.23	0.386	4.540
	vitamin transport (GO:0051180)	3	3	1.93	0.417	4.090
	peroxisomal transport (GO:0043574)	9	7	1.5	0.436	3.660
	primary metabolic process (GO:0044238)	2755	1459	1.02	0.453	3.320
	regulation of biological process (GO:0050789)	694	376	1.04	0.482	3.240
	mRNA splicing, via spliceosome (GO:0000398)	121	70	1.12	0.490	3.030
	sensory perception of sound (GO:0007605)	5	4	1.54	0.502	2.950
	cytokine-mediated signaling pathway (GO:0019221)	5	4	1.54	0.502	2.880
	polysaccharide metabolic process (GO:0005976)	57	34	1.15	0.508	2.780
	lipid metabolic process (GO:0006629)	205	115	1.08	0.512	2.740
	protein glycosylation (GO:0006486)	60	36	1.16	0.517	2.700
	mRNA processing (GO:0006397)	174	98	1.09	0.519	2.650
	carbohydrate metabolic process (GO:0005975)	215	120	1.08	0.522	2.550
	response to stimulus (GO:0050896)	768	413	1.04	0.543	2.610
	circadian rhythm (GO:0007623)	7	5	1.38	0.558	2.340
	response to toxic substance (GO:0009636)	7	5	1.38	0.558	2.300
	female gamete generation (GO:0007292)	7	5	1.38	0.558	2.260

	rhythmic process (GO:0048511)	7	5	1.38	0.558	2.220
	mRNA polyadenylation (GO:0006378)	18	12	1.29	0.564	2.100
	anion transport (GO:0006820)	109	62	1.1	0.571	2.100
	sensory perception of chemical stimulus (GO:0007606)	19	12	1.22	0.575	2.050
	calcium-mediated signaling (GO:0019722)	20	13	1.25	0.582	1.950
	metabolic process (GO:0008152)	3307	1738	1.01	0.582	1.980
	chromosome segregation (GO:0007059)	60	35	1.12	0.588	1.920
	cell proliferation (GO:0008283)	21	13	1.19	0.593	1.910
	chromatin remodeling (GO:0006338)	21	13	1.19	0.593	1.880
	lysosomal transport (GO:0007041)	9	6	1.29	0.599	1.880
	regulation of binding (GO:0051098)	2	2	1.93	0.610	1.840
	sensory perception of pain (GO:0019233)	2	2	1.93	0.610	1.810
	prosthetic group metabolic process (GO:0051189)	2	2	1.93	0.610	1.790
	protein ADP-ribosylation (GO:0006471)	2	2	1.93	0.610	1.770
	disaccharide metabolic process (GO:0005984)	2	2	1.93	0.610	1.750
	biological regulation (GO:0065007)	885	472	1.03	0.610	1.860
	sensory perception of taste (GO:0050909)	12	8	1.29	0.639	1.750
	phosphate-containing compound metabolic process (GO:0006796)	830	442	1.03	0.643	1.720
	cellular amino acid biosynthetic process (GO:0008652)	55	32	1.12	0.650	1.700
	RNA metabolic process (GO:0016070)	891	474	1.03	0.653	1.670
	acyl-CoA metabolic process (GO:0006637)	14	9	1.24	0.662	1.670
	apoptotic process (GO:0006915)	96	54	1.08	0.665	1.610
	death (GO:0016265)	99	55	1.07	0.670	1.610
	cell death (GO:0008219)	99	55	1.07	0.670	1.590
	RNA splicing, via transesterification reactions (GO:0000375)	106	59	1.07	0.680	1.570
	response to stress (GO:0006950)	310	167	1.04	0.695	1.510
	muscle organ development (GO:0007517)	4	3	1.45	0.697	1.500
	fertilization (GO:0009566)	4	3	1.45	0.697	1.490
	heart development (GO:0007507)	4	3	1.45	0.697	1.480

	anterior/posterior axis specification (GO:0009948)	4	3	1.45	0.697	1.460
	fatty acid beta-oxidation (GO:0006635)	18	11	1.18	0.697	1.450
	steroid metabolic process (GO:0008202)	43	25	1.12	0.701	1.440
	protein acetylation (GO:0006473)	45	26	1.11	0.707	1.410
	regulation of transcription from RNA polymerase II promoter (GO:0006357)	256	138	1.04	0.707	1.420
	nucleobase-containing compound metabolic process (GO:0006139)	1602	843	1.01	0.709	1.400
	mitosis (GO:0007067)	131	72	1.06	0.710	1.390
	cellular protein modification process (GO:0006464)	470	251	1.03	0.717	1.380
	cytokinesis (GO:0000910)	57	32	1.08	0.737	1.400
	homeostatic process (GO:0042592)	106	58	1.05	0.741	1.380
	catabolic process (GO:0009056)	722	382	1.02	0.741	1.370
	sensory perception of smell (GO:0007608)	6	4	1.29	0.743	1.360
	transcription, DNA-dependent (GO:0006351)	456	243	1.03	0.744	1.350
	mRNA 3'-end processing (GO:0031124)	29	17	1.13	0.756	1.370
	DNA catabolic process (GO:0006308)	8	5	1.21	0.774	1.340
	organelle organization (GO:0006996)	690	364	1.02	0.787	1.330
	regulation of cell cycle (GO:0051726)	93	51	1.06	0.791	1.330
	tRNA metabolic process (GO:0006399)	100	54	1.04	0.798	1.330
	phospholipid metabolic process (GO:0006644)	102	55	1.04	0.800	1.320
	regulation of phosphate metabolic process (GO:0019220)	200	107	1.03	0.808	1.320
	pteridine-containing compound metabolic process (GO:0042558)	12	7	1.12	0.812	1.310
	RNA catabolic process (GO:0006401)	63	34	1.04	0.831	1.290
	mitochondrial translation (GO:0032543)	16	9	1.08	0.835	1.290
	cellular component organization (GO:0016043)	974	511	1.01	0.840	1.270
	cellular amino acid metabolic process (GO:0006520)	177	94	1.02	0.846	1.280
	ion transport (GO:0006811)	178	95	1.03	0.847	1.270
	transmembrane receptor protein serine/threonine kinase signaling pathway (GO:0007178)	19	11	1.12	0.848	1.260

	fatty acid metabolic process (GO:0006631)	83	45	1.05	0.852	1.260
	glycogen metabolic process (GO:0005977)	25	14	1.08	0.866	1.260
	protein methylation (GO:0006479)	31	17	1.06	0.879	1.240
	localization (GO:0051179)	949	496	1.01	0.884	1.240
	cell differentiation (GO:0030154)	145	77	1.02	0.887	1.230
	anatomical structure morphogenesis (GO:0009653)	37	20	1.04	0.889	1.210
	transcription from RNA polymerase II promoter (GO:0006366)	350	184	1.01	0.889	1.220
	protein lipidation (GO:0006497)	41	22	1.03	0.895	1.210
	negative regulation of apoptotic process (GO:0043066)	42	23	1.06	0.896	1.200
	proteolysis (GO:0006508)	257	135	1.01	0.914	1.210
	cellular process (GO:0009987)	3836	1994	1	0.929	1.210
	nitrogen compound metabolic process (GO:0006807)	1508	784	1	0.962	1.240
	JAK-STAT cascade (GO:0007259)	1	1	1.93	1.000	1.190
	hemopoiesis (GO:0030097)	1	1	1.93	1.000	1.190
	angiogenesis (GO:0001525)	1	1	1.93	1.000	1.180
	pinocytosis (GO:0006907)	1	1	1.93	1.000	1.170
	endoderm development (GO:0007492)	1	1	1.93	1.000	1.170
	regulation of sequence-specific DNA binding transcription factor activity (GO:0051090)	9	5	1.07	1.000	1.100
	asymmetric protein localization (GO:0008105)	9	5	1.07	1.000	1.090
	induction of apoptosis (GO:0006917)	11	6	1.05	1.000	1.050
	transcription initiation from RNA polymerase II promoter (GO:0006367)	30	16	1.03	1.000	1.040
	glycolysis (GO:0006096)	15	8	1.03	1.000	1.140
	tricarboxylic acid cycle (GO:0006099)	19	10	1.01	1.000	1.040
	sensory perception (GO:0007600)	57	30	1.01	1.000	1.050
	vesicle-mediated transport (GO:0016192)	339	176	1	1.000	1.280
	protein localization (GO:0008104)	314	163	1	1.000	1.280
	chromatin organization (GO:0006325)	160	83	1	1.000	1.270
Cellular Component	nucleus (GO:0005634)	1092	602	1.06	0.200	3.210
	nucleolus (GO:0005730)	106	64	1.16	0.330	2.110

	Golgi apparatus (GO:0005794)	165	95	1.11	0.429	1.960
	coated pit (GO:0005905)	4	4	1.93	0.458	1.830
	neuronal cell body (GO:0043025)	5	4	1.54	0.502	1.780
	presynaptic membrane (GO:0042734)	5	4	1.54	0.502	1.690
	lipid particle (GO:0005811)	6	5	1.61	0.527	1.680
	neuron projection (GO:0043005)	80	46	1.11	0.572	1.530
	axon (GO:0030424)	16	10	1.21	0.681	1.410
	cell part (GO:0044464)	3149	1649	1.01	0.710	1.420
	SNARE complex (GO:0031201)	20	12	1.16	0.711	1.380
	peroxisome (GO:0005777)	22	13	1.14	0.723	1.360
	dendrite (GO:0030425)	25	15	1.16	0.739	1.350
	organelle (GO:0043226)	2231	1169	1.01	0.757	1.350
	intracellular (GO:0005622)	3052	1594	1.01	0.793	1.300
	integral to membrane (GO:0016021)	387	205	1.02	0.825	1.320
	vesicle coat (GO:0030120)	23	13	1.09	0.861	1.220
	cell projection (GO:0042995)	134	71	1.02	0.882	1.180
	plasma membrane (GO:0005886)	465	243	1.01	0.935	1.130
	Unclassified (UNCLASSIFIED)	5982	3104	1	0.986	1.170
	basal part of cell (GO:0045178)	1	1	1.93	1.000	1.100
	microvillus (GO:0005902)	1	1	1.93	1.000	1.080
	neuromuscular junction (GO:0031594)	3	2	1.29	1.000	1.070
	postsynaptic membrane (GO:0045211)	11	6	1.05	1.000	1.030
Molecular Function	transferase activity, transferring glycosyl groups (GO:0016757)	93	64	1.33	0.090	4.110
	hydrolase activity, acting on ester bonds (GO:0016788)	266	160	1.16	0.133	4.030
	catalytic activity (GO:0003824)	2434	1314	1.04	0.181	3.660
	RNA helicase activity (GO:0003724)	24	19	1.53	0.197	3.580
	transferase activity (GO:0016740)	742	414	1.08	0.219	3.330
	hydrolase activity, hydrolyzing N-glycosyl compounds (GO:0016799)	12	10	1.61	0.268	3.480
	transmembrane transporter activity (GO:0022857)	389	221	1.1	0.276	3.140
	nucleotidyltransferase activity (GO:0016779)	66	42	1.23	0.309	2.960
	poly(A) RNA binding (GO:0008143)	5	5	1.93	0.325	2.960
	endoribonuclease activity (GO:0004521)	34	23	1.3	0.329	2.860

	transporter activity (GO:0005215)	479	267	1.07	0.342	2.700
	hydrolase activity, hydrolyzing O-glycosyl compounds (GO:0004553)	38	25	1.27	0.354	2.680
	helicase activity (GO:0004386)	57	36	1.22	0.380	2.770
	kinase regulator activity (GO:0019207)	35	23	1.27	0.406	2.840
	nuclease activity (GO:0004518)	104	62	1.15	0.410	2.670
	DNA N-glycosylase activity (GO:0019104)	9	7	1.5	0.436	2.730
	galactosidase activity (GO:0015925)	4	4	1.93	0.458	2.690
	phosphorylase activity (GO:0004645)	4	4	1.93	0.458	2.600
	chromatin binding (GO:0003682)	104	61	1.13	0.458	2.520
	ubiquitin-protein ligase activity (GO:0004842)	47	29	1.19	0.468	2.440
	motor activity (GO:0003774)	49	30	1.18	0.477	2.410
	DNA-methyltransferase activity (GO:0009008)	5	4	1.54	0.502	2.470
	DNA helicase activity (GO:0003678)	40	25	1.21	0.512	2.450
	transmembrane receptor protein serine/threonine kinase activity (GO:0004675)	6	5	1.61	0.527	2.460
	enzyme regulator activity (GO:0030234)	188	105	1.08	0.534	2.370
	translation initiation factor activity (GO:0003743)	45	27	1.16	0.536	2.320
	ligase activity (GO:0016874)	156	88	1.09	0.540	2.230
	SNAP receptor activity (GO:0005484)	16	11	1.33	0.543	2.200
	phosphatase activity (GO:0016791)	96	55	1.1	0.547	2.160
	mRNA binding (GO:0003729)	99	57	1.11	0.553	2.010
	non-membrane spanning protein tyrosine kinase activity (GO:0004715)	7	5	1.38	0.558	1.920
	RNA-directed DNA polymerase activity (GO:0003964)	7	5	1.38	0.558	1.880
	nucleotide phosphatase activity (GO:0019204)	7	5	1.38	0.558	1.850
	DNA replication origin binding (GO:0003688)	8	6	1.45	0.575	1.870
	DNA-directed RNA polymerase activity (GO:0003899)	35	21	1.16	0.576	1.840
	hydrolase activity (GO:0016787)	993	529	1.03	0.607	1.780
	lipid transporter activity (GO:0005319)	23	14	1.17	0.608	1.760

	receptor regulator activity (GO:0030545)	2	2	1.93	0.610	1.730
	calcium channel regulator activity (GO:0005246)	2	2	1.93	0.610	1.710
	DNA primase activity (GO:0003896)	2	2	1.93	0.610	1.680
	deaminase activity (GO:0019239)	12	8	1.29	0.639	1.740
	guanyl-nucleotide exchange factor activity (GO:0005085)	53	31	1.13	0.645	1.730
	kinase activity (GO:0016301)	289	157	1.05	0.648	1.710
	transmembrane receptor protein tyrosine kinase activity (GO:0004714)	32	19	1.14	0.658	1.690
	antioxidant activity (GO:0016209)	36	21	1.12	0.676	1.600
	oxidoreductase activity (GO:0016491)	360	194	1.04	0.681	1.590
	microtubule motor activity (GO:0003777)	39	23	1.14	0.687	1.580
	signal transducer activity (GO:0004871)	168	92	1.06	0.692	1.560
	protein kinase activity (GO:0004672)	168	92	1.06	0.692	1.540
	phosphoric diester hydrolase activity (GO:0008081)	18	11	1.18	0.697	1.530
	methyltransferase activity (GO:0008168)	87	48	1.06	0.716	1.530
	lyase activity (GO:0016829)	131	71	1.04	0.765	1.530
	small GTPase regulator activity (GO:0005083)	76	42	1.07	0.770	1.520
	amino acid transmembrane transporter activity (GO:0015171)	38	21	1.07	0.785	1.520
	RNA methyltransferase activity (GO:0008173)	41	23	1.08	0.792	1.520
	intramolecular transferase activity (GO:0016866)	12	7	1.12	0.812	1.510
	voltage-gated potassium channel activity (GO:0005249)	13	8	1.19	0.818	1.500
	exodeoxyribonuclease activity (GO:0004529)	14	8	1.1	0.825	1.500
	DNA-directed DNA polymerase activity (GO:0003887)	14	8	1.1	0.825	1.490
	phosphoprotein phosphatase activity (GO:0004721)	63	34	1.04	0.831	1.480
	nucleotide kinase activity (GO:0019201)	21	12	1.1	0.855	1.500
	peroxidase activity (GO:0004601)	24	13	1.04	0.864	1.500
	damaged DNA binding (GO:0003684)	25	14	1.08	0.866	1.470
	lipase activity (GO:0016298)	31	17	1.06	0.879	1.480

	isomerase activity (GO:0016853)	62	33	1.03	0.914	1.500
	transferase activity, transferring acyl groups (GO:0016746)	114	60	1.01	0.936	1.480
	DNA ligase activity (GO:0003909)	1	1	1.93	1.000	1.360
	core RNA polymerase I binding transcription factor activity (GO:0001181)	1	1	1.93	1.000	1.350
	RNA polymerase III transcription factor binding transcription factor activity (GO:0001007)	1	1	1.93	1.000	1.340
	sequence-specific DNA binding RNA polymerase I transcription factor activity (GO:0001167)	1	1	1.93	1.000	1.330
	receptor inhibitor activity (GO:0030547)	1	1	1.93	1.000	1.320
	receptor activator activity (GO:0030546)	1	1	1.93	1.000	1.310
	potassium channel regulator activity (GO:0015459)	1	1	1.93	1.000	1.300
	DNA polymerase processivity factor activity (GO:0030337)	1	1	1.93	1.000	1.290
	RNA polymerase I transcription factor binding transcription factor activity (GO:0001082)	1	1	1.93	1.000	1.280
	phosphatase activator activity (GO:0019211)	1	1	1.93	1.000	1.270
	sequence-specific DNA binding RNA polymerase III transcription factor activity (GO:0001034)	1	1	1.93	1.000	1.260
	aspartic-type endopeptidase activity (GO:0004190)	3	2	1.29	1.000	1.130
	carbohydrate phosphatase activity (GO:0019203)	3	2	1.29	1.000	1.120
	aldehyde-lyase activity (GO:0016832)	3	2	1.29	1.000	1.120
	transforming growth factor beta-activated receptor activity (GO:0005024)	3	2	1.29	1.000	1.110
	cytokine activity (GO:0005125)	5	3	1.16	1.000	1.240
	guanylate cyclase activity (GO:0004383)	5	3	1.16	1.000	1.230
	protein disulfide isomerase activity (GO:0003756)	7	4	1.1	1.000	1.260
	transforming growth factor beta receptor binding (GO:0005160)	9	5	1.07	1.000	1.100
	endodeoxyribonuclease activity (GO:0004520)	13	7	1.04	1.000	1.220
	kinase activator activity (GO:0019209)	13	7	1.04	1.000	1.210

	glucosidase activity (GO:0015926)	15	8	1.03	1.000	1.210
	ligand-activated sequence-specific DNA binding RNA polymerase II transcription factor activity (GO:0004879)	15	8	1.03	1.000	1.200
	serine-type endopeptidase inhibitor activity (GO:0004867)	15	8	1.03	1.000	1.190
	deacetylase activity (GO:0019213)	15	8	1.03	1.000	1.180
	exoribonuclease activity (GO:0004532)	17	9	1.02	1.000	1.140
	neurotransmitter transporter activity (GO:0005326)	17	9	1.02	1.000	1.140
	carboxy-lyase activity (GO:0016831)	19	10	1.01	1.000	1.050
	acetyltransferase activity (GO:0016407)	57	30	1.01	1.000	1.060
	translation regulator activity (GO:0045182)	80	42	1.01	1.000	1.170
	enzyme activator activity (GO:0008047)	61	32	1.01	1.000	1.030
	ligand-gated ion channel activity (GO:0015276)	44	23	1.01	1.000	1.010

Table 9-10 Gene ontology overrepresentation analysis of genes differentially expressed based on an interaction between embryo type and time (h) post-ovulation after false discovery rate correction ($p < 0.1$). Background contained 19,244 genes, and original input contained 3013 genes

Domain	PANTHER GO-Slim term	Background	Input	Fold enrichment	raw *p*-value	FDR
Biological process	reproduction (GO:0000003)	74	36	1.54	0.033	3.910
	cell cycle (GO:0007049)	361	140	1.23	0.037	2.910
	meiosis (GO:0007126)	36	20	1.76	0.058	3.400
	fatty acid metabolic process (GO:0006631)	83	38	1.45	0.068	3.200
	fatty acid biosynthetic process (GO:0006633)	22	13	1.88	0.075	2.930
	protein folding (GO:0006457)	67	31	1.47	0.076	2.540
	lipid metabolic process (GO:0006629)	205	81	1.25	0.092	2.170
	circadian rhythm (GO:0007623)	7	6	2.72	0.095	2.040
	rhythmic process (GO:0048511)	7	6	2.72	0.095	1.870
	metabolic process (GO:0008152)	3307	1092	1.05	0.101	1.830
	carbohydrate metabolic process (GO:0005975)	215	83	1.23	0.114	1.790
	phospholipid metabolic process (GO:0006644)	102	43	1.34	0.117	1.620
	phosphate-containing compound metabolic process (GO:0006796)	830	289	1.11	0.133	1.640
	regulation of cell cycle (GO:0051726)	93	38	1.3	0.181	1.850
	catabolic process (GO:0009056)	722	250	1.1	0.184	1.800
	cellular amino acid catabolic process (GO:0009063)	44	20	1.44	0.186	1.750
	embryo development (GO:0009790)	26	13	1.59	0.188	1.700
	tricarboxylic acid cycle (GO:0006099)	19	10	1.67	0.193	1.680
	cytokinesis (GO:0000910)	57	25	1.39	0.193	1.620
	cellular amino acid metabolic process (GO:0006520)	177	67	1.2	0.198	1.500
	DNA metabolic process (GO:0006259)	254	93	1.16	0.225	1.610
	cytokine-mediated signaling pathway (GO:0019221)	5	4	2.54	0.232	1.600
	chromatin assembly (GO:0031497)	18	9	1.59	0.261	1.530
	fatty acid beta-oxidation (GO:0006635)	18	9	1.59	0.261	1.500
	response to stress (GO:0006950)	310	110	1.13	0.270	1.440
	tRNA metabolic process (GO:0006399)	100	39	1.24	0.271	1.420

	negative regulation of apoptotic process (GO:0043066)	42	18	1.36	0.288	1.380
	response to toxic substance (GO:0009636)	7	4	1.81	0.308	1.420
	gluconeogenesis (GO:0006094)	22	10	1.44	0.308	1.390
	acyl-CoA metabolic process (GO:0006637)	14	7	1.59	0.311	1.380
	glycogen metabolic process (GO:0005977)	25	11	1.4	0.334	1.430
	biological regulation (GO:0065007)	885	297	1.07	0.334	1.400
	primary metabolic process (GO:0044238)	2755	896	1.03	0.334	1.380
	regulation of sequence-specific DNA binding transcription factor activity (GO:0051090)	9	5	1.76	0.345	1.380
	lysosomal transport (GO:0007041)	9	5	1.76	0.345	1.350
	fertilization (GO:0009566)	4	3	2.38	0.369	1.360
	heart development (GO:0007507)	4	3	2.38	0.369	1.330
	polyphosphate catabolic process (GO:0006798)	4	3	2.38	0.369	1.310
	sensory perception of chemical stimulus (GO:0007606)	19	9	1.5	0.373	1.310
	cell-matrix adhesion (GO:0007160)	10	5	1.59	0.374	1.270
	mitosis (GO:0007067)	131	48	1.16	0.378	1.270
	cell proliferation (GO:0008283)	21	9	1.36	0.400	1.290
	DNA replication (GO:0006260)	97	36	1.18	0.414	1.300
	coenzyme metabolic process (GO:0006732)	74	28	1.2	0.415	1.280
	DNA repair (GO:0006281)	128	46	1.14	0.422	1.240
	organelle organization (GO:0006996)	690	231	1.06	0.423	1.210
	response to stimulus (GO:0050896)	768	256	1.06	0.423	1.230
	homeostatic process (GO:0042592)	106	39	1.17	0.433	1.200
	chromosome segregation (GO:0007059)	60	23	1.22	0.439	1.200
	glycolysis (GO:0006096)	15	7	1.48	0.452	1.220
	neuromuscular synaptic transmission (GO:0007274)	15	7	1.48	0.452	1.210
	7-methylguanosine mRNA capping (GO:0006370)	6	3	1.59	0.455	1.190
	sensory perception of smell (GO:0007608)	6	3	1.59	0.455	1.180
	tRNA aminoacylation for protein translation (GO:0006418)	44	17	1.23	0.455	1.200
	purine nucleobase metabolic process (GO:0006144)	32	13	1.29	0.483	1.220

	response to external stimulus (GO:0009605)	103	37	1.14	0.487	1.220
	gamete generation (GO:0007276)	33	13	1.25	0.490	1.210
	neuron-neuron synaptic transmission (GO:0007270)	33	13	1.25	0.490	1.200
	transmembrane receptor protein serine/threonine kinase signaling pathway (GO:0007178)	19	8	1.34	0.500	1.210
	polysaccharide metabolic process (GO:0005976)	57	21	1.17	0.509	1.210
	regulation of biological process (GO:0050789)	694	229	1.05	0.522	1.230
	lipid transport (GO:0006869)	39	15	1.22	0.523	1.220
	monosaccharide metabolic process (GO:0005996)	62	23	1.18	0.524	1.210
	asymmetric protein localization (GO:0008105)	9	4	1.41	0.526	1.200
	peroxisomal transport (GO:0043574)	9	4	1.41	0.526	1.190
	cytoskeleton organization (GO:0007010)	177	61	1.09	0.540	1.190
	developmental process (GO:0032502)	416	139	1.06	0.542	1.170
	induction of apoptosis (GO:0006917)	11	5	1.44	0.557	1.180
	pteridine-containing compound metabolic process (GO:0042558)	12	5	1.32	0.575	1.180
	sensory perception of taste (GO:0050909)	12	5	1.32	0.575	1.170
	porphyrin-containing compound metabolic process (GO:0006778)	15	6	1.27	0.612	1.190
	DNA recombination (GO:0006310)	38	14	1.17	0.626	1.210
	carbohydrate transport (GO:0008643)	21	8	1.21	0.664	1.210
	chromatin remodeling (GO:0006338)	21	8	1.21	0.664	1.200
	apoptotic process (GO:0006915)	96	33	1.09	0.678	1.180
	cellular amino acid biosynthetic process (GO:0008652)	55	19	1.1	0.684	1.180
	sensory perception (GO:0007600)	57	20	1.11	0.688	1.180
	protein glycosylation (GO:0006486)	60	21	1.11	0.695	1.180
	RNA splicing (GO:0008380)	7	3	1.36	0.711	1.180
	female gamete generation (GO:0007292)	7	3	1.36	0.711	1.170
	mitochondrion organization (GO:0007005)	69	24	1.1	0.714	1.170

	regulation of phosphate metabolic process (GO:0019220)	200	66	1.05	0.717	1.160
	neurological system process (GO:0050877)	202	67	1.05	0.718	1.160
	protein complex biogenesis (GO:0070271)	141	47	1.06	0.731	1.150
	cell adhesion (GO:0007155)	79	27	1.08	0.732	1.150
	biological adhesion (GO:0022610)	79	27	1.08	0.732	1.140
	intracellular signal transduction (GO:0035556)	336	110	1.04	0.735	1.130
	transmembrane receptor protein tyrosine kinase signaling pathway (GO:0007169)	36	13	1.15	0.737	1.120
	cell-cell adhesion (GO:0016337)	37	13	1.12	0.740	1.120
	response to endogenous stimulus (GO:0009719)	85	29	1.08	0.741	1.120
	phagocytosis (GO:0006909)	40	14	1.11	0.750	1.110
	death (GO:0016265)	99	33	1.06	0.759	1.120
	cell death (GO:0008219)	99	33	1.06	0.759	1.110
	cellular process (GO:0009987)	3836	1218	1.01	0.766	1.120
	amino acid transport (GO:0006865)	47	16	1.08	0.768	1.110
	cellular protein modification process (GO:0006464)	470	152	1.03	0.773	1.110
	protein complex assembly (GO:0006461)	140	46	1.04	0.795	1.130
	cell differentiation (GO:0030154)	145	48	1.05	0.799	1.110
	segment specification (GO:0007379)	18	6	1.06	0.816	1.120
	pyrimidine nucleobase metabolic process (GO:0006206)	20	7	1.11	0.822	1.120
	chromatin organization (GO:0006325)	160	52	1.03	0.871	1.150
	steroid metabolic process (GO:0008202)	43	14	1.03	0.877	1.140
	exocytosis (GO:0006887)	77	25	1.03	0.907	1.170
	MAPK cascade (GO:0000165)	91	29	1.01	0.915	1.160
	anion transport (GO:0006820)	109	35	1.02	0.922	1.160
	system process (GO:0003008)	226	72	1.01	0.945	1.180
	behavior (GO:0007610)	9	3	1.06	1.000	1.040
	cellular glucose homeostasis (GO:0001678)	15	5	1.06	1.000	1.190
	RNA catabolic process (GO:0006401)	63	20	1.01	1.000	1.190
	protein lipidation (GO:0006497)	41	13	1.01	1.000	1.210
	multicellular organismal process (GO:0032501)	358	113	1	1.000	1.240

	system development (GO:0048731)	92	29	1	1.000	1.240
Cellular component	nuclear outer membrane-endoplasmic reticulum membrane network (GO:0042175)	143	66	1.46	0.011	0.358
	endoplasmic reticulum (GO:0005783)	229	94	1.3	0.034	0.549
	protein-DNA complex (GO:0032993)	29	15	1.64	0.115	0.814
	neuromuscular junction (GO:0031594)	3	3	3.17	0.153	0.981
	nuclear chromosome (GO:0000228)	96	39	1.29	0.187	1.090
	microtubule (GO:0005874)	68	29	1.35	0.188	1.000
	chromosome (GO:0005694)	148	57	1.22	0.216	1.060
	cell junction (GO:0030054)	22	11	1.59	0.221	1.010
	SNARE complex (GO:0031201)	20	10	1.59	0.282	1.060
	peroxisome (GO:0005777)	22	10	1.44	0.308	1.090
	Unclassified (UNCLASSIFIED)	5982	1916	1.02	0.312	1.050
	cytoplasmic membrane-bounded vesicle (GO:0016023)	88	32	1.15	0.519	1.110
	Golgi apparatus (GO:0005794)	165	57	1.1	0.527	1.090
	proton-transporting ATP synthase complex (GO:0045259)	11	5	1.44	0.557	1.020
	membrane (GO:0016020)	905	294	1.03	0.644	1.060
	mitochondrion (GO:0005739)	255	84	1.05	0.699	1.070
	endosome (GO:0005768)	66	23	1.11	0.708	1.050
	neuron projection (GO:0043005)	80	27	1.07	0.734	1.070
	plasma membrane (GO:0005886)	465	151	1.03	0.735	1.050
	integral to membrane (GO:0016021)	387	126	1.03	0.751	1.020
	microtubule organizing center (GO:0005815)	62	21	1.07	0.796	1.060
Molecular function	catalytic activity (GO:0003824)	2434	840	1.1	0.009	1.620
	transferase activity (GO:0016740)	742	279	1.19	0.009	0.847
	transferase activity, transferring glycosyl groups (GO:0016757)	93	46	1.57	0.016	0.971
	microtubule binding (GO:0008017)	55	30	1.73	0.021	0.948
	damaged DNA binding (GO:0003684)	25	15	1.9	0.061	1.390
	nucleotide kinase activity (GO:0019201)	21	13	1.96	0.068	1.370
	cytoskeletal protein binding (GO:0008092)	123	52	1.34	0.075	1.140

	chromatin binding (GO:0003682)	104	45	1.37	0.082	1.150
	exodeoxyribonuclease activity (GO:0004529)	14	9	2.04	0.092	1.110
	kinase activity (GO:0016301)	289	110	1.21	0.095	1.080
	phosphorylase activity (GO:0004645)	4	4	3.17	0.100	1.070
	oxidoreductase activity (GO:0016491)	360	133	1.17	0.118	1.130
	endodeoxyribonuclease activity (GO:0004520)	13	8	1.95	0.131	1.190
	SNAP receptor activity (GO:0005484)	16	9	1.79	0.163	1.340
	isomerase activity (GO:0016853)	62	27	1.38	0.170	1.350
	nucleotidyltransferase activity (GO:0016779)	66	28	1.35	0.183	1.390
	endoribonuclease activity (GO:0004521)	34	16	1.49	0.186	1.350
	carboxy-lyase activity (GO:0016831)	19	10	1.67	0.193	1.350
	ligase activity (GO:0016874)	156	59	1.2	0.227	1.480
	acetyltransferase activity (GO:0016407)	57	24	1.34	0.240	1.510
	protein kinase activity (GO:0004672)	168	63	1.19	0.243	1.480
	nuclease activity (GO:0004518)	104	40	1.22	0.281	1.500
	single-stranded DNA binding (GO:0003697)	42	18	1.36	0.288	1.500
	transferase activity, transferring acyl groups (GO:0016746)	114	43	1.2	0.302	1.530
	RNA-directed DNA polymerase activity (GO:0003964)	7	4	1.81	0.308	1.510
	protein disulfide isomerase activity (GO:0003756)	7	4	1.81	0.308	1.470
	calmodulin binding (GO:0005516)	35	15	1.36	0.320	1.500
	motor activity (GO:0003774)	49	20	1.3	0.324	1.470
	DNA-methyltransferase activity (GO:0009008)	5	3	1.9	0.407	1.580
	double-stranded DNA binding (GO:0003690)	26	11	1.34	0.440	1.600
	pyrophosphatase activity (GO:0016462)	367	124	1.07	0.484	1.570
	transmembrane transporter activity (GO:0022857)	389	131	1.07	0.496	1.580
	kinase regulator activity (GO:0019207)	35	14	1.27	0.502	1.570
	DNA-directed RNA polymerase activity (GO:0003899)	35	14	1.27	0.502	1.550
	protein binding (GO:0005515)	1004	329	1.04	0.519	1.580
	microtubule motor activity (GO:0003777)	39	15	1.22	0.523	1.540

	DNA helicase activity (GO:0003678)	40	15	1.19	0.530	1.510
	signal transducer activity (GO:0004871)	168	58	1.1	0.530	1.530
	aminoacyl-tRNA ligase activity (GO:0004812)	42	16	1.21	0.537	1.500
	lipid transporter activity (GO:0005319)	23	9	1.24	0.540	1.490
	proton-transporting ATP synthase activity, rotational mechanism (GO:0046933)	11	5	1.44	0.557	1.510
	transmembrane receptor protein tyrosine kinase activity (GO:0004714)	32	12	1.19	0.597	1.490
	deacetylase activity (GO:0019213)	15	6	1.27	0.612	1.430
	hydrolase activity, hydrolyzing O-glycosyl compounds (GO:0004553)	38	14	1.17	0.626	1.440
	neurotransmitter transporter activity (GO:0005326)	17	7	1.31	0.631	1.420
	enzyme regulator activity (GO:0030234)	188	63	1.06	0.655	1.440
	peroxidase activity (GO:0004601)	24	9	1.19	0.683	1.400
	RNA helicase activity (GO:0003724)	24	9	1.19	0.683	1.380
	helicase activity (GO:0004386)	57	20	1.11	0.688	1.380
	extracellular matrix structural constituent (GO:0005201)	7	3	1.36	0.711	1.330
	non-membrane spanning protein tyrosine kinase activity (GO:0004715)	7	3	1.36	0.711	1.320
	growth factor activity (GO:0008083)	8	3	1.19	0.732	1.330
	DNA replication origin binding (GO:0003688)	8	3	1.19	0.732	1.320
	amino acid transmembrane transporter activity (GO:0015171)	38	13	1.09	0.745	1.330
	hydrolase activity, acting on ester bonds (GO:0016788)	266	87	1.04	0.752	1.330
	receptor activity (GO:0004872)	188	62	1.05	0.765	1.340
	carbohydrate transmembrane transporter activity (GO:0015144)	13	5	1.22	0.782	1.330
	G-protein coupled receptor activity (GO:0004930)	56	19	1.08	0.786	1.330
	DNA-directed DNA polymerase activity (GO:0003887)	14	5	1.13	0.790	1.320
	anion channel activity (GO:0005253)	16	6	1.19	0.802	1.320
	transaminase activity (GO:0008483)	17	6	1.12	0.808	1.300
	small GTPase regulator activity (GO:0005083)	76	25	1.04	0.816	1.290

	carbohydrate kinase activity (GO:0019200)	21	7	1.06	0.828	1.300
	antioxidant activity (GO:0016209)	36	12	1.06	0.866	1.310
	GTPase activity (GO:0003924)	173	56	1.03	0.876	1.310
	transporter activity (GO:0005215)	479	153	1.01	0.886	1.300
	ion channel activity (GO:0005216)	86	28	1.03	0.912	1.320
	glucosidase activity (GO:0015926)	15	5	1.06	1.000	1.390
	ligand-activated sequence-specific DNA binding RNA polymerase II transcription factor activity (GO:0004879)	15	5	1.06	1.000	1.380
	intramolecular transferase activity (GO:0016866)	12	4	1.06	1.000	1.340
	ubiquitin-protein ligase activity (GO:0004842)	47	15	1.01	1.000	1.370
	ligand-gated ion channel activity (GO:0015276)	44	14	1.01	1.000	1.400

9.6. Investigation of genes related to glutathione

Table 9-11 Table of all genes with OrthoDB gene families associated with glutathione that were differentially expressed with time (h) post ovulation of the interaction of embryo type and time (h) post-ovulation.

Differential expression model	Gene ID	*p*-adjusted	OrthoDB gene family	InterPro database ID
Time	Dapma7bEVm011307	2.43E-25	EOG091G0K2E	IPR003081:Glutathione S-transferase, Mu class
Time	Dapma7bEVm003943	3.76E-23	EOG091G0MBB	IPR004045:Glutathione S-transferase, N-terminal
Time	Dapma7bEVm004350	1.20E-08	EOG091G0FIA	IPR014186:S-formylglutathione hydrolase
Time	Dapma7bEVm008206	4.00E-08	EOG091G0MZM	IPR006840:Glutathione-specific gamma-glutamylcyclotransferase
Time	Dapma7bEVm010893	4.04E-08	EOG091G0MS3	IPR004045:Glutathione S-transferase, N-terminal
Time	Dapma7bEVm002915	1.41E-07	EOG091G0MBB	IPR004045:Glutathione S-transferase, N-terminal
Time	Dapma7bEVm018419	4.07E-07	EOG091G0FIA	IPR014186:S-formylglutathione hydrolase
Time	Dapma7bEVm028610	4.19E-05	EOG091G0K2E	IPR003081:Glutathione S-transferase, Mu class
Time	Dapma7bEVm005711	0.0001878	EOG091G0IEA	IPR005442:Glutathione S-transferase, omega-class
Time	Dapma7bEVm013092	0.000214145	EOG091G0OJG	IPR000889:Glutathione peroxidase
Time	Dapma7bEVm003922	0.000453993	EOG091G0QFQ	IPR005955:Glutathione S-transferases, class Zeta
Time	Dapma7bEVm002909	0.000820496	EOG091G0MBB	IPR004045:Glutathione S-transferase, N-terminal
Time	Dapma7bEVm011267	0.001255242	EOG091G0PCG	IPR004045:Glutathione S-transferase, N-terminal
Time	Dapma7bEVm004944	0.002356558	EOG091G0K2E	IPR003081:Glutathione S-transferase, Mu class
Time	Dapma7bEVm000589	0.005489411	EOG091G0MS3	IPR004045:Glutathione S-transferase, N-terminal
Time	Dapma7bEVm010649	0.016458929	EOG091G0MS3	IPR004045:Glutathione S-transferase, N-terminal
Time *x* Embryo type	Dapma7bEVm003922	4.24E-09	EOG091G0QFQ	IPR005955:Glutathione S-transferases, class Zeta
Time *x* Embryo type	Dapma7bEVm010893	4.77E-07	EOG091G0MS3	IPR004045:Glutathione S-transferase, N-terminal
Time *x* Embryo type	Dapma7bEVm000589	6.15E-06	EOG091G0MS3	IPR004045:Glutathione S-transferase, N-terminal
Time *x* Embryo type	Dapma7bEVm010990	6.76E-05	EOG091G0K2E	IPR003081:Glutathione S-transferase, Mu class

Time *x* Embryo type	Dapma7bEVm029274	0.001204435	EOG091G10LN	IPR000889:Glutathione peroxidase
Time *x* Embryo type	Dapma7bEVm005744	0.002692011	EOG091G0OJG	IPR000889:Glutathione peroxidase
Time *x* Embryo type	Dapma7bEVm013092	0.005078232	EOG091G0OJG	IPR000889:Glutathione peroxidase
Time *x* Embryo type	Dapma7bEVm000673	0.005809749	EOG091G03IU	IPR006338:Thioredoxin/glutathione reductase selenoprotein
Time *x* Embryo type	Dapma7bEVm002915	0.007132587	EOG091G0MBB	IPR004045:Glutathione S-transferase, N-terminal
Time *x* Embryo type	Dapma7bEVm002909	0.007364314	EOG091G0MBB	IPR004045:Glutathione S-transferase, N-terminal

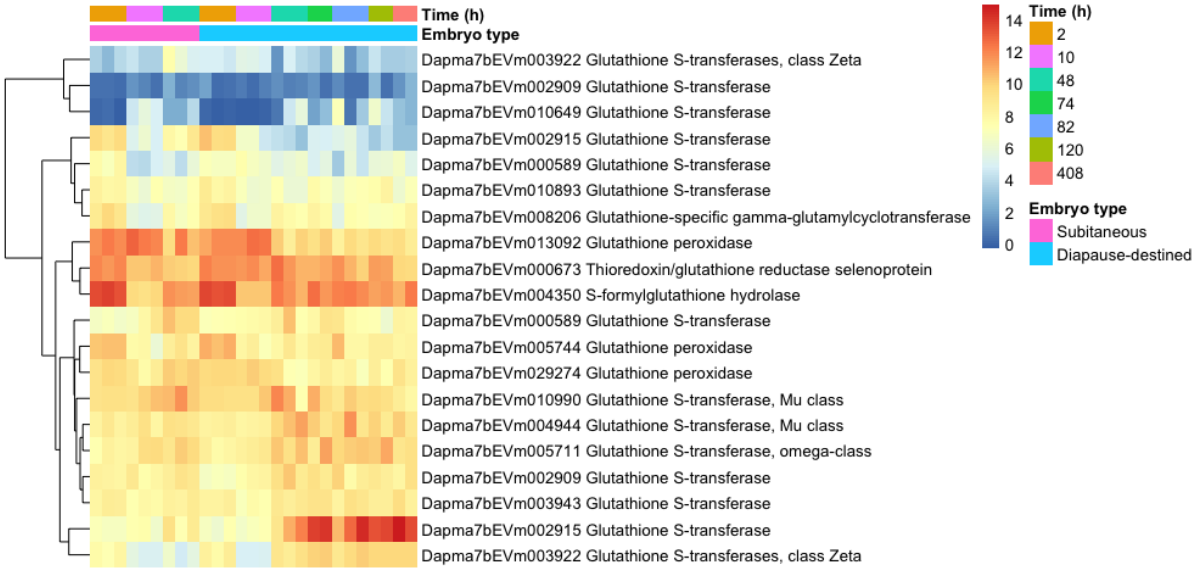


Figure 9-11 Heat map and hierarchical clustering of genes relating to glutathione that displayed differential expression (p-adjusted < 0.01) in samples of subitaneous and diapause-destined *Daphnia magna* embryos collected between 2 h and 48 h post-ovulation. Row labels indicate gene name with corresponding annotation based on OrthoDB orthologous gene family group and coloured column top panels indicate time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale**.**

10. Appendix B

10.1. Metabolite fold changes at 24 h versus 10 h and 48 h versus 24 h

Several additional time points were collected for the metabolomics analyses that were not present in the transcriptomic dataset. In order to maintain continuity for the findings described in Section 5, fold changes were calculated at 48 h versus 10 h. Here, the additional 24 h time point present only in the metabolomics datasets is considered. Metabolite fold changes between sexually produced embryos were calculated in the same manner described in Section 5.2 and are presented in Figure 10-1.

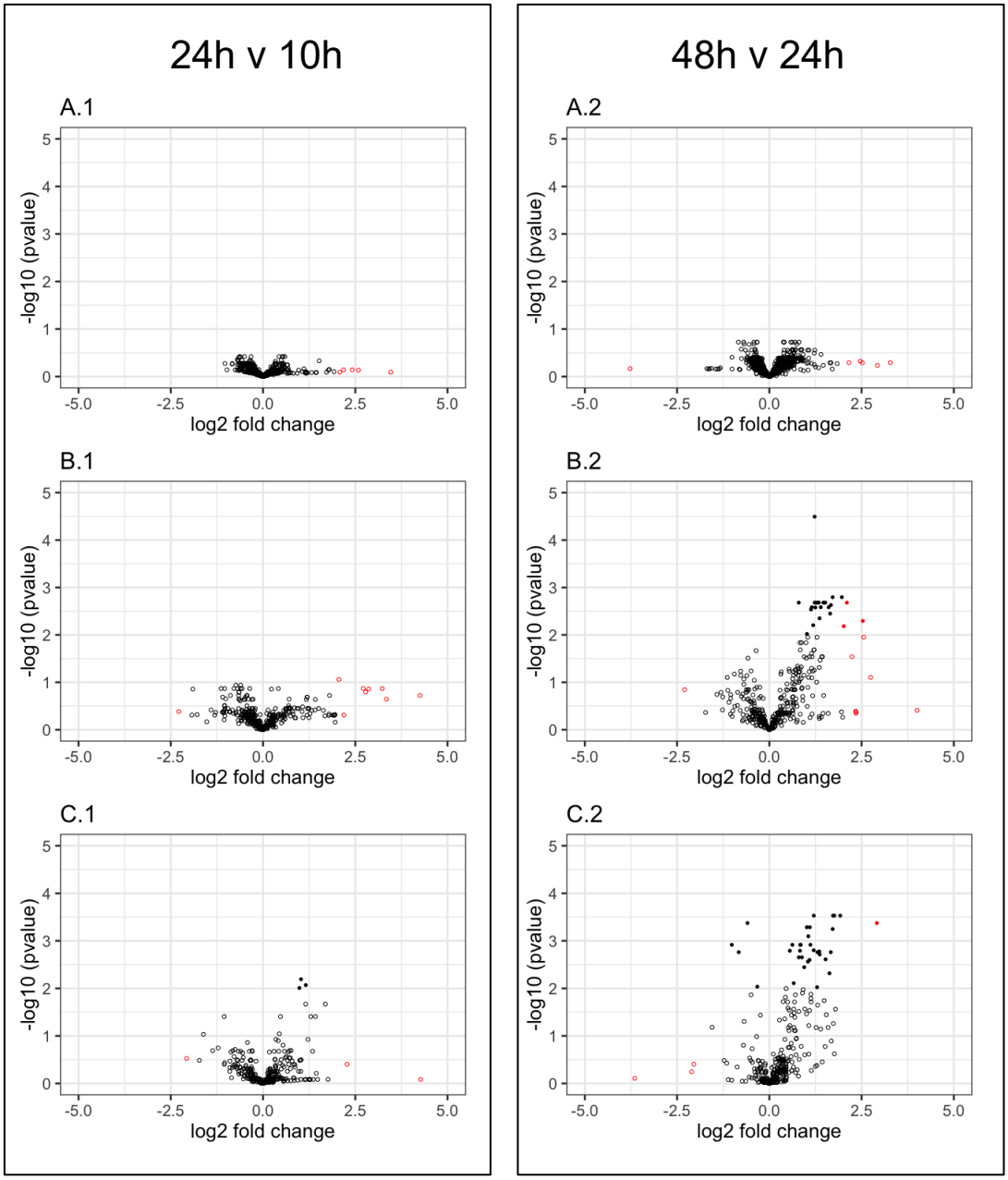


Figure 10-1 Volcano plots of metabolites with significant intensity fold changes between 24 h v10 h and 48 h v 24 h post-ovulation in relation to the onset of diapause (ca. 48 h) in embryos of *Daphnia magna.* Time points compared are indicated by surrounding boxes. Metabolite analysis implemented nESI-DIMS analysis of non-polar extracts (A; negative ion mode) and polar extracts analysed in negative ion mode (B) and positive ion mode (C). Filled points represent $p < 0.01$ (after Benjamini-Hochberg false discovery rate correction), red points indicate a $\log_2$ fold change ratio > 2.

In all three analyses only a small number of metabolites were observed to change significantly in intensity. There were no significantly changing peaks identified in either of the lipidomic comparisons, which is consistent with observations made when

just 48 h and 10 h embryos are compared. For both polar datasets, there were no peaks with a significant, large $\log_2$ intensity fold change at 24 h versus intensity levels at 10 h. Between 24 h and 48 h, three peaks showed large positive increases in the negative ion dataset and just a single peak in the positive ion dataset. Of the three peaks with significant fold changes in negative ion analysis, two were assigned to the putative glutathione peaks discussed in Chapter 4 and a third was not assigned a putative compound ID, but had 25 possible molecular formulae (Table 10-1). The only peak identified as significantly changing in positive ion mode was the putative trehalose peak (*m/z* = 381.079983) also identified and discussed in section 5.2.4.

Table 10-1 Table showing putative annotation of spectral peaks from postive and negative ion analysis of polar sexual *Daphnia magna* embryo extracts with signifcant ($p < 0.01$; after false discovery rate correction) large peak intensity fold change ratios (> 2) between 24 h and 48 h post-ovulation. Other putative formulae without compound identity assignment are indicated, for peaks with a high number of possible formulae only numbers are indicated for brevity.

Analysis	*m/z*	*p*-adjust	Fold change	Formula	Number of possible species	Ion form	Compound identity	Theoretical mass	Mass error
Negative ion	306.076599	0.005	2.53	C10H17N3O6S		[M-H]-	Glutathione	306.0765326	-0.30
				C18H13NO4	2	[M-H]-	Aristolodione	306.077183	1.82
							Hallacridone		
				4	–	–	–	–	–
	316.0597304	0.007	2.02	25	–	–	–	–	–
	328.058396	0.002	2.10	C10H17N3O6S	1	[M+Na-2H]-	Glutathione	328.0584776	0.23
				C18H13NO4	2	[M+Na-2H]-	Aristolodione	328.0591276	2.22
							Hallacridone		
				21	–	–	–	–	–
Positive ion	381.079983	0.000	2.91	C12H22O11	26	[M+K]+	3-b-Galactopyranosyl glucose	381.079374	-1.597305
							Allolactose		
							Alpha-Lactose		
							Cellobiose		
							D-Maltose		
							DEAE-cellulose		
							Epimelibiose		
							Fagopyritol A1		
							Galabiose		
							Galactinol		
							Glucinol		
							Glucose-1,3-mannose oligosaccharide		
							Inulobiose		
							Isomaltose		
							Kojibiose		

							Lactulose		
							Maltulose		
							Mannobiose		
							Melibiose		
							Neotrehalose		
							Sakebiose		
							Sucrose		
							Trehalose		
							Trehalulose		
							Turanose		
							beta-Lactose		
				39	–	–	–	–	–

None of the peaks are specific to the diapause only dataset and somewhat unexpectedly do not appear in the dataset reported in Section 5.2.2 comparing intensity values at 48 h versus 10 h. However, the very small number of peaks identified by this analysis is in line with the minimal changes in metabolism observed when embryos at 10 h and 48 h are compared and highlight the stark contrast with the much larger number of metabolite peaks with significant changes in intensity between 48 h and 74 h (Section 5.2.2).

10.2. Investigation of metabolite peaks that are significantly changing around the onset of diapause developmental arrest that are unique only to diapause trajectory dataset

To investigate the relevance of additional peaks present in the metabolomics dataset formed only of sexually produced embryos to the onset of diapause, significantly changing peaks were investigated to determine if any, how many peaks were diapause development specific. These peaks are indicated for negative ion analysis in Table 10-2 and for positive ion analysis in Table 10-3.

Table 10-2 Table showing spectral peaks from the negative ion analysis of polar extracts of sexually produced *Daphnia magna* embryos with signifcant ($p < 0.01$; after false discovery rate correction), large peak intensity log_2 fold change ratios (> 2) in relation to the onset of diapause (ca. 48 h post-ovulation). Peaks indicated in bold were unique only to diapause developmental samples when subitaneous embryos were excluded from analysis.

		m/z	*p*	*p*-adjust	fold change (log_2)
Negative polar	48 h v 10 h	**283.0821353**	8.06E-06	0.0003	2.20
		283.0882717	9.32E-06	0.0003	2.89
		284.0889469	1.07E-05	0.0003	2.35
		318.0610568	5.09E-05	0.0010	2.01
		353.0454938	0.0005	0.0051	3.04
		365.0910179	4.12E-07	3.64E-05	2.05
		380.0524944	8.21E-06	0.0003	2.21
		440.0736973	1.72E-05	0.0005	2.34
		446.0909565	3.02E-05	0.0007	2.34
	74 h v 48 h	**240.0806983**	1.68E-05	0.0005	2.96
		344.1164393	0.0002	0.0027	2.96
		404.1372965	0.0004	0.0038	3.04
		415.1458562	7.85E-07	3.96E-05	2.98
		423.0996209	0.0018	0.0096	2.83
		425.1159017	0.0005	0.0044	2.49
		456.1726082	0.0008	0.0056	2.90
		457.1929716	4.00E-05	0.0007	3.36
		471.0763489	0.0003	0.0033	2.76
		472.2041376	0.0014	0.0088	2.95
		473.5922387	1.01E-05	0.0003	3.21
		474.0939921	2.42E-05	0.0005	3.19
		477.0330939	0.0020	0.0097	2.55
		484.5824166	6.49E-07	3.82E-05	2.95
		499.2398356	0.0017	0.0096	2.57
		552.1245501	2.47E-05	0.0005	2.52
		559.0359971	5.55E-05	0.0009	2.70
		563.1126877	0.0012	0.0077	2.40
		567.1411589	3.97E-08	9.49E-06	3.47
		568.1433173	1.20E-07	1.06E-05	2.85
		569.0926547	7.06E-05	0.0010	2.40
		578.1314973	0.0004	0.0033	3.38
		579.0526412	3.91E-07	2.76E-05	2.16
		583.1148184	0.0003	0.0031	2.25
		587.0297446	0.0005	0.0040	-2.46
		587.1064332	8.06E-08	9.49E-06	2.82
		595.0844108	5.85E-08	9.49E-06	2.63
		598.3523645	0.0006	0.0045	2.32

Table 10-3 Table showing spectral peaks from the positive ion analysis of polar extracts of sexually produced *Daphnia magna* embryos with signifcant ($p < 0.01$; after false discovery rate correction), large peak intensity log_2 fold change ratios (> 2) in relation to the onset of diapause (ca. 48 h post-ovulation). Peaks indicated in bold were unique only to diapause developmental samples when subitaneous embryos were excluded from analysis.

		m/z	*p*	p-adjust	fold change (log_2)
Positive polar	48 h v 10 h	134.027165	1.07E-07	1.82E-05	2.52
		284.0994293	1.63E-07	1.85E-05	2.92
		306.0812489	5.44E-07	4.28E-05	2.89
		328.0633212	2.98E-06	0.0001	2.73
	74 h v 48 h	157.5378455	0.0004	0.0030	2.54
		182.0339773	0.0011	0.0056	4.11
		183.0504096	0.0007	0.0042	2.68
		191.5408531	0.0008	0.0047	3.90
		192.0556836	0.0002	0.0022	4.11
		200.5460564	0.0012	0.0058	4.37
		203.563695	0.0006	0.0037	3.84
		209.0498724	0.0011	0.0056	4.57
		211.5525371	0.0020	0.0081	4.33
		249.0461707	0.0003	0.0022	2.11
		258.1105988	5.15E-06	0.0001	2.83
		275.1515738	4.62E-06	0.0001	2.73
		356.1777409	3.36E-06	0.0001	4.55
		356.6794062	3.32E-06	0.0001	4.35
		360.1505037	8.91E-05	0.0013	4.08
		367.110097	3.57E-06	0.0001	3.90
		370.1933764	4.25E-06	0.0001	4.43
		370.6950165	2.77E-06	0.0001	4.26
		381.0799831	0.0007	0.0038	4.12
		382.083031	0.0005	0.0035	3.81
		384.1255341	0.0014	0.0066	3.56
		387.0876498	5.18E-06	0.0001	3.88
		404.1766512	0.0002	0.0022	4.34
		411.0985452	0.0005	0.0034	3.84
		423.0644118	6.50E-05	0.0010	3.49
		433.0805711	0.0002	0.0022	3.51
		434.0962698	1.05E-05	0.0002	3.75
		449.0666777	0.0016	0.0073	3.88
		450.0611668	3.01E-05	0.0006	3.62
		450.0700994	0.0006	0.0037	2.73
		450.0829515	6.13E-05	0.0010	3.68
		455.0622602	0.0003	0.0022	3.32
		471.0397442	9.41E-06	0.0002	3.65
		489.3261404	2.36E-06	0.0001	2.20
		495.0506106	0.0006	0.0037	3.23
		495.0631599	5.57E-05	0.0010	3.75
		569.0671265	0.0002	0.0018	3.50

10.3. Investigation of the role of trehalose in the diapausing embryos of *Daphnia magna*

Table 10-4 Mean intensity of putatively annotated trehalose peak (*m/z* 381.0799831) in *Daphnia magna* sexual embryos collected between 2 h and 2 y (17592 h) post-ovulation.

Time (h)	*n*	Mean intensity	Standard deviation	Fold change from previous time point	log_2 fold change from previous time point
2	3	138306.87	95714.32	–	–
10	5	43788.43	25242.98	0.32	-1.66
24	2	13237.76	3321.85	0.30	-1.73
48	8	99692.08	22627.90	7.53	2.91
74	7	1729879.86	669272.65	17.35	4.12
82	8	2676854.15	688158.96	1.55	0.63
98	8	3276735.99	923731.65	1.22	0.29
120	8	3373232.92	1176833.40	1.03	0.04
408	8	3511615.17	783616.47	1.04	0.06
17592	7	3874240.63	879061.25	1.10	0.14

Table 10-5 Table showing log_2 fold changes between 10 h, 48 h and 74 h post-ovulation, significance (*p*-values adjusted for false discovery rate), and relationship between mean log-normalised gene expression and trehalose metabolite intensity assessed with Spearman's rank correlation coefficient (ρ) of genes related to trehalose metabolism in the sexual embryos of *Daphnia magna.* NA represents genes that with outlying counts (in sample replicates) determined by Cook's distance (cut off = 0.99 quantile), (–) represents genes not detected in the dataset.

			Correlation with Trehalose intensity		10h v 48h		48h v 74h	
Gene ID	OrthoDB orthologous group	Putative identity	ρ	*p*	Fold change	*p*-adjust	Fold change	*p*-adjust
Dapma7bEVm004057	EOG091G00R4	Trehalose-phosphatase	–	–	–	–	–	–
Dapma7bEVm005872	EOG091G0341	Trehalase	-0.036	0.963	-0.288	0.784	-0.353	1.000
Dapma7bEVm005873	EOG091G0341	Trehalase	0.821	0.034	-0.086	0.932	0.439	1.000
Dapma7bEVm009716	EOG091G00R4	Trehalose-phosphatase	0.107	0.840	4.381	5.44E-24	-0.999	0.488
Dapma7bEVm015481	EOG091G00R4	Trehalose-phosphatase	0.714	0.088	2.347	NA	0.774	NA
Dapma7bEVm029000	EOG091G00R4	Trehalose-phosphatase	0.714	0.088	0.518	0.765	0.810	1.000

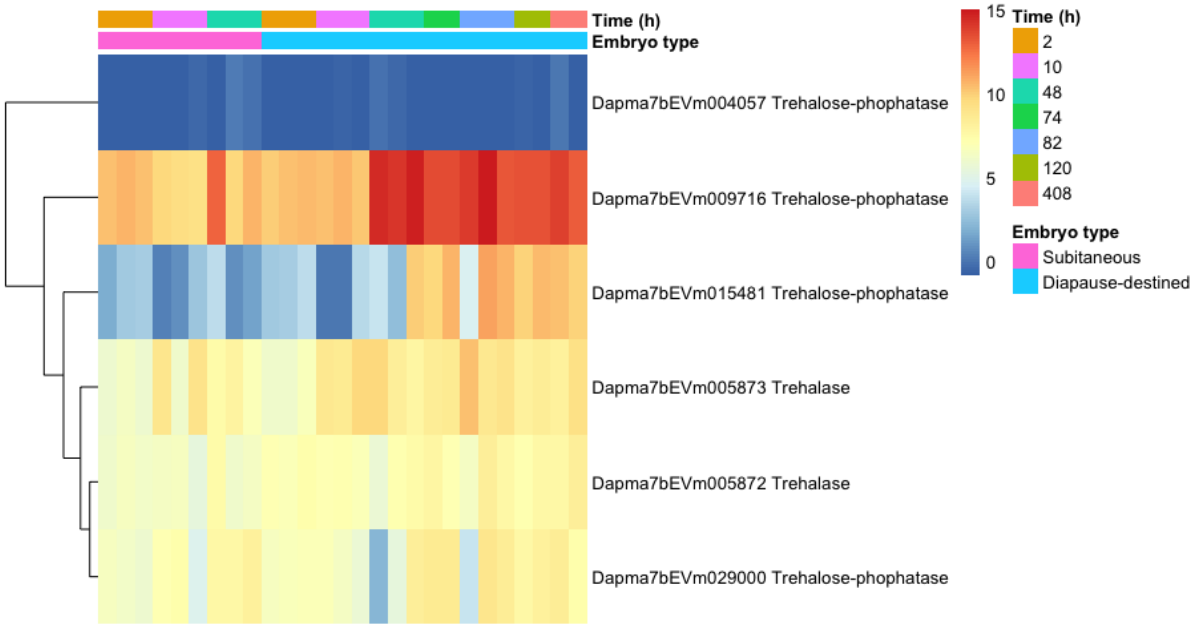


Figure 10-2 Heat map and hierarchical clustering of genes relating to trehalose in samples of subitaneous and diapause-destined *Daphnia* embryos collected between 2 h and 2 weeks (408 h) post-ovulation. Row labels indicate gene name with corresponding annotation based on OrthoDB orthologous gene family group and coloured column top panels indicate time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale**.**

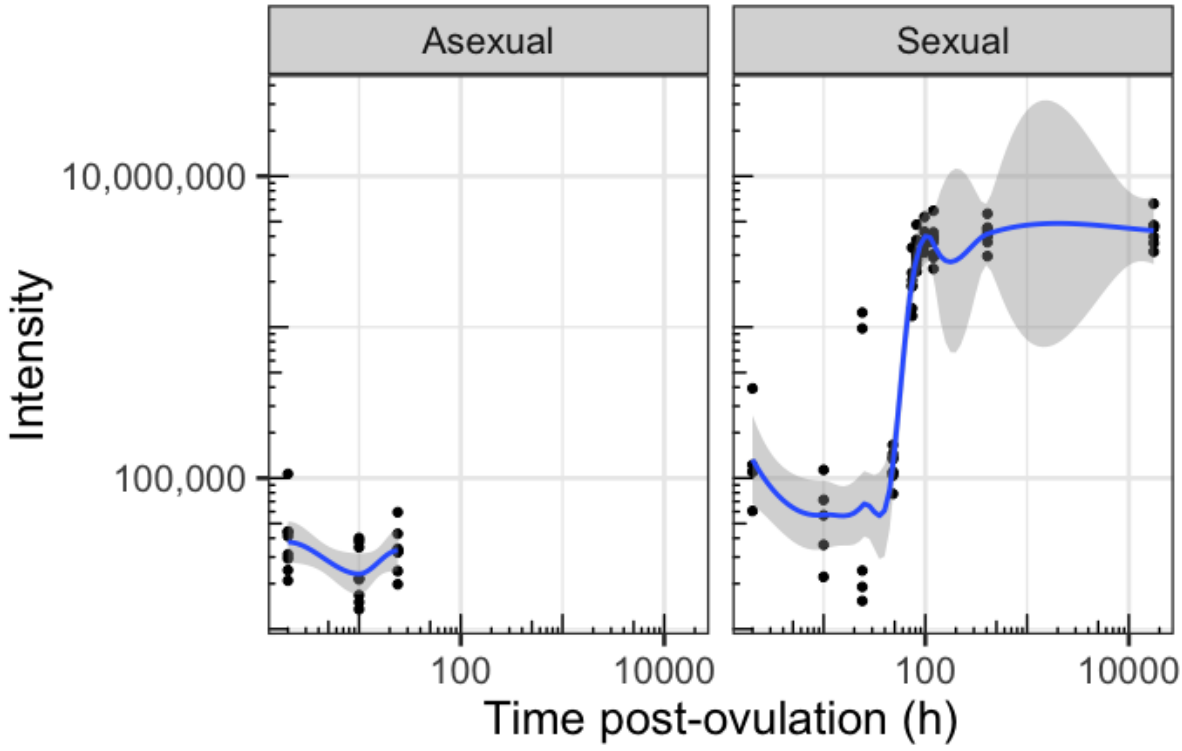


Figure 10-3 Intensity of putatively annotated trehalose peak identified utilising nESI-DIMS analysis of polar (positive ion analysis) extracts of *Daphnia magna* subitaneous (Asexual) and diapause-destined (Sexual) embryos. Points indicate individual biological replicates, lines fitted with locally weighted scatterplot smoothing (LOESS; span = 0.5), grey shading indicate model 95% confidence interval.

10.4. Gene ontology annotation and overrepresentation analysis of genes differentially expressed in relation to the initiation of diapause in sexually produced embryos of *Daphnia magna*

Table 10-6 Table of gene ontology terms significantly overrepresented in genes differentially expressed between sexually produced *Daphnia magna* embryos collected 48 h and 10 h post-ovulation after false discovery rate correction ($p < 0.1$). Background contained 19,244 genes, and original input contained 5,127 genes (2681 up-regulated, 2,446 down-regulated; direction refers to gene expression at 48 h relative to 10 h).

	Domain	PANTHER GO-Slim term	Background	Input	Fold enrichment	raw *p*-value	*p*-adjust
Down-regulated	Biological process	DNA metabolic process	254	95	2.25	1.33E-10	1.56E-08
		nucleobase-containing compound metabolic process	1602	376	1.41	9.54E-11	2.24E-08
		cell cycle	361	117	1.95	1.14E-09	8.94E-08
		DNA replication	97	43	2.67	5.15E-07	3.02E-05
		nitrogen compound metabolic process	1508	326	1.3	3.39E-06	1.59E-04
		RNA metabolic process	891	208	1.41	6.40E-06	2.51E-04
		mitosis	131	47	2.16	1.80E-05	0.001
		chromatin organization	160	52	1.96	5.92E-05	0.002
		reproduction	74	30	2.44	9.92E-05	0.002
		DNA repair	128	44	2.07	9.37E-05	0.002
		chromosome segregation	60	26	2.61	1.28E-04	0.003
		meiosis	36	19	3.18	1.23E-04	0.003
		regulation of nucleobase-containing compound metabolic process	230	62	1.62	0.001	0.014
		primary metabolic process	2755	523	1.14	0.001	0.015
		metabolic process	3307	615	1.12	0.001	0.019
		organelle organization	690	151	1.32	0.002	0.021
		regulation of cell cycle	93	31	2.01	0.002	0.022
		tRNA metabolic process	100	31	1.87	0.004	0.039
	Cellular component	nucleus	1092	269	1.48	2.20E-09	7.03E-08

		chromosome	148	64	2.61	1.72E-09	1.10E-07
		nuclear chromosome	96	45	2.83	6.90E-08	1.10E-06
		nucleoplasm	238	69	1.75	8.71E-05	0.001
		protein-DNA complex	29	16	3.33	2.85E-04	0.002
		nuclear envelope	57	20	2.11	0.008	0.038
		microtubule organizing center	62	21	2.04	0.007	0.039
	Molecular function	helicase activity	57	24	2.54	3.15E-04	0.014
		DNA helicase activity	40	20	3.01	2.42E-04	0.015
		transferase activity	742	169	1.37	1.65E-04	0.015
		DNA-directed RNA polymerase activity	35	17	2.93	0.001	0.028
		nucleotidyltransferase activity	66	25	2.28	0.001	0.029
		chromatin binding	104	33	1.91	0.002	0.049
Up-regulated	Biological process	lipid metabolic process	205	62	1.66	0.001	0.036
	Molecular function	oxidoreductase activity	360	105	1.6	3.24E-05	0.003

Table 10-7 Table of gene ontology terms overrepresented in genes differentially expressed between sexually produced *Daphnia magna* embryos collected 74 h and 48 h post-ovulation after false discovery rate correction ($p < 0.1$). Background contained 19,244 genes, and original input contained 229 genes (135 up-regulated, 94 down-regulated; direction refers to gene expression at 74 h relative to 48 h).

	Domain	PANTHER GO-Slim term	Background	Input	Fold enrichment	raw *p*-value	*p*-adjust
Down-regulated	Biological process	protein glycosylation (GO:0006486)	60	4	9.03	0.001	0.284
		carbohydrate metabolic process (GO:0005975)	215	6	3.78	0.005	0.644
		coenzyme metabolic process (GO:0006732)	74	1	1.83	0.425	2.560
		steroid metabolic process (GO:0008202)	43	1	3.15	0.277	2.710
		synaptic transmission (GO:0007268)	123	1	1.1	0.601	2.710
		polysaccharide metabolic process (GO:0005976)	57	2	4.75	0.069	2.720
		fatty acid metabolic process (GO:0006631)	83	1	1.63	0.462	2.720
		acyl-CoA metabolic process (GO:0006637)	14	1	9.68	0.105	2.730
		RNA splicing, via transesterification reactions (GO:0000375)	106	1	1.28	0.547	2.730
		mRNA splicing, via spliceosome (GO:0000398)	121	1	1.12	0.595	2.740
		protein targeting (GO:0006605)	118	1	1.15	0.585	2.750
		biosynthetic process (GO:0009058)	1065	9	1.14	0.703	2.750
		regulation of transcription from RNA polymerase II promoter (GO:0006357)	256	2	1.06	0.714	2.750
		protein localization (GO:0008104)	314	3	1.29	0.505	2.760
		immune system process (GO:0002376)	106	1	1.28	0.547	2.790
		multicellular organismal process (GO:0032501)	358	3	1.14	0.748	2.790
		respiratory electron transport chain (GO:0022904)	93	1	1.46	0.501	2.800
		tRNA metabolic process (GO:0006399)	100	1	1.35	0.526	2.810
		protein transport (GO:0015031)	397	4	1.37	0.540	2.820
		cation transport (GO:0006812)	43	1	3.15	0.277	2.830
		single-multicellular organism process (GO:0044707)	355	3	1.14	0.747	2.830
		lipid metabolic process (GO:0006629)	205	3	1.98	0.195	2.870

		system process (GO:0003008)	226	3	1.8	0.236	2.910
		mesoderm development (GO:0007498)	65	1	2.08	0.386	2.920
		cellular protein modification process (GO:0006464)	470	5	1.44	0.398	2.930
		cellular amino acid metabolic process (GO:0006520)	177	2	1.53	0.379	2.970
		intracellular protein transport (GO:0006886)	375	4	1.45	0.360	3.020
		fatty acid biosynthetic process (GO:0006633)	22	1	6.16	0.156	3.050
		Unclassified (UNCLASSIFIED)	4687	40	1.16	0.233	3.050
		neuron-neuron synaptic transmission (GO:0007270)	33	1	4.11	0.222	3.060
		glycogen metabolic process (GO:0005977)	25	1	5.42	0.174	3.150
	Cellular component	extracellular region (GO:0005576)	179	4	3.03	0.045	0.576
		extracellular space (GO:0005615)	135	3	3.01	0.080	0.735
		Unclassified (UNCLASSIFIED)	5982	51	1.16	0.110	0.782
		Golgi apparatus (GO:0005794)	165	3	2.46	0.125	0.801
		endoplasmic reticulum (GO:0005783)	229	3	1.77	0.241	1.290
		cytoplasmic membrane-bounded vesicle (GO:0016023)	88	1	1.54	0.482	1.710
		plasma membrane (GO:0005886)	465	4	1.17	0.777	2.160
	Molecular function	transferase activity, transferring glycosyl groups (GO:0016757)	93	3	4.37	0.033	1.510
		transmembrane receptor protein serine/threonine kinase activity (GO:0004675)	6	1	22.58	0.050	1.520
		phosphatase regulator activity (GO:0019208)	9	1	15.05	0.071	1.840
		antioxidant activity (GO:0016209)	36	2	7.53	0.031	1.910
		transporter activity (GO:0005215)	479	6	1.7	0.169	2.200
		neurotransmitter transporter activity (GO:0005326)	17	1	7.97	0.124	2.260
		carboxy-lyase activity (GO:0016831)	19	1	7.13	0.137	2.270
		peroxidase activity (GO:0004601)	24	1	5.64	0.168	2.350

		transmembrane transporter activity (GO:0022857)	389	5	1.74	0.214	2.430
		lipid transporter activity (GO:0005319)	23	1	5.89	0.162	2.460
		transaminase activity (GO:0008483)	17	1	7.97	0.124	2.510
		receptor activity (GO:0004872)	188	2	1.44	0.651	2.760
		lyase activity (GO:0016829)	131	1	1.03	0.624	3.070
		cytoskeletal protein binding (GO:0008092)	123	1	1.1	0.601	3.120
		protein binding transcription factor activity (GO:0000988)	129	1	1.05	0.618	3.120
		calcium ion binding (GO:0005509)	69	1	1.96	0.403	3.190
		transcription factor binding transcription factor activity (GO:0000989)	122	1	1.11	0.598	3.200
		isomerase activity (GO:0016853)	62	1	2.19	0.372	3.220
		Unclassified (UNCLASSIFIED)	5192	41	1.07	0.552	3.240
		ATPase activity, coupled to transmembrane movement of substances (GO:0042626)	67	1	2.02	0.395	3.260
		transferase activity, transferring acyl groups (GO:0016746)	114	1	1.19	0.573	3.260
		transcription cofactor activity (GO:0003712)	105	1	1.29	0.543	3.300
		enzyme activator activity (GO:0008047)	61	1	2.22	0.367	3.340
		cation transmembrane transporter activity (GO:0008324)	154	2	1.76	0.317	3.400
		signal transducer activity (GO:0004871)	168	2	1.61	0.355	3.400
		phosphatase activity (GO:0016791)	96	1	1.41	0.512	3.450
		oxidoreductase activity (GO:0016491)	360	4	1.51	0.342	3.460
		sequence-specific DNA binding RNA polymerase II transcription factor activity (GO:0000981)	93	1	1.46	0.501	3.510
		transferase activity (GO:0016740)	742	7	1.28	0.499	3.630
		amino acid kinase activity (GO:0019202)	3	1	45.16	0.029	5.280
Up-regul	Biological proc ess	lipid metabolic process (GO:0006629)	205	2	3.35	0.121	2.830

		negative regulation of apoptotic process (GO:0043066)	42	1	8.18	0.118	3.070
		protein phosphorylation (GO:0006468)	42	1	8.18	0.118	3.460
		cyclic nucleotide metabolic process (GO:0009187)	40	1	8.59	0.113	3.780
		response to stimulus (GO:0050896)	768	4	1.79	0.278	3.850
		DNA repair (GO:0006281)	128	1	2.68	0.314	3.890
		intracellular signal transduction (GO:0035556)	336	1	1.02	1.000	3.980
		regulation of catalytic activity (GO:0050790)	108	1	3.18	0.273	4.010
		cell differentiation (GO:0030154)	145	1	2.37	0.348	4.090
		regulation of molecular function (GO:0065009)	128	1	2.68	0.314	4.100
		homeostatic process (GO:0042592)	106	1	3.24	0.269	4.210
		response to abiotic stimulus (GO:0009628)	39	1	8.81	0.110	4.310
		phosphate-containing compound metabolic process (GO:0006796)	830	3	1.24	0.730	4.400
		protein metabolic process (GO:0019538)	968	3	1.06	0.757	4.450
		cellular protein modification process (GO:0006464)	470	2	1.46	0.646	4.470
		proteolysis (GO:0006508)	257	1	1.34	0.532	4.470
		regulation of phosphate metabolic process (GO:0019220)	200	1	1.72	0.446	4.560
		MAPK cascade (GO:0000165)	91	1	3.78	0.236	4.610
		regulation of transcription from RNA polymerase II promoter (GO:0006357)	256	1	1.34	0.531	4.620
		Unclassified (UNCLASSIFIED)	4687	17	1.25	0.256	4.630
		fatty acid metabolic process (GO:0006631)	83	1	4.14	0.217	4.650
		DNA metabolic process (GO:0006259)	254	1	1.35	0.528	4.960
		ion transport (GO:0006811)	178	2	3.86	0.095	5.610
		response to stress (GO:0006950)	310	3	3.32	0.061	7.160
		anion transport (GO:0006820)	109	2	6.3	0.041	9.600
	Cellular component	Unclassified (UNCLASSIFIED)	5982	18	1.03	1.000	1.160
		extracellular space (GO:0005615)	135	2	5.09	0.060	1.270

		plasma membrane (GO:0005886)	465	4	2.96	0.045	1.430
		extracellular region (GO:0005576)	179	2	3.84	0.096	1.540
		integral to membrane (GO:0016021)	387	4	3.55	0.025	1.610
		membrane (GO:0016020)	905	4	1.52	0.330	3.020
		mitochondrion (GO:0005739)	255	1	1.35	0.530	3.080
	Molecular function	metallopeptidase activity (GO:0008237)	39	1	8.81	0.110	2.220
		adenylate cyclase activity (GO:0004016)	37	1	9.28	0.105	2.380
		antioxidant activity (GO:0016209)	36	1	9.54	0.102	2.660
		kinase regulator activity (GO:0019207)	35	1	9.82	0.100	3.020
		transmembrane receptor protein tyrosine kinase activity (GO:0004714)	32	1	10.74	0.092	3.330
		carboxy-lyase activity (GO:0016831)	19	1	18.08	0.057	3.430
		lipase activity (GO:0016298)	31	1	11.08	0.089	4.050
		protein binding transcription factor activity (GO:0000988)	129	1	2.66	0.316	4.110
		receptor binding (GO:0005102)	147	1	2.34	0.352	4.270
		receptor activity (GO:0004872)	188	1	1.83	0.426	4.310
		sequence-specific DNA binding RNA polymerase II transcription factor activity (GO:0000981)	93	1	3.69	0.240	4.370
		transcription cofactor activity (GO:0003712)	105	1	3.27	0.266	4.410
		peptidase activity (GO:0008233)	180	1	1.91	0.412	4.410
		transmembrane transporter activity (GO:0022857)	389	2	1.77	0.315	4.410
		signal transducer activity (GO:0004871)	168	1	2.04	0.391	4.450
		transcription factor binding transcription factor activity (GO:0000989)	122	1	2.82	0.302	4.580
		hydrolase activity, acting on ester bonds (GO:0016788)	266	1	1.29	0.545	4.720
		sequence-specific DNA binding transcription factor activity (GO:0003700)	307	1	1.12	0.597	4.730
		kinase activity (GO:0016301)	289	1	1.19	0.575	4.760

		protein binding (GO:0005515)	1004	4	1.37	0.529	4.810
		catalytic activity (GO:0003824)	2434	8	1.13	0.667	4.850
		transporter activity (GO:0005215)	479	2	1.43	0.649	4.920
		lyase activity (GO:0016829)	131	2	5.24	0.057	5.140
		hydrolase activity (GO:0016787)	993	3	1.04	0.763	5.140
		transaminase activity (GO:0008483)	17	1	20.21	0.051	9.290

10.5. Investigation of the mRNA expression of putatively annotated genes within the 'response to stress' gene ontology term

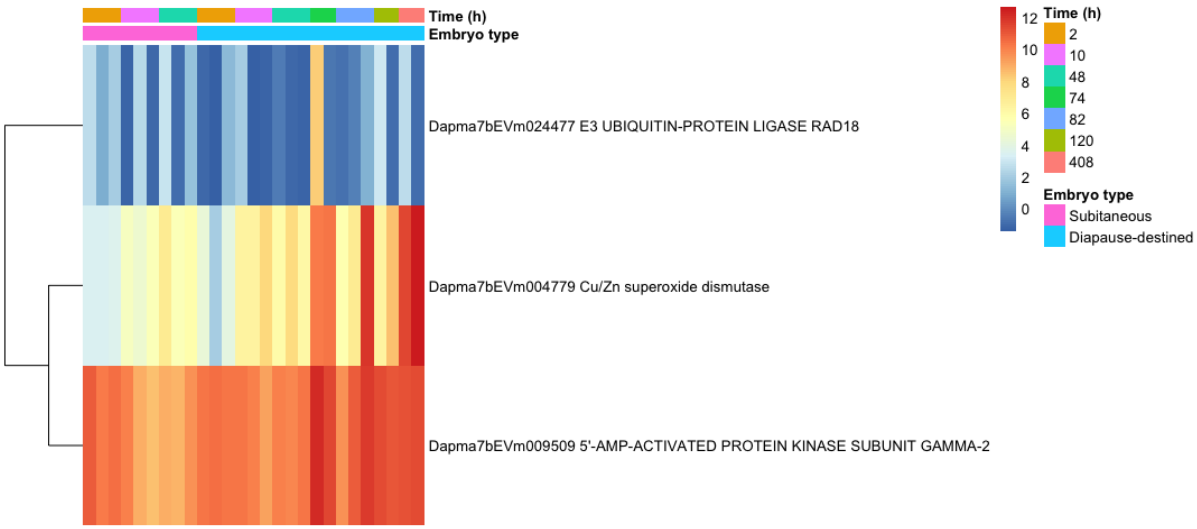


Figure 10-4 Heat map and hierarchical clustering of putatively annotated genes within the 'response to stress' gene ontology term in samples of subitaneous and diapause-destined *Daphnia magna* embryos collected between 2 h and 2 weeks (408 h) post-ovulation. Row labels indicate gene name with corresponding annotation based on OrthoDB orthologous gene family group and coloured column top panels indicate time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale.

10.6. Investigation of putatively annotated superoxide dismutase genes in the sexually produced embryos of *Daphnia magna.*

Table 10-8 Table showing log_2 fold changes between 10 h, 48 h and 74 h post-ovulation, significance (*p*-values adjusted for false discovery rate) of putatively annotated superoxide dismutase genes in the sexually produced embryos of *Daphnia magna.* NA represents genes with outlying counts (in sample replicates) determined by Cook's distance.

			10 h v 48 h		48 h v 74 h	
Gene ID	OrthoDB orthologous group	Identity	Fold change	*p*-adjust	Fold change	*p*-adjust
Dapma7bEVm002040	EOG091G0OG2	Superoxide dismutase, copper/zinc	-1.527	0.335	-0.281	NA
Dapma7bEVm002574	EOG091G014U	Superoxide dismutase, copper/zinc	0.522	0.481	0.071	1.000
Dapma7bEVm003936	EOG091G0MVL	Manganese/iron superoxide dismutase	6.036	3.72E-16	0.826	1.000
Dapma7bEVm004455	EOG091G0MVL	Manganese/iron superoxide dismutase	1.091	0.169	0.742	1.000
Dapma7bEVm004779	EOG091G09XI	Superoxide dismutase, copper/zinc	-0.265	0.821	2.827	0.002
Dapma7bEVm006441	EOG091G0OG2	Superoxide dismutase, copper/zinc	0.802	0.312	-0.303	1.000
Dapma7bEVm008098	EOG091G0OG2	Superoxide dismutase, copper/zinc	6.469	2.85E-33	1.324	0.585
Dapma7bEVm008201	EOG091G0FQ1	Superoxide dismutase, copper/zinc	1.055	0.373	0.044	1.000
Dapma7bEVm028228	EOG091G0OG2	Superoxide dismutase, copper/zinc	6.504	1.25E-24	1.057	0.844

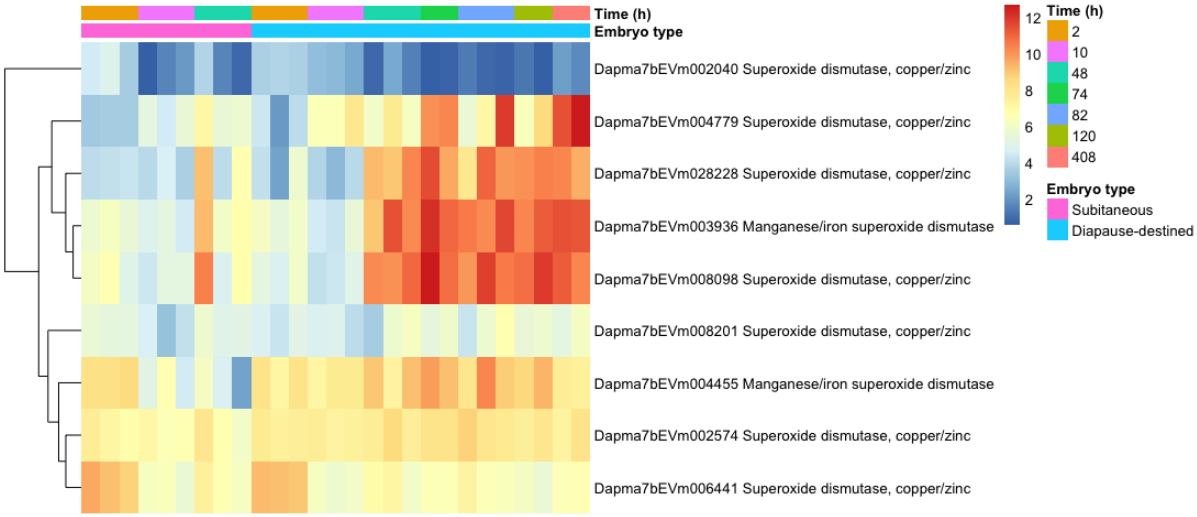


Figure 10-5 Heat map and hierarchical clustering of expression of putatively annotated superoxide dismutase genes in samples of subitaneous and diapause-destined *Daphnia magna* embryos collected between 2 h and 2 weeks (408 h) post-ovulation. Row labels indicate gene name with corresponding annotation based on OrthoDB orthologous gene family group and coloured column top panels indicate time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale**.**

10.7. Investigation of genes annotated with 'oxidoreductase activity' and/or 'lipid metabolic process' gene ontology

Table 10-9 Table of *Daphnia magna* genes with significant (p-adjust < 0.1) increased expression at 48 h versus 10 h post-ovulation that were assigned putative identities and membership to 'lipid metabolic process' gene ontology during Protein ANalysis THrough Evolutionary Relationships (PANTHER) overrepresentation analysis.

Gene ID	Putative identity
Dapma7bEVm000016	CARNITINE PALMITOYLTRANSFERASE 1AB (LIVER)
Dapma7bEVm000246	BETA-HEXOSAMINIDASE SUBUNIT ALPHA
Dapma7bEVm000284	MICROSOMAL TRIACYLGLYCEROL TRANSFER PROTEIN
Dapma7bEVm000400	CARNITINE O-ACETYL-TRANSFERASE, ISOFORM A
Dapma7bEVm000508	ACYL-COA SYNTHETASE LONG-CHAIN, ISOFORM J
Dapma7bEVm000551	CYTOCHROME P450 4V2
Dapma7bEVm000680	VERY LONG-CHAIN SPECIFIC ACYL-COA DEHYDROGENASE, MITOCHONDRIAL
Dapma7bEVm000715	CYTOCHROME P450 4AE1-RELATED
Dapma7bEVm001330	SORTILIN-RELATED RECEPTOR
Dapma7bEVm001432	PROPIONYL-COA CARBOXYLASE ALPHA CHAIN, MITOCHONDRIAL
Dapma7bEVm001491	SUBFAMILY NOT NAMED
Dapma7bEVm001517	ACETYL-COA ACETYLTRANSFERASE, MITOCHONDRIAL
Dapma7bEVm001551	LONG-CHAIN SPECIFIC ACYL-COA DEHYDROGENASE, MITOCHONDRIAL
Dapma7bEVm001798	DIHYDROXYACETONE PHOSPHATE ACYLTRANSFERASE
Dapma7bEVm001880	SUBFAMILY NOT NAMED
Dapma7bEVm002241	ACYL-COA DEHYDROGENASE FAMILY MEMBER 11
Dapma7bEVm002265	BETA-HEXOSAMINIDASE SUBUNIT ALPHA
Dapma7bEVm002369	MONOACYLGLYCEROL LIPASE ABHD2
Dapma7bEVm002472	FAMILY NOT NAMED
Dapma7bEVm003085	TRIFUNCTIONAL ENZYME SUBUNIT ALPHA, MITOCHONDRIAL
Dapma7bEVm003191	SUBFAMILY NOT NAMED
Dapma7bEVm003216	SUBFAMILY NOT NAMED
Dapma7bEVm003238	BRUMMER, ISOFORM B-RELATED
Dapma7bEVm003301	LIPOYL SYNTHASE, MITOCHONDRIAL
Dapma7bEVm003385	SUBFAMILY NOT NAMED
Dapma7bEVm003395	ENOYL-COA DELTA ISOMERASE 1, MITOCHONDRIAL
Dapma7bEVm003505	FAMILY NOT NAMED
Dapma7bEVm003876	PEROXISOMAL FARNESYLATED PROTEIN
Dapma7bEVm004064	PEROXISOMAL ACYL-COENZYME A OXIDASE 3
Dapma7bEVm004083	ACYL-COA SYNTHETASE LONG-CHAIN, ISOFORM J
Dapma7bEVm004168	FAMILY NOT NAMED
Dapma7bEVm004284	ELONGATION OF VERY LONG CHAIN FATTY ACIDS PROTEIN
Dapma7bEVm004563	BETA-HEXOSAMINIDASE A
Dapma7bEVm004663	SUBFAMILY NOT NAMED
Dapma7bEVm004826	ENOYL-[ACYL-CARRIER-PROTEIN] REDUCTASE, MITOCHONDRIAL
Dapma7bEVm005035	3-HYDROXY-3-METHYLGLUTARYL-COENZYME A REDUCTASE
Dapma7bEVm005231	CERAMIDE GLUCOSYLTRANSFERASE
Dapma7bEVm005658	LAMBDA-CRYSTALLIN HOMOLOG

Dapma7bEVm005865	PHOSPHOENOLPYRUVATE CARBOXYKINASE, CYTOSOLIC [GTP]
Dapma7bEVm006375	3-OXO-5-ALPHA-STEROID 4-DEHYDROGENASE 1
Dapma7bEVm006854	SC2
Dapma7bEVm006877	SI:DKEY-12E7.4
Dapma7bEVm007221	CHOLINE O-ACETYLTRANSFERASE
Dapma7bEVm008018	LD18904P
Dapma7bEVm008370	ORM1-LIKE PROTEIN 3
Dapma7bEVm008565	DELTA(7)-STEROL 5(6)-DESATURASE
Dapma7bEVm008584	SUBFAMILY NOT NAMED
Dapma7bEVm009049	GLUTATHIONE PEROXIDASE-RELATED
Dapma7bEVm009473	DIACYLGLYCEROL KINASE ZETA
Dapma7bEVm009616	BCDNA.GH02901
Dapma7bEVm009688	SUBFAMILY NOT NAMED
Dapma7bEVm010041	SN1-SPECIFIC DIACYLGLYCEROL LIPASE BETA
Dapma7bEVm010112	SUBFAMILY NOT NAMED
Dapma7bEVm010137	COUP TRANSCRIPTION FACTOR 2
Dapma7bEVm010248	BCDNA.GH02901
Dapma7bEVm011365	3-OXO-5-ALPHA-STEROID 4-DEHYDROGENASE 1
Dapma7bEVm011832	PLECKSTRIN HOMOLOGY DOMAIN-CONTAINING FAMILY A MEMBER 3
Dapma7bEVm013092	SUBFAMILY NOT NAMED
Dapma7bEVm014983	SUBFAMILY NOT NAMED
Dapma7bEVm014984	SUBFAMILY NOT NAMED
Dapma7bEVm015143	FATTY ACID (LONG CHAIN) TRANSPORT PROTEIN, ISOFORM F-RELATED
Dapma7bEVm015378	SUBFAMILY NOT NAMED
Dapma7bEVm015478	ACYL-COA SYNTHETASE SHORT-CHAIN FAMILY MEMBER 3, MITOCHONDRIAL
Dapma7bEVm015590	PROTEIN NPC2 HOMOLOG
Dapma7bEVm015741	CYTOCHROME B
Dapma7bEVm016721	SUBFAMILY NOT NAMED
Dapma7bEVm019204	SUBFAMILY NOT NAMED
Dapma7bEVm023103	POLYPRENOL REDUCTASE
Dapma7bEVm023119	SUBFAMILY NOT NAMED
Dapma7bEVm023277	TA01441P
Dapma7bEVm023366	SUBFAMILY NOT NAMED
Dapma7bEVm026608	ZGC:113142
Dapma7bEVm027047	ACYL-COENZYME A BINDING PROTEIN
Dapma7bEVm027888	LAMIN-B RECEPTOR
Dapma7bEVm028406	2-HYDROXYACYL-COA LYASE 1
Dapma7bEVm029009	SUBFAMILY NOT NAMED
Dapma7bEVm029274	GLUTATHIONE PEROXIDASE-RELATED

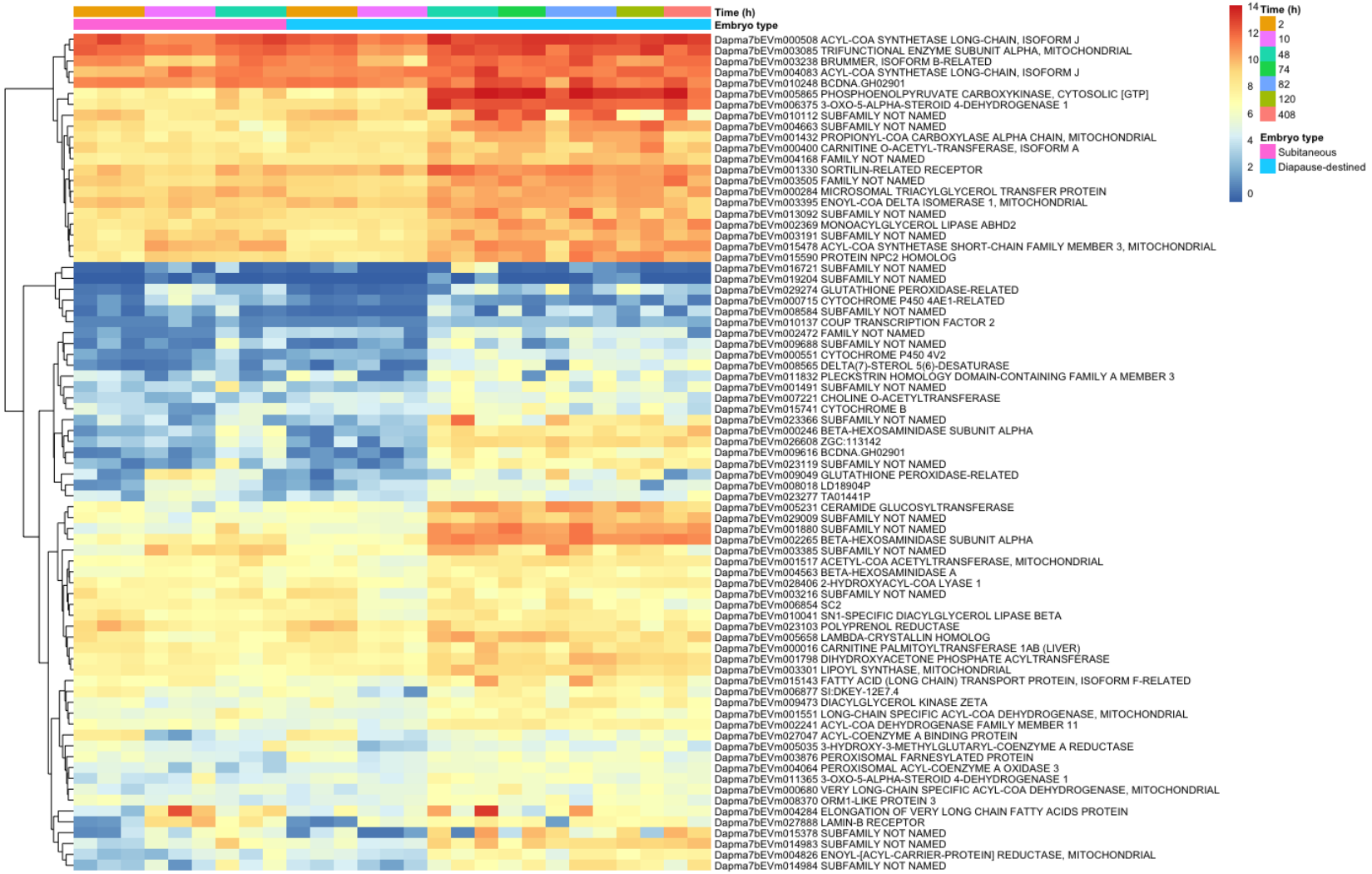


Figure 10-6 Heat map and hierarchical clustering of expression of genes assigned to the 'lipid metabolic process' gene ontology term that were significantly up-regulated in expression in samples of sexual *Daphnia magna* embryos collected at 48 h versus 10 h post-ovulation. Column labels individual biological replicates (see key for embryo type and collection time). Colours indicate low expression (blue) and high expression (red) on log-normalised scale

Table 10-10 Table of *Daphnia magna* genes with significant (p-adjusted < 0.1) increased expression at 48 h versus 10 h post-ovulation that were assigned putative identities and membership to 'oxidoreductase activity' gene ontology during Protein ANalysis THrough Evolutionary Relationships (PANTHER) overrepresentation analysis.

Gene ID	Putative identity
Dapma7bEVm000314	ALKYLATED DNA REPAIR PROTEIN ALKB HOMOLOG 8
Dapma7bEVm000333	PEPTIDYL-ALPHA-HYDROXYGLYCINE ALPHA-AMIDATING LYASE 1-RELATED
Dapma7bEVm000389	NADH-UBIQUINONE OXIDOREDUCTASE 75 KDA SUBUNIT, MITOCHONDRIAL
Dapma7bEVm000551	CYTOCHROME P450 4V2
Dapma7bEVm000680	VERY LONG-CHAIN SPECIFIC ACYL-COA DEHYDROGENASE, MITOCHONDRIAL
Dapma7bEVm000715	CYTOCHROME P450 4AE1-RELATED
Dapma7bEVm000732	ERO1-RELATED
Dapma7bEVm000747	MICROSOMAL GLUTATHIONE S-TRANSFERASE
Dapma7bEVm000798	SUBFAMILY NOT NAMED
Dapma7bEVm000941	CYTOCHROME C OXIDASE ASSEMBLY PROTEIN COX15 HOMOLOG
Dapma7bEVm000957	ALCOHOL DEHYDROGENASE CLASS-3
Dapma7bEVm000958	GLYCEROL-3-PHOSPHATE DEHYDROGENASE [NAD(+)]-RELATED
Dapma7bEVm000980	SUBFAMILY NOT NAMED
Dapma7bEVm001079	SUBFAMILY NOT NAMED
Dapma7bEVm001190	MALE STERILITY PROTEIN 2-RELATED
Dapma7bEVm001214	SUBFAMILY NOT NAMED
Dapma7bEVm001242	4-HYDROXYPHENYLPYRUVATE DIOXYGENASE
Dapma7bEVm001386	GH11385P
Dapma7bEVm001491	SUBFAMILY NOT NAMED
Dapma7bEVm001499	ALPHA-AMINOADIPIC SEMIALDEHYDE DEHYDROGENASE
Dapma7bEVm001529	PYRIDINE NUCLEOTIDE-DISULFIDE OXIDOREDUCTASE DOMAIN-CONTAINING PROTEIN 1
Dapma7bEVm001546	SUBFAMILY NOT NAMED
Dapma7bEVm001551	LONG-CHAIN SPECIFIC ACYL-COA DEHYDROGENASE, MITOCHONDRIAL
Dapma7bEVm001716	FAMILY NOT NAMED
Dapma7bEVm001876	GLUCOSE-6-PHOSPHATE 1-DEHYDROGENASE
Dapma7bEVm001880	SUBFAMILY NOT NAMED
Dapma7bEVm001943	TRYPTOPHAN 2,3-DIOXYGENASE
Dapma7bEVm001959	HISTONE DEACETYLASE
Dapma7bEVm001983	MALATE DEHYDROGENASE
Dapma7bEVm002027	PERSULFIDE DIOXYGENASE ETHE1, MITOCHONDRIAL
Dapma7bEVm002209	SUCCINATE DEHYDROGENASE [UBIQUINONE] FLAVOPROTEIN SUBUNIT, MITOCHONDRIAL
Dapma7bEVm002241	ACYL-COA DEHYDROGENASE FAMILY MEMBER 11
Dapma7bEVm002294	D-LACTATE DEHYDROGENASE, MITOCHONDRIAL-RELATED
Dapma7bEVm002353	OXIDOREDUCTASE AIM17-RELATED

Dapma7bEVm002433	ZGC:113142
Dapma7bEVm002597	PEPTIDYL-ALPHA-HYDROXYGLYCINE ALPHA-AMIDATING LYASE 1-RELATED
Dapma7bEVm002652	NADPH OXIDASE 5
Dapma7bEVm002733	ZGC:66484
Dapma7bEVm002780	ISOCITRATE DEHYDROGENASE [NAD] SUBUNIT GAMMA, MITOCHONDRIAL
Dapma7bEVm002894	SURFEIT LOCUS PROTEIN 1
Dapma7bEVm003024	MONOFUNCTIONAL C1-TETRAHYDROFOLATE SYNTHASE, MITOCHONDRIAL
Dapma7bEVm003085	TRIFUNCTIONAL ENZYME SUBUNIT ALPHA, MITOCHONDRIAL
Dapma7bEVm003191	SUBFAMILY NOT NAMED
Dapma7bEVm003714	MALE STERILITY PROTEIN 2-RELATED
Dapma7bEVm003872	ZGC:101765-RELATED
Dapma7bEVm003936	SUPEROXIDE DISMUTASE [MN], MITOCHONDRIAL
Dapma7bEVm004064	PEROXISOMAL ACYL-COENZYME A OXIDASE 3
Dapma7bEVm004249	SUBFAMILY NOT NAMED
Dapma7bEVm004492	NADH-UBIQUINONE OXIDOREDUCTASE 13 KD-B SUBUNIT
Dapma7bEVm004639	MONODEHYDROASCORBATE REDUCTASE 4, PEROXISOMAL
Dapma7bEVm004798	FAMILY NOT NAMED
Dapma7bEVm004825	ISOCITRATE DEHYDROGENASE [NAD] SUBUNIT ALPHA, MITOCHONDRIAL
Dapma7bEVm004826	ENOYL-[ACYL-CARRIER-PROTEIN] REDUCTASE, MITOCHONDRIAL
Dapma7bEVm004941	PEROXIREDOXIN-6
Dapma7bEVm004987	UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX UBIQUINONE-BINDING PROTEIN QP-C
Dapma7bEVm005035	3-HYDROXY-3-METHYLGLUTARYL-COENZYME A REDUCTASE
Dapma7bEVm005293	SUBFAMILY NOT NAMED
Dapma7bEVm005301	SUBFAMILY NOT NAMED
Dapma7bEVm005658	LAMBDA-CRYSTALLIN HOMOLOG
Dapma7bEVm005735	HEME OXYGENASE
Dapma7bEVm005811	FAD NADPH DEHYDROGENASE/OXIDOREDUCTASE
Dapma7bEVm005844	CARDINAL
Dapma7bEVm006079	SUBFAMILY NOT NAMED
Dapma7bEVm006113	NADH DEHYDROGENASE [UBIQUINONE] IRON-SULFUR PROTEIN 4, MITOCHONDRIAL
Dapma7bEVm006337	3-HYDROXYANTHRANILATE 3,4-DIOXYGENASE
Dapma7bEVm006375	3-OXO-5-ALPHA-STEROID 4-DEHYDROGENASE 1
Dapma7bEVm006495	4-TRIMETHYLAMINOBUTYRALDEHYDE DEHYDROGENASE
Dapma7bEVm006616	PROTEIN SCO1 HOMOLOG, MITOCHONDRIAL
Dapma7bEVm006694	MICROSOMAL GLUTATHIONE S-TRANSFERASE
Dapma7bEVm006854	SC2
Dapma7bEVm006877	SI:DKEY-12E7.4
Dapma7bEVm007051	CARBONYL REDUCTASE, ISOFORM A
Dapma7bEVm007175	FAMILY NOT NAMED
Dapma7bEVm007287	DIHYDROPTERIDINE REDUCTASE
Dapma7bEVm007534	METHIONINE SULFOXIDE REDUCTASE

Dapma7bEVm007542	SUPEROXIDE DISMUTASE [CU-ZN]
Dapma7bEVm007632	METHYLENETETRAHYDROFOLATE DEHYDROGENASE (NADP+-DEPENDENT) 1A
Dapma7bEVm007646	ALLYL ALCOHOL DEHYDROGENASE-LIKE PROTEIN-RELATED
Dapma7bEVm008044	FRATAXIN, MITOCHONDRIAL
Dapma7bEVm008073	MICROSOMAL GLUTATHIONE S-TRANSFERASE
Dapma7bEVm008098	SUPEROXIDE DISMUTASE [CU-ZN]
Dapma7bEVm008286	NADH-UBIQUINONE OXIDOREDUCTASE SUBUNIT B14.5B
Dapma7bEVm008361	1,3-PROPANEDIOL DEHYDROGENASE-RELATED
Dapma7bEVm008489	GEO09626P1
Dapma7bEVm008565	DELTA(7)-STEROL 5(6)-DESATURASE
Dapma7bEVm008622	UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX 6.4 KD PROTEIN
Dapma7bEVm009049	GLUTATHIONE PEROXIDASE-RELATED
Dapma7bEVm009650	SUBFAMILY NOT NAMED
Dapma7bEVm009688	SUBFAMILY NOT NAMED
Dapma7bEVm009918	ISOCITRATE DEHYDROGENASE [NAD] SUBUNIT BETA, MITOCHONDRIAL
Dapma7bEVm010033	MITOCHONDRIAL 10-FORMYLTETRAHYDROFOLATE DEHYDROGENASE
Dapma7bEVm010075	FAMILY NOT NAMED
Dapma7bEVm010261	METHYLMALONATE-SEMIALDEHYDE DEHYDROGENASE [ACYLATING], MITOCHONDRIAL
Dapma7bEVm010300	SI:DKEY-10O6.2
Dapma7bEVm010418	SUBFAMILY NOT NAMED
Dapma7bEVm010506	SUBFAMILY NOT NAMED
Dapma7bEVm010518	SI:DKEY-10O6.2
Dapma7bEVm010549	NADH DEHYDROGENASE [UBIQUINONE] 1 BETA SUBCOMPLEX SUBUNIT 11, MITOCHONDRIAL
Dapma7bEVm010828	D-3-PHOSPHOGLYCERATE DEHYDROGENASE
Dapma7bEVm010848	UREIDOGLYCOLATE DEHYDROGENASE (NAD(+))
Dapma7bEVm011006	CYTOCHROME C OXIDASE SUBUNIT 6A, MITOCHONDRIAL
Dapma7bEVm011240	HOMOGENTISATE 1,2-DIOXYGENASE
Dapma7bEVm011365	3-OXO-5-ALPHA-STEROID 4-DEHYDROGENASE 1
Dapma7bEVm012058	ZGC:113142
Dapma7bEVm012296	SUBFAMILY NOT NAMED
Dapma7bEVm012650	UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX, SUBUNIT X
Dapma7bEVm012911	GLUTAREDOXIN 2
Dapma7bEVm012943	SUBFAMILY NOT NAMED
Dapma7bEVm013092	SUBFAMILY NOT NAMED
Dapma7bEVm013358	PEROXIREDOXIN-5, MITOCHONDRIAL
Dapma7bEVm014983	SUBFAMILY NOT NAMED
Dapma7bEVm014984	SUBFAMILY NOT NAMED
Dapma7bEVm015164	DIHYDROFOLATE REDUCTASE
Dapma7bEVm015378	SUBFAMILY NOT NAMED
Dapma7bEVm015475	GH11385P
Dapma7bEVm015478	ACYL-COA SYNTHETASE SHORT-CHAIN FAMILY MEMBER 3, MITOCHONDRIAL

Dapma7bEVm015577	SUCCINATE DEHYDROGENASE
Dapma7bEVm015741	YTOCHROME B
Dapma7bEVm015753	SUBFAMILY NOT NAMED
Dapma7bEVm016056	ZGC:113142
Dapma7bEVm016601	ZGC:113142
Dapma7bEVm016721	SUBFAMILY NOT NAMED
Dapma7bEVm018544	SUBFAMILY NOT NAMED
Dapma7bEVm019204	SUBFAMILY NOT NAMED
Dapma7bEVm023103	POLYPRENOL REDUCTASE
Dapma7bEVm023119	SUBFAMILY NOT NAMED
Dapma7bEVm023143	PEROXIREDOXIN-5, MITOCHONDRIAL
Dapma7bEVm023366	SUBFAMILY NOT NAMED
Dapma7bEVm026608	ZGC:113142
Dapma7bEVm027306	PEROXIREDOXIN
Dapma7bEVm027569	L-2-HYDROXYGLUTARATE DEHYDROGENASE, MITOCHONDRIAL
Dapma7bEVm027655	GLUCOSE-6-PHOSPHATE 1-DEHYDROGENASE
Dapma7bEVm027706	ZGC:113142
Dapma7bEVm027888	LAMIN-B RECEPTOR
Dapma7bEVm027936	FAMILY NOT NAMED
Dapma7bEVm028220	MICROSOMAL GLUTATHIONE S-TRANSFERASE
Dapma7bEVm028228	SUPEROXIDE DISMUTASE [CU-ZN] 3
Dapma7bEVm028517	SUBFAMILY NOT NAMED
Dapma7bEVm028539	SUBFAMILY NOT NAMED
Dapma7bEVm029274	GLUTATHIONE PEROXIDASE-RELATED
Dapma7bEVm029635	GH11385P

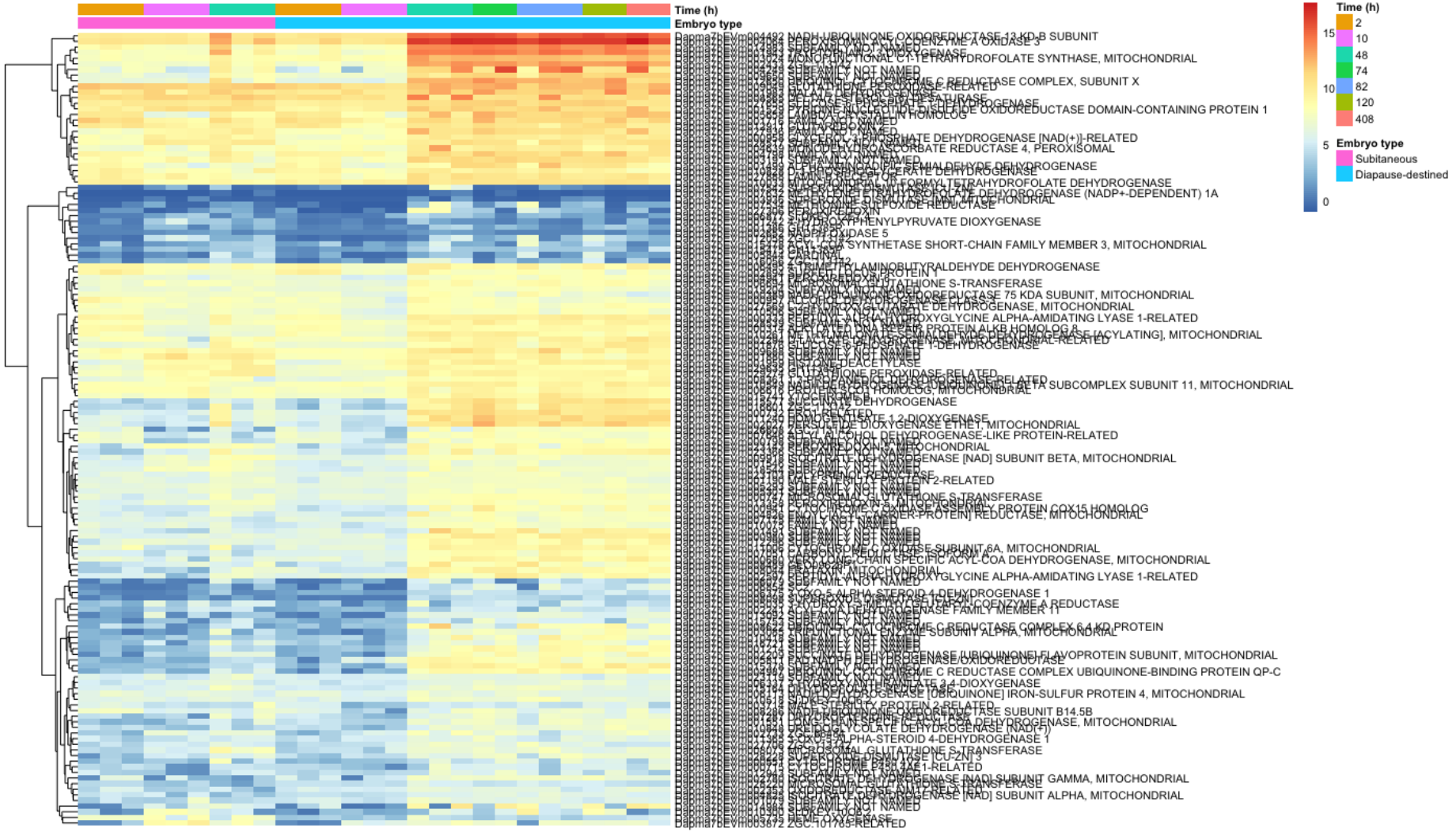


Figure 10-7 Heat map and hierarchical clustering of expression of genes assigned to the 'oxidoreductase activity' gene ontology term that were significantly up-regulated in expression in samples of sexual *Daphnia magna* embryos collected at 48 h versus 10 h post-ovulation. Column labels individual biological replicates (see key for embryo type and collection time). Colours indicate low expression (blue) and high expression (red) on log-normalised scale

10.8. Investigation of genes down-regulated upon initiation of diapause (48 h versus 10 h post-ovulation) annotated within the gene ontology terms ‘DNA replication’ and ‘cell cycle’

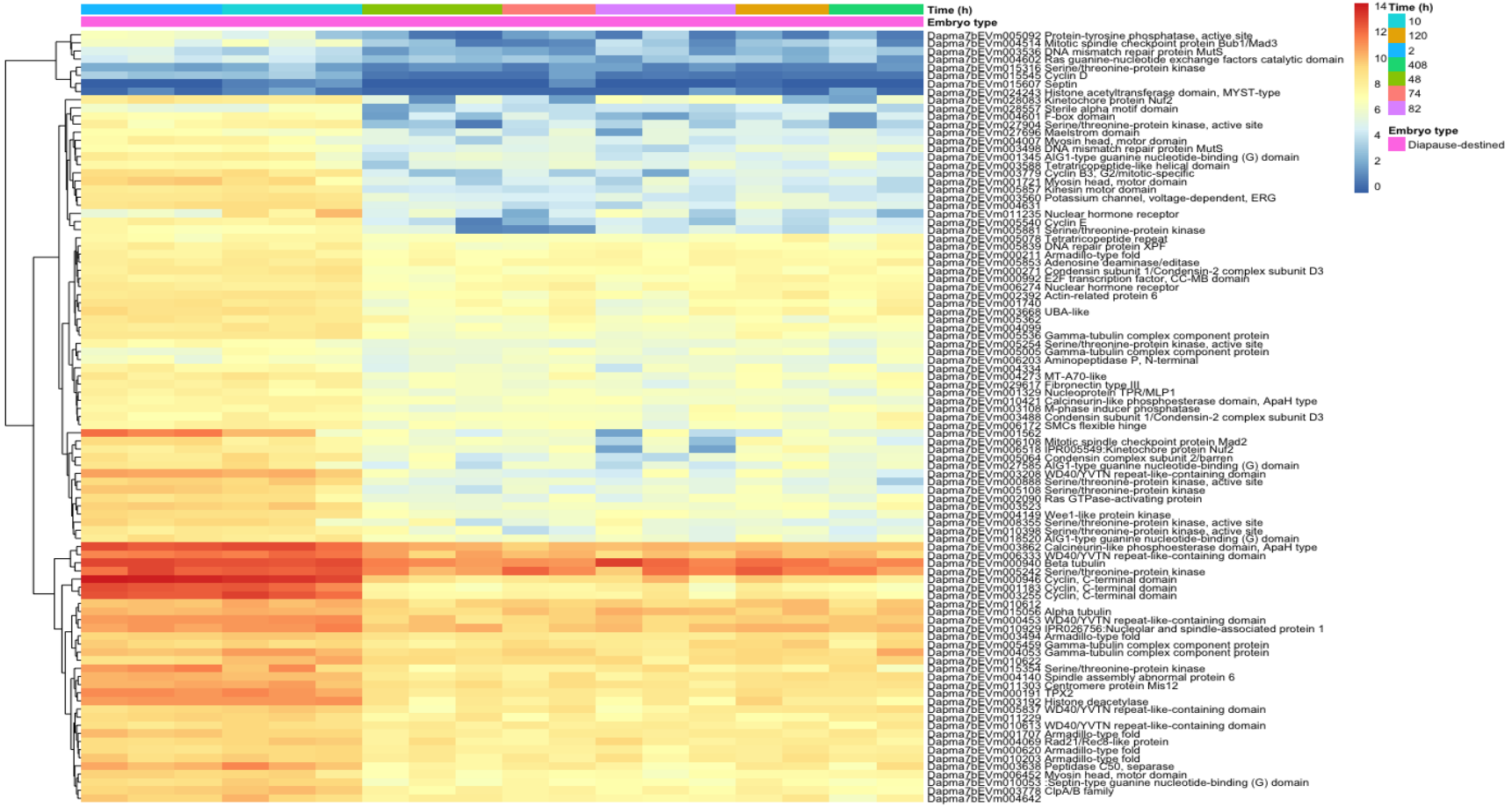


Figure 10-8 Heat map and hierarchical clustering of expression of genes assigned to the ‘cell cycle’ gene ontology term that were significantly down-regulated in expression in samples of sexual *Daphnia magna* embryos collected at 48 h versus 10 h post-ovulation. Row labels indicate gene name with corresponding annotation based on OrthoDB orthologous gene family group (also listed in **Table 10-11**) and coloured column top panels indicate time post-ovulation (refer to legend).. Colours indicate low expression (blue) and high expression (red) on log-normalised scale.

Table 10-11 Table of *Daphnia magna* genes assigned to the 'cell cycle' gene ontology term, their corresponding OrthodDB gene family and putative protein identity that were significantly up-regulated in expression in samples of sexual embryos collected at 74 h versus 48 h post-ovulation (see Figure 10-8 for expression).

Gene ID	OrthoDB orthologous group	Putative identity
Dapma7bEVm000191	EOG091G0MIV	TPX2
Dapma7bEVm000211	EOG091G026T	Armadillo-type fold
Dapma7bEVm000271	EOG091G00PU	Condensin subunit 1/Condensin-2 complex subunit D3
Dapma7bEVm000292	EOG091G020Q	Mini-chromosome maintenance, conserved site
Dapma7bEVm000453	EOG091G01L0	WD40/YVTN repeat-like-containing domain
Dapma7bEVm000620	EOG091G0195	Armadillo-type fold
Dapma7bEVm000830	EOG091G01BH	MCM domain
Dapma7bEVm000888	EOG091G0D89	Serine/threonine-protein kinase, active site
Dapma7bEVm000940	EOG091G06U2	Beta tubulin
Dapma7bEVm000946	EOG091G0E9B	Cyclin, C-terminal domain
Dapma7bEVm000992	EOG091G0CJQ	E2F transcription factor, CC-MB domain
Dapma7bEVm001031	EOG091G09WT	CWF11 family
Dapma7bEVm001090	EOG091G02B0	DNA replication licensing factor Mcm3
Dapma7bEVm001183	EOG091G0E9B	Cyclin, C-terminal domain
Dapma7bEVm001329	EOG091G006U	Nucleoprotein TPR/MLP1
Dapma7bEVm001345	EOG091G0I0X	AIG1-type guanine nucleotide-binding (G) domain
Dapma7bEVm001562	EOG091G1E1J	Putative cyclin B
Dapma7bEVm001707	EOG091G042I	Armadillo-type fold
Dapma7bEVm001718	EOG091G01N1	Dna2
Dapma7bEVm001721	EOG091G0136	Myosin head, motor domain
Dapma7bEVm001740	EOG091G09E2	lin-54 DREAM MuvB core complex component
Dapma7bEVm001769	EOG091G07YA	chromatin licensing and DNA replication factor 1
Dapma7bEVm002090	EOG091G00ET	Ras GTPase-activating protein
Dapma7bEVm002105	EOG091G01KA	DNA repair protein Rad50, eukaryotes
Dapma7bEVm002392	EOG091G09HG	Actin-related protein 6
Dapma7bEVm002619	EOG091G0GP7	Protein kinase Mps1 family
Dapma7bEVm002789	EOG091G0C0E	Flap endonuclease 1
Dapma7bEVm003108	EOG091G0LFI	M-phase inducer phosphatase
Dapma7bEVm003124	EOG091G09IO	minichromosome maintenance complex binding protein
Dapma7bEVm003192	EOG091G067J	Histone deacetylase
Dapma7bEVm003208	EOG091G06FJ	WD40/YVTN repeat-like-containing domain
Dapma7bEVm003255	EOG091G090I	Cyclin, C-terminal domain
Dapma7bEVm003436	EOG091G19OS	Geminin/Multicilin
Dapma7bEVm003488	EOG091G00PU	Condensin subunit 1/Condensin-2 complex subunit D3
Dapma7bEVm003494	EOG091G00ZG	Armadillo-type fold
Dapma7bEVm003498	EOG091G018J	DNA mismatch repair protein MutS
Dapma7bEVm003523	EOG091G01HH	
Dapma7bEVm003536	EOG091G026W	DNA mismatch repair protein MutS

Dapma7bEVm003560	EOG091G024C	Potassium channel, voltage-dependent, ERG
Dapma7bEVm003588	EOG091G05K7	Tetratricopeptide-like helical domain
Dapma7bEVm003638	EOG091G03VI	Peptidase C50, separase
Dapma7bEVm003668	EOG091G0DFP	UBA-like
Dapma7bEVm003778	EOG091G0D83	ClpA/B family
Dapma7bEVm003779	EOG091G0L84	Cyclin B3, G2/mitotic-specific
Dapma7bEVm003850	EOG091G00YT	Nucleosome assembly protein (NAP)
Dapma7bEVm003862	EOG091G0B6S	Calcineurin-like phosphoesterase domain, ApaH type
Dapma7bEVm003973	EOG091G006J	DNA-directed DNA polymerase, family B
Dapma7bEVm004007	EOG091G0136	Myosin head, motor domain
Dapma7bEVm004053	EOG091G01V2	Gamma-tubulin complex component protein
Dapma7bEVm004069	EOG091G03QW	Rad21/Rec8-like protein
Dapma7bEVm004099	EOG091G0QVI	N-acetyltransferase ESCO, zinc finger
Dapma7bEVm004140	EOG091G09W4	Spindle assembly abnormal protein 6
Dapma7bEVm004149	EOG091G049H	Wee1-like protein kinase
Dapma7bEVm004273	EOG091G0BBQ	MT-A70-like
Dapma7bEVm004334	EOG091G02HO	centrosomal protein 85kDa-like
Dapma7bEVm004514	EOG091G00MB	Mitotic spindle checkpoint protein Bub1/Mad3
Dapma7bEVm004524	EOG091G04KI	zinc finger, NFX-type containing 1
Dapma7bEVm004558	EOG091G026H	DNA replication licensing factor Mcm2
Dapma7bEVm004590	EOG091G02LB	DNA replication licensing factor Mcm7
Dapma7bEVm004601	EOG091G1ER5	F-box domain
Dapma7bEVm004602	EOG091G0NUM	Ras guanine-nucleotide exchange factors catalytic domain
Dapma7bEVm004631	EOG091G046M	NDC80 kinetochore complex component
Dapma7bEVm004642	EOG091G0MME	
Dapma7bEVm004742	EOG091G05VI	Polymerase, nucleotidyl transferase domain
Dapma7bEVm005005	EOG091G0ABS	Gamma-tubulin complex component protein
Dapma7bEVm005064	EOG091G01HE	Condensin complex subunit 2/barren
Dapma7bEVm005078	EOG091G0HW1	Tetratricopeptide repeat
Dapma7bEVm005092	EOG091G09VG	Protein-tyrosine phosphatase, active site
Dapma7bEVm005108	EOG091G0D89	Serine/threonine-protein kinase
Dapma7bEVm005242	EOG091G0D89	Serine/threonine-protein kinase
Dapma7bEVm005254	EOG091G0CI9	Serine/threonine-protein kinase, active site
Dapma7bEVm005362	EOG091G0LDA	MOB kinase activator 3C
Dapma7bEVm005459	EOG091G01V2	Gamma-tubulin complex component protein
Dapma7bEVm005460	EOG091G02V8	Mini-chromosome maintenance complex protein 4
Dapma7bEVm005470	EOG091G0BA1	Bromo adjacent homology (BAH) domain
Dapma7bEVm005536	EOG091G01V2	Gamma-tubulin complex component protein
Dapma7bEVm005540	EOG091G0DAW	Cyclin E
Dapma7bEVm005837	EOG091G04F0	WD40/YVTN repeat-like-containing domain
Dapma7bEVm005839	EOG091G0554	DNA repair protein XPF
Dapma7bEVm005853	EOG091G0MZP	Adenosine deaminase/editase
Dapma7bEVm005857	EOG091G0DJX	Kinesin motor domain
Dapma7bEVm005881	EOG091G0OE6	Serine/threonine-protein kinase
Dapma7bEVm006108	EOG091G0JBM	Mitotic spindle checkpoint protein Mad2
Dapma7bEVm006172	EOG091G00R5	SMCs flexible hinge

Dapma7bEVm006203	EOG091G0BAV	Aminopeptidase P, N-terminal
Dapma7bEVm006274	EOG091G0YX6	Nuclear hormone receptor
Dapma7bEVm006333	EOG091G0AB8	WD40/YVTN repeat-like-containing domain
Dapma7bEVm006452	EOG091G00DE	Myosin head, motor domain
Dapma7bEVm006518	EOG091G0YDU	Kinetochore protein Nuf2
Dapma7bEVm008355	EOG091G0CH0	Serine/threonine-protein kinase, active site
Dapma7bEVm010053	EOG091G0IKM	Septin-type guanine nucleotide-binding (G) domain
Dapma7bEVm010203	EOG091G00YR	Armadillo-type fold
Dapma7bEVm010398	EOG091G048P	Serine/threonine-protein kinase, active site
Dapma7bEVm010421	EOG091G0B6S	Calcineurin-like phosphoesterase domain, ApaH type
Dapma7bEVm010612	EOG091G0B3V	coilin
Dapma7bEVm010613	EOG091G09BB	WD40/YVTN repeat-like-containing domain
Dapma7bEVm010622	EOG091G0A1S	
Dapma7bEVm010920	EOG091G027N	DNA replication licensing factor Mcm
Dapma7bEVm010929	EOG091G17LQ	Nucleolar and spindle-associated protein 1
Dapma7bEVm011229	EOG091G085F	
Dapma7bEVm011235	EOG091G0AV3	Nuclear hormone receptor
Dapma7bEVm011303	EOG091G0O0P	Centromere protein Mis12
Dapma7bEVm011705	EOG091G07ZH	XPG/Rad2 endonuclease
Dapma7bEVm015056	EOG091G0736	Alpha tubulin
Dapma7bEVm015316	EOG091G0CH0	Serine/threonine-protein kinase
Dapma7bEVm015354	EOG091G0I1Q	Serine/threonine-protein kinase
Dapma7bEVm015545	EOG091G0URX	Cyclin D
Dapma7bEVm015607	EOG091G03TM	Septin
Dapma7bEVm018520	EOG091G0I0X	AIG1-type guanine nucleotide-binding (G) domain
Dapma7bEVm024243	EOG091G00D2	Histone acetyltransferase domain, MYST-type
Dapma7bEVm027585	EOG091G0I0X	AIG1-type guanine nucleotide-binding (G) domain
Dapma7bEVm027696	EOG091G0BYM	Maelstrom domain
Dapma7bEVm027856	EOG091G084W	WD40/YVTN repeat-like-containing domain
Dapma7bEVm027904	EOG091G00MB	Serine/threonine-protein kinase, active site
Dapma7bEVm028083	EOG091G0YDU	Kinetochore protein Nuf2
Dapma7bEVm028557	EOG091G0WL3	Sterile alpha motif domain
Dapma7bEVm029617	EOG091G1V4B	Fibronectin type III

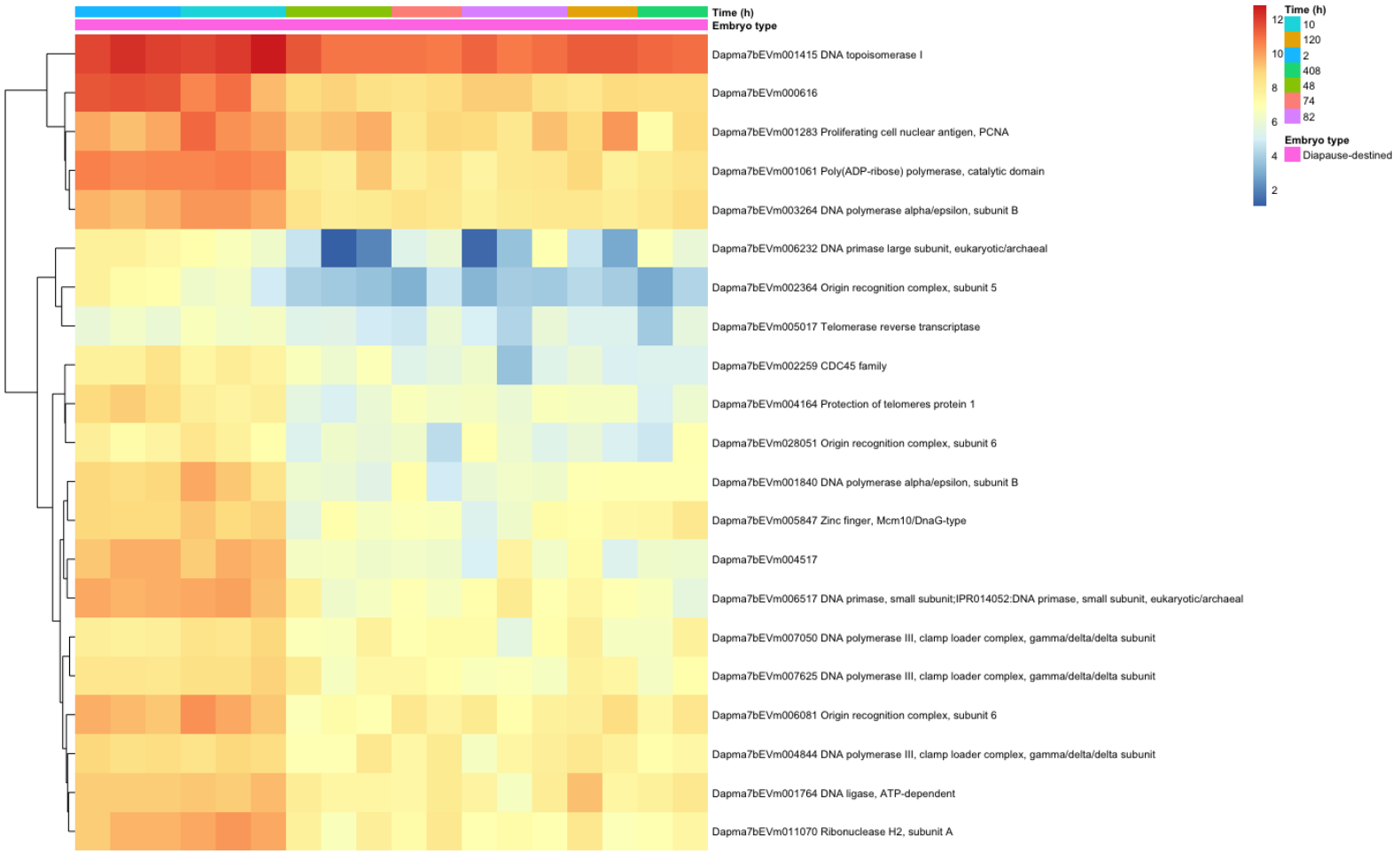


Figure 10-9 Heat map and hierarchical clustering of expression of genes assigned to the 'DNA replication' gene ontology term that were significantly down-regulated in expression in samples of sexual *Daphnia magna* embryos collected at 48 h versus 10 h post-ovulation. Row labels indicate gene name with corresponding annotation based on OrthoDB orthologous gene family group (also listed in **Table 10-12**) and coloured column top panels indicate time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale.

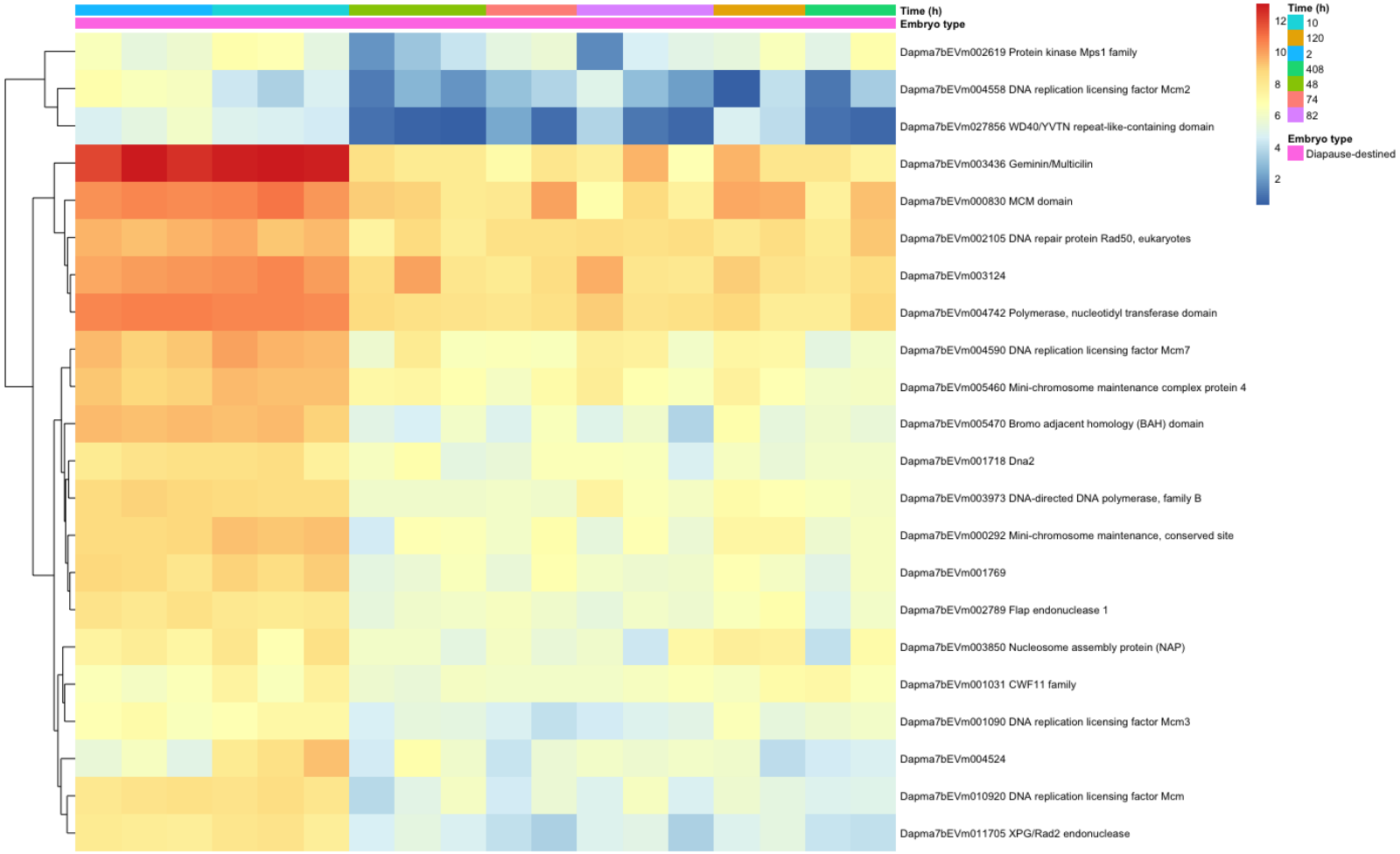


Figure 10-10 **Heat map and hierarchical clustering of expression of genes assigned to both 'cell cycle' and 'DNA replication' gene ontology terms that were that were significantly down-regulated in expression in samples of sexual *Daphnia magna* embryos collected at 48 h versus 10 h post-ovulation.** Row labels indicate gene name with corresponding annotation based on OrthoDB orthologous gene family group (also listed in **Table 10-11** and **Table 10-12**) and coloured column top panels indicate time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale.

Table 10-12 Table of *Daphnia magna* genes assigned to the 'DNA replication' gene ontology term, their corresponding OrthodDB gene family and putative protein identity that were significantly up-regulated in expression in samples of sexual embryos collected at 74 h versus 48 h post-ovulation.

Gene ID	OrthoDB orthologous group	Putative identity
Dapma7bEVm000292	EOG091G020Q	DNA replication licensing factor Mcm6
Dapma7bEVm000616	EOG091G01CH	
Dapma7bEVm000830	EOG091G01BH	MCM domain
Dapma7bEVm001031	EOG091G09WT	CWF11 family
Dapma7bEVm001061	EOG091G13H1	Poly(ADP-ribose) polymerase, catalytic domain
Dapma7bEVm001090	EOG091G02B0	DNA replication licensing factor Mcm3
Dapma7bEVm001283	EOG091G0GQ7	Proliferating cell nuclear antigen, PCNA
Dapma7bEVm001415	EOG091G02KT	DNA topoisomerase I
Dapma7bEVm001718	EOG091G01N1	Dna2
Dapma7bEVm001764	EOG091G06OS	DNA ligase, ATP-dependent
Dapma7bEVm001769	EOG091G07YA	
Dapma7bEVm001840	EOG091G05Q6	DNA polymerase alpha/epsilon, subunit B
Dapma7bEVm002105	EOG091G01KA	DNA repair protein Rad50, eukaryotes
Dapma7bEVm002259	EOG091G05EU	CDC45 family
Dapma7bEVm002364	EOG091G0D0I	Origin recognition complex, subunit 5
Dapma7bEVm002619	EOG091G0GP7	Protein kinase Mps1 family
Dapma7bEVm002789	EOG091G0C0E	Flap endonuclease 1
Dapma7bEVm003124	EOG091G09IO	
Dapma7bEVm003264	EOG091G094I	DNA polymerase alpha/epsilon, subunit B
Dapma7bEVm003436	EOG091G19OS	Geminin/Multicilin
Dapma7bEVm003850	EOG091G00YT	Nucleosome assembly protein (NAP)
Dapma7bEVm003973	EOG091G006J	DNA-directed DNA polymerase, family B
Dapma7bEVm004164	EOG091G01T8	Protection of telomeres protein 1
Dapma7bEVm004517	EOG091G00YS	
Dapma7bEVm004524	EOG091G04KI	
Dapma7bEVm004558	EOG091G026H	DNA replication licensing factor Mcm2;IPR001208:MCM domain
Dapma7bEVm004590	EOG091G02LB	DNA replication licensing factor Mcm7;IPR001208:MCM domain
Dapma7bEVm004742	EOG091G05VI	Polymerase, nucleotidyl transferase domain
Dapma7bEVm004844	EOG091G0CKT	DNA polymerase III, clamp loader complex, gamma/delta/delta subunit
Dapma7bEVm005017	EOG091G04DO	Telomerase reverse transcriptase
Dapma7bEVm005460	EOG091G02V8	Mini-chromosome maintenance complex protein 4
Dapma7bEVm005470	EOG091G0BA1	Bromo adjacent homology (BAH) domain
Dapma7bEVm005847	EOG091G053R	Zinc finger, Mcm10/DnaG-type
Dapma7bEVm006081	EOG091G0OLA	Origin recognition complex, subunit 6
Dapma7bEVm006232	EOG091G08G7	DNA primase large subunit, eukaryotic/archaeal
Dapma7bEVm006517	EOG091G0EAF	DNA primase, small subunit;IPR014052:DNA primase, small subunit, eukaryotic/archaeal
Dapma7bEVm007050	EOG091G0CKT	DNA polymerase III, clamp loader complex, gamma/delta/delta subunit
Dapma7bEVm007625	EOG091G0BQF	DNA polymerase III, clamp loader complex, gamma/delta/delta subunit
Dapma7bEVm010920	EOG091G027N	DNA replication licensing factor Mcm

Dapma7bEVm011070	EOG091G0E06	Ribonuclease H2, subunit A
Dapma7bEVm011705	EOG091G07ZH	XPG/Rad2 endonuclease
Dapma7bEVm027856	EOG091G084W	WD40/YVTN repeat-like-containing domain
Dapma7bEVm028051	EOG091G0OLA	Origin recognition complex, subunit 6

11. Appendix C

11.1. Sparse principal component analysis of RNA-Seq data obtained from sexually produced *Daphnia magna* embryos collected between 2 h and 408 h post-ovulation

Table 11-1 Table of genes with non-zero loadings for sparse principal components (sPC) 1–4 from the sparse principal component analysis of RNA-Seq data from sexually-produced *Daphnia magna* embryos collected between 2 h and 408 h (2 weeks) post-ovulation. *Daphnia magna* gene ID is indicated with corresponding annotation with OrthoDB gene family designation and identity.

sPC1	sPC2	sPC3	sPC4	Gene ID	OrthoDB Orthologous group	Identity
-0.215	0	0	0	Dapma7bEVm000024	EOG091G017D	Perilipin 4
-0.160	0	0	0	Dapma7bEVm000198	EOG091G0I0S	Transmembrane: Helical
-0.410	0	0	0	Dapma7bEVm000246	EOG091G04NA	Glycoside hydrolase family 20, catalytic domain
0	0.047	0	0	Dapma7bEVm000502	EOG091G05ML	SLC26A/SulP transporter
0	0.146	0	0	Dapma7bEVm000766	NA	
0.211	0	0	0	Dapma7bEVm001268	EOG091G0P0L	
0.188	0	0	0	Dapma7bEVm001383	EOG091G02XX	
0	0.518	0	0	Dapma7bEVm001668	NA	
0.103	0	0	0	Dapma7bEVm002006	NA	
0	0.360	0	0	Dapma7bEVm002882	NA	
-0.187	0	0	0	Dapma7bEVm002915	EOG091G0MBB	Glutathione S-transferase, N-terminal
0	0.296	0	0	Dapma7bEVm003677	EOG091G0PPX	Chitin binding domain
0	0	0.430	0	Dapma7bEVm004290	EOG091G0D5F	Sulfotransferase domain
-0.279	0	0	0	Dapma7bEVm004331	EOG091G086D	Innexin
0	0.307	0	0	Dapma7bEVm004432	NA	
-0.066	0	0	0	Dapma7bEVm004536	EOG091G0NA2	suprabasin
0.620	0	0	0	Dapma7bEVm004775	EOG091G0BTD	4-hydroxyphenylpyruvate dioxygenase
-0.033	0	0	0	Dapma7bEVm004918	EOG091G0OTG	Protein-L-isoaspartate(D-aspartate) O-methyltransferase
0.067	0	0	0	Dapma7bEVm004979	NA	
-0.190	0	0	0	Dapma7bEVm005083	NA	
-0.316	0	0	0	Dapma7bEVm005309	EOG091G0HDX	CD20-like family
0	0	0	0	Dapma7bEVm005865	EOG091G02YK	Phosphoenolpyruvate carboxykinase, GTP-utilising
-0.395	0	0	0	Dapma7bEVm006050	EOG091G0C3E	EMI domain
-0.091	0	0	0	Dapma7bEVm006339	EOG091G0EEJ	Quinone oxidoreductase/zeta-crystallin, conserved site
-0.481	0	0	0	Dapma7bEVm006430	NA	
-0.073	0	0	0	Dapma7bEVm006625	EOG091G0UCP	breast cancer metastasis-suppressor 1-like
-0.242	0	0	0	Dapma7bEVm006659	NA	
-0.052	0	0	0	Dapma7bEVm006961	EOG091G04AX	
-0.142	0	0	0	Dapma7bEVm007095	EOG091G1O41	
-0.130	0	0	0	Dapma7bEVm007102	EOG091G0NJA	Tetraspanin
0	0.197	0	0	Dapma7bEVm007132	EOG091G0TQQ	Invertebrate-type lysozyme
-0.021	0	0	0	Dapma7bEVm007137	NA	
0	0	0.673	0	Dapma7bEVm007171	NA	
0	0	0.119	0	Dapma7bEVm007268	NA	
-0.643	0	0	0	Dapma7bEVm007647	EOG091G009U	Peptidase M12A
-0.209	0	0	0	Dapma7bEVm007713	NA	
0	0.288	0	0	Dapma7bEVm007743	NA	
-0.064	0	0	0	Dapma7bEVm007748	EOG091G09TM	BTB/POZ domain

-0.212	0	0	0	Dapma7bEVm007878	NA	
-0.056	0	0	0	Dapma7bEVm008073	EOG091G0TU0	Transmembrane: Helical
-0.036	0	0	0	Dapma7bEVm008098	EOG091G0OG2	Superoxide dismutase, copper/zinc binding domain
-0.403	0	0	0	Dapma7bEVm008142	EOG091G0ION	Serpin family
-0.019	0	0	0	Dapma7bEVm008215	EOG091G0T79	Protein-lysine N-methyltransferase Efm4
-0.424	0	0	0	Dapma7bEVm008280	NA	
-0.129	0	0	0	Dapma7bEVm008306	NA	
-0.176	0	0	0	Dapma7bEVm008409	NA	
-0.324	0	0	0	Dapma7bEVm008418	EOG091G0J20	Ferritin
-0.155	0	0	0	Dapma7bEVm008540	NA	
-0.025	0	0	0	Dapma7bEVm008590	NA	
-0.598	0	0	0	Dapma7bEVm008619	NA	
-0.221	0	0	0	Dapma7bEVm008639	NA	
0	0.224	0	0	Dapma7bEVm008651	NA	
-0.278	0	0	0	Dapma7bEVm009045	EOG091G063D	mRNA capping enzyme, catalytic domain
-0.283	0	0	0	Dapma7bEVm009587	EOG091G0FMS	Major intrinsic protein
-0.104	0	0	0	Dapma7bEVm009616	EOG091G06A1	AMP-dependent synthetase/ligase
0.320	0	0	0	Dapma7bEVm009662	EOG091G0P0L	
0	0.103	0	0	Dapma7bEVm009838	NA	
0	0	0	-0.337	Dapma7bEVm010179	EOG091G1600	Sugar/inositol transporter
0	0.039	0	0	Dapma7bEVm010318	EOG091G0UOF	Tumour necrosis factor-like domain
-0.004	0	0	0	Dapma7bEVm010488	EOG091G07YB	Malate synthase
0	0.672	0	0	Dapma7bEVm010544	NA	
0.105	0	0	0	Dapma7bEVm010645	EOG091G0E18	Claudin, conserved site
-0.155	0	0	0	Dapma7bEVm010842	EOG091G0196	Armadillo-type fold
-0.086	0	0	0	Dapma7bEVm011332	EOG091G05ML	SLC26A/SulP transporter
-0.074	0	0	0	Dapma7bEVm011614	EOG091G0Q2M	Basic-leucine zipper domain
-0.337	0	0	0	Dapma7bEVm012058	NA	
-0.195	0	0	0	Dapma7bEVm012181	NA	
-0.190	0	0	0	Dapma7bEVm012303	NA	
-0.284	0	0	0	Dapma7bEVm012383	EOG091G08ES	Zinc finger, ZZ-type
-0.124	0	0	0	Dapma7bEVm012625	NA	
-0.380	0	0	0	Dapma7bEVm012695	NA	
-0.269	0	0	0	Dapma7bEVm012800	NA	
-0.077	0	0	0	Dapma7bEVm012881	EOG091G0ET2	Caenorhabditis elegans Protein K08E4.3
-0.307	0	0	0	Dapma7bEVm012943	EOG091G0BT8	Cytochrome P450, E-class, group I
-0.123	0	0	0	Dapma7bEVm012992	EOG091G0D6Q	Perlipin
-0.262	0	0	0	Dapma7bEVm013071	EOG091G0DB2	
-0.112	0	0	0	Dapma7bEVm013098	NA	Sinc finger protein 414
-0.195	0	0	0	Dapma7bEVm013849	EOG091G0QZU	Transmembrane: Helical
-0.128	0	0	0	Dapma7bEVm013920	NA	
-0.297	0	0	0	Dapma7bEVm014768	NA	
-0.043	0	0	0	Dapma7bEVm015240	NA	
-0.194	0	0	0	Dapma7bEVm015255	EOG091G09CU	Membrane dipeptidase, active site
0.269	0	0	0	Dapma7bEVm015263	EOG091G0E18	Claudin, conserved site
0.149	0	0	0	Dapma7bEVm015264	NA	
0	0.462	0	0	Dapma7bEVm015390	NA	
-0.317	0	0	0	Dapma7bEVm015481	EOG091G00R4	Glycosyl transferase, family 20
-0.209	0	0	0	Dapma7bEVm015647	EOG091G0B2X	Heme-NO binding
-0.205	0	0	0	Dapma7bEVm015753	NA	
-0.288	0	0	0	Dapma7bEVm015787	NA	
0.260	0	0	0	Dapma7bEVm015966	EOG091G0AQW	Transmembrane: Helical
0	0	0.216	0	Dapma7bEVm015992	EOG091G0HC1	Stanniocalcin
-0.289	0	0	0	Dapma7bEVm016000	NA	
-0.028	0	0	0	Dapma7bEVm016056	NA	
-0.403	0	0	0	Dapma7bEVm016131	EOG091G0BP6	Transmembrane: Helical
0	0	0.232	0	Dapma7bEVm016149	NA	
-0.282	0	0	0	Dapma7bEVm016216	EOG091G0DDW	Cytidyltransferase-like domain
-0.195	0	0	0	Dapma7bEVm017088	EOG091G0ET2	Caenorhabditis elegans Protein K08E4.3

-0.119	0	0	0	Dapma7bEVm018290	NA	
-0.324	0	0	0	Dapma7bEVm018382	NA	
0	0.458	0	0	Dapma7bEVm018445	EOG091G0AH5	Peptidase S1A, chymotrypsin family
0.480	0	0	0	Dapma7bEVm018468	EOG091G0P0L	
-0.282	0	0	0	Dapma7bEVm018865	EOG091G0E1F	Major facilitator superfamily
-0.166	0	0	0	Dapma7bEVm019101	EOG091G02DA	Zinc finger C2H2-type
-0.166	0	0	0	Dapma7bEVm019300	EOG091G0IEO	
0	0	0.237	0	Dapma7bEVm019328	NA	
-0.174	0	0	0	Dapma7bEVm019337	EOG091G0EEJ	Quinone oxidoreductase/zeta-crystallin, conserved site
0.198	0	0	0	Dapma7bEVm019349	NA	
0.499	0	0	0	Dapma7bEVm019397	EOG091G0E18	Claudin, conserved site
-0.036	0	0	0	Dapma7bEVm019501	EOG091G0IEO	
0	0	0.232	0	Dapma7bEVm019957	EOG091G0UW4	Similarity:Contains 1 BPTI/Kunitz inhibitor domain.
0	0	0.330	0	Dapma7bEVm020076	NA	
0	0	0.486	0	Dapma7bEVm020105	EOG091G0HQD	Phosphatidylinositol 4,5-bisphosphate 4-phosphatase
-0.082	0	0	0	Dapma7bEVm020711	NA	
0	0	-0.003	0	Dapma7bEVm022793	EOG091G099F	Fork head domain
-0.236	0	0	0	Dapma7bEVm023143	EOG091G0OSA	Redoxin
0	0.825	0	0	Dapma7bEVm023209	EOG091G0AH5	Serine proteases, trypsin family, histidine active site
0	0.072	0	0	Dapma7bEVm023349	EOG091G0M9O	Transmembrane: Helical
0	0	0.449	0	Dapma7bEVm023361	EOG091G0HC1	Stanniocalcin
0	0	0.718	0	Dapma7bEVm023362	EOG091G0HC1	Stanniocalcin
-0.008	0	0	0	Dapma7bEVm023455	EOG091G09TM	BTB/POZ domain
-0.209	0	0	0	Dapma7bEVm023457	EOG091G1600	Sugar/inositol transporter
-0.050	0	0	0	Dapma7bEVm023466	EOG091G0D5F	Sulfotransferase domain
-0.040	0	0	0	Dapma7bEVm023820	EOG091G09TM	BTB/POZ domain
-0.295	0	0	0	Dapma7bEVm024492	EOG091G04GK	Protein kinase C-like, phorbol ester/diacylglycerol-binding domain
-0.486	0	0	0	Dapma7bEVm024554	NA	
0	0	0	-0.067	Dapma7bEVm024896	NA	
0	0	0.089	0	Dapma7bEVm025097	NA	
-0.030	0	0	0	Dapma7bEVm025265	EOG091G1QMV	
-0.167	0	0	0	Dapma7bEVm025387	NA	
-0.378	0	0	0	Dapma7bEVm025509	EOG091G0SR6	Succinate dehydrogenase/fumarate reductase type B, transmembrane subunit
-0.085	0	0	0	Dapma7bEVm025634	NA	
0.157	0	0	0	Dapma7bEVm025970	NA	
-0.374	0	0	0	Dapma7bEVm026608	NA	
-0.185	0	0	0	Dapma7bEVm027306	EOG091G0OSA	Redoxin
0.174	0	0	0	Dapma7bEVm027483	NA	
0.503	0	0	0	Dapma7bEVm027541	NA	
0	0.546	0	0	Dapma7bEVm027626	NA	
-0.252	0	0	0	Dapma7bEVm028186	NA	
0	0	0.597	0	Dapma7bEVm028216	NA	
-0.118	0	0	0	Dapma7bEVm028228	EOG091G0OG2	Superoxide dismutase, copper/zinc
-0.018	0	0	0	Dapma7bEVm028814	NA	
-0.747	0	0	0	Dapma7bEVm029061	NA	
0.095	0	0	0	Dapma7bEVm029190	EOG091G0P0L	
0.397	0	0	0	Dapma7bEVm029285	EOG091G0P0L	
0.034	0	0	0	Dapma7bEVm029333	EOG091G0P0L	
0.009	0	0	0	Dapma7bEVm029420	EOG091G0YCF	
-0.078	0	0	0	Dapma7bEVm030633	EOG091G0C5N	UBA-like

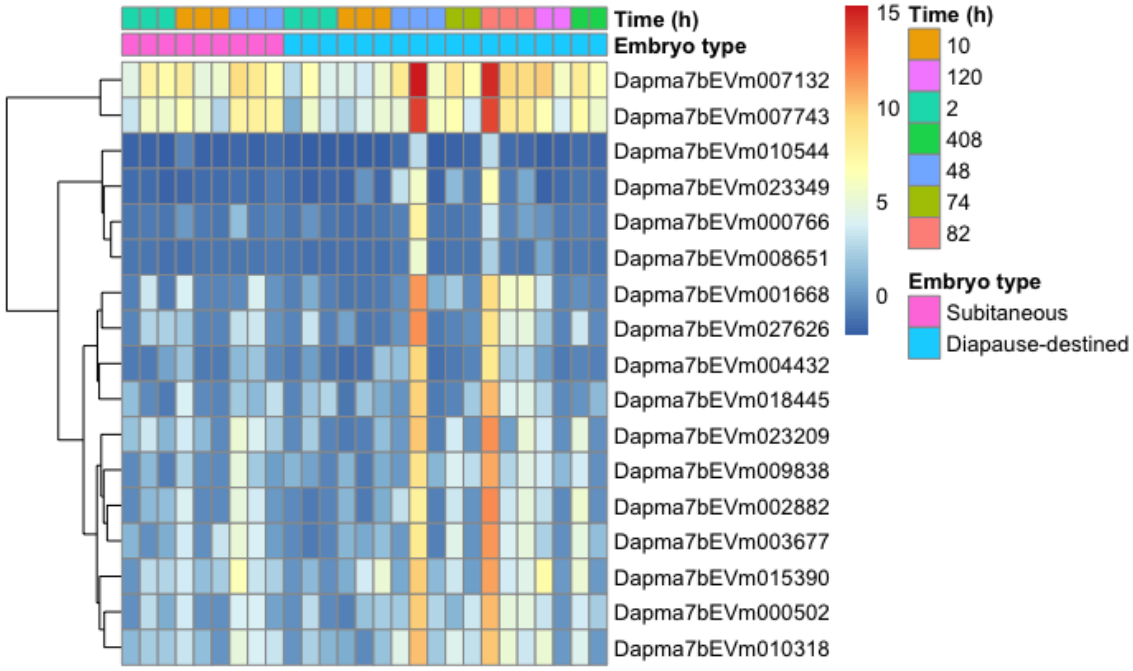


Figure 11-1 Heat map and hierarchical clustering of expression of genes with non-zero loadings in the second sparse principal component in sparse principal component analysis of RNA-Seq data from diapause-destined *Daphnia magna* embryos collected between 2 h and 2 weeks (408 h) post-ovulation, subitaneous embryos are shown for comparison only. Row labels indicate gene name and coloured column top panels indicate time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale**.**

11.2. Sparse principal component analysis of nESI-DIMS data from negative ion analysis of non-polar extracts of *Daphnia magna* embryos collected between 2 h and 17,592 h (2y) post-ovulation

Table 11-2 Table of non-polar metabolite spectral peaks with non-zero loadings for sparse principal components (sPC) 1–4 from the sparse principal component analysis of nESI-DIMS data from negative ion analysis of non-polar extracts of *Daphnia magna* embryos collected between 2h and 17,592 h (2 y) post-ovulation. Observed peak *m/z* is shown in addition to sPC loadings, bold values indicate peaks not detected in subitaneous embryo sample dataset. Formula indicates putative empirical formulae assigned to each peak using BEAMS, with corresponding annotation including number of possible species attributable to the aforementioned formula, respective putative compound identity with theoretical mass. Mass error (ppm) is the difference between theoretical and experimental *m/z*, normalised to experimental *m/z.* Other putative formula without compound ID assignment are indicated, for peaks with a high number of possible formulae only number are indicated for brevity.

sPC1	sPC2	sPC3	sPC4	*m/z*	Formula		Ion form	Compound identity	Theoretical mass	Mass error
0	0	0	-0.662	313.0583743	21	–	–	–	–	–
0	0	0	0.203	**327.1884086**	8	–	–	–	–	–
0	0	0.419	0	331.1890899	24	–	–	–	–	–
0	0.047	0	0	350.2393294	3	–	–	–	–	–
0	0	0	-0.003	357.0729023	40	–	–	–	–	–
0	0	0	-0.203	514.0618621	–	–	–	–	–	–
0	0	1.140	0	708.4604129	–	–	–	–	–	–
0	0	0	-0.073	720.4549592	–	–	–	–	–	–
0	0	0.104	0	773.5159682	–	–	–	–	–	–
0	0	0.080	0	793.5788977	–	–	–	–	–	–
0	0	0	0.606	812.500067	C44H76NO8P	23	[M+Cl]-	PC/PE	812.5002586	0.236
					C45H78NO7P	1	[M+K-2H]-	PE(P-18:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	812.5001986	0.162
0	0	-1.830	0	845.5171288	–	–	–	–	–	–
0	0	-0.606	0	863.5643078	–	–	–	–	–	–
0	0	0	0.033	868.5700342	–	–	–	–	–	–
0	0	-0.300	0	868.5972221	–	–	–	–	–	–
0	0	0	-0.244	872.63823	C48H92NO10P	4	[M-H]-	PS	872.638611	0.436
					C46H88NO8P	39	[M+Hac-H]-	PC/PE		
0	-1.135	0	0	873.707984	C47H95N2O6P		[M+Hac-H]-	SM(d18:0/24:1(15Z))	873.70663	

								SM(d18:1/24:0)		-1.550
0	0	0	0.550	881.6417022	–	–	–	–	–	–
0	0	0	0.779	882.644765	–	–	–	–	–	–
0	0	0	0.054	883.6570781	–	–	–	–	–	–
0	0	0	0.170	884.6603481	–	–	–	–	–	–
0	0	0	0.132	887.6702189	–	–	–	–	–	–
0	0	0	0.016	**887.6795157**	–	–	–	–	–	–
0	0	0	0.224	**889.6860285**	–	–	–	–	–	–
0	-0.308	0	0	**892.7038542**	–	–	–	–	–	–
0	0	0	-1.044	896.637089	C48H88NO8P	26	[M+Hac-H]-	PC/PE	896.6386106	1.698
					C48H94NO10P	3	[M+Na-2H]-	PS(21:0/21:0) PS(22:0/20:0) PS(20:0/22:0)	896.636206	-0.985
					C50H92NO10P	2	[M-H]-	PS(22:2(13Z,16Z)/22:1(11Z)) PS(22:1(11Z)/22:2(13Z,16Z))	896.638611	1.698
0	0	0	-0.416	897.6408854	C49H97O8PS		[M+Na-2H]-	–	897.6388486	-2.269
0	-0.050	0	0	900.7366324	–	–	–	–	–	–
0.033	0	0	0	905.5328754	–	–	–	–	–	–
-0.163	0	0	0	906.6895016	–	–	–	–	–	–
-0.350	-0.237	0	0	907.702135	C48H97O9P	1	[M+Hac-H]-	PG(O-20:0/22:0)	907.7008766	-1.387
					C50H101O11P	1	[M-H]-	Archaetidylglycerol-myo-inositol	907.7008766	-1.387
					C55H92O6	36	[M+Hac-H]-	TG	907.7032436	1.221
-0.219	-0.142	0	0	908.7052631	–	–	–	–	–	–
-0.277	0	0	0	**909.7176343**	C55H94O6	51	[M+Hac-H]-	–	909.7188936	1.384
0	0	0	0.367	**910.6750613**	–	–	–	–	–	–
0	0	0	-0.083	912.4849181	–	–	–	–	–	–
0	-0.807	0	0	921.6421073	–	–	–	–	–	–
0	0	0	0.038	938.7074435	–	–	–	–	–	–
0	0	-0.038	0	939.520293	–	–	–	–	–	–
0	0	-0.053	0	**942.738446**	–	–	–	–	–	–
0.185	0	0	0	950.5335979	–	–	–	–	–	–

0	0	-0.214	0.161	**957.7168413**	–	–	–	–	–	–
0	-0.633	-0.415	0	971.7337883	–	–	–	–	–	–
0	0.060	0	0	972.7224675	–	–	–	–	–	–
0.074	0	0	0	**974.5343902**	–	–	–	–	–	–
0.052	0.100	0	0	982.5005645	–	–	–	–	–	–
0.160	0	0	0	982.5980589	–	–	–	–	–	–
0	-0.564	0	0	983.8281816	C60H108O6	59	[M+Hac-H]-	TG	983.8284436	0.266
0.339	0	0	0	984.5159173	–	–	–	–	–	–
0	0.001	0	0	**984.7224649**	–	–	–	–	–	–
0.198	0	0	0	986.533283	–	–	–	–	–	–
0.116	0	0	0	987.5358914	–	–	–	–	–	–
0.276	0	0	0	988.5496142	–	–	–	–	–	–
0	0.040	0	0	993.702465	C51H99O12P		[M+Hac-H]-	PI(O-20:0/22:1(11Z)) PI(P-20:0/22:0)	993.701272	-1.201
					C54H102O13		[M+Cl]-	DAT(18:0/22:0(2Me[S],4Me[S])) DAT(16:0/24:0(2Me[S],4Me[S]))	993.701447	-1.025
0	0.104	0	0	997.73388	–	–	–	–	–	–
0	0.011	0	0	999.7492613	–	–	–	–	–	–
0	0.020	0	0	1001.765336	–	–	–	–	–	–
0.557	0	0	0	**1003.567308**	C53H91O13P	4	[M+K-2H]-	PI	1003.56834	1.028
0.021	0	0	0	1006.597093	–	–	–	–	–	–
0	-0.040	0	0	1006.814752	–	–	–	–	–	–
0	0.291	0	0	1007.717946	–	–	–	–	–	–
0	0.312	0	0	**1011.750061**	–	–	–	–	–	–
0	0.202	0	0	**1012.753671**	–	–	–	–	–	–
0	0.196	0	0	1013.765345	–	–	–	–	–	–
0.516	0	0	0	1046.55293	–	–	–	–	–	–
0.082	0	0	0	**1056.536438**	–	–	–	–	–	–
0.299	0	0	0	1058.552618	–	–	–	–	–	–
0	0.089	0	0	1059.879464	–	–	–	–	–	–
0.345	0	0	0	1060.501563	–	–	–	–	–	–
0.207	0	0	0	1060.568244	–	–	–	–	–	–
0.076	0	0	0	1061.571089	–	–	–	–	–	–
0.383	0	0	0	1062.584864	–	–	–	–	–	–

0.074	0	0	0	1063.586414	–	–	–	–	–	–
0.238	0	0	0	1064.50374	–	–	–	–	–	–
0.038	0	0	0	1068.536417	–	–	–	–	–	–
0.533	0	0	0	1072.568806	–	–	–	–	–	–
0.447	0	0	0	1073.571797	–	–	–	–	–	–
0.300	0	0	0	1074.584798	–	–	–	–	–	–
0	0.509	-0.029	0	**1080.671003**	–	–	–	–	–	–
0.040	0	0	0	1086.585532	–	–	–	–	–	–
0.180	0	0	0	1088.601596	–	–	–	–	–	–
0.395	0	0	0	1089.604601	–	–	–	–	–	–
0.178	0	0	0	1090.61633	–	–	–	–	–	–
0.019	0	0	0	1136.55589	–	–	–	–	–	–
0.036	0	0	0	1140.557833	–	–	–	–	–	–
0.223	0	0	0	1144.589786	–	–	–	–	–	–
0	0.167	-0.068	0	**1145.59156**	–	–	–	–	–	–
0	0.162	0	0	1146.53927	–	–	–	–	–	–
0.034	0	0	0	**1147.607582**	–	–	–	–	–	–
0.470	0	0	-0.029699235	**1154.57254**	–	–	–	–	–	–
0.005	0	0	0	1166.572274	–	–	–	–	–	–
0.122	0	0	0	1168.588397	–	–	–	–	–	–
0.187	0	0	0	1170.604772	–	–	–	–	–	–

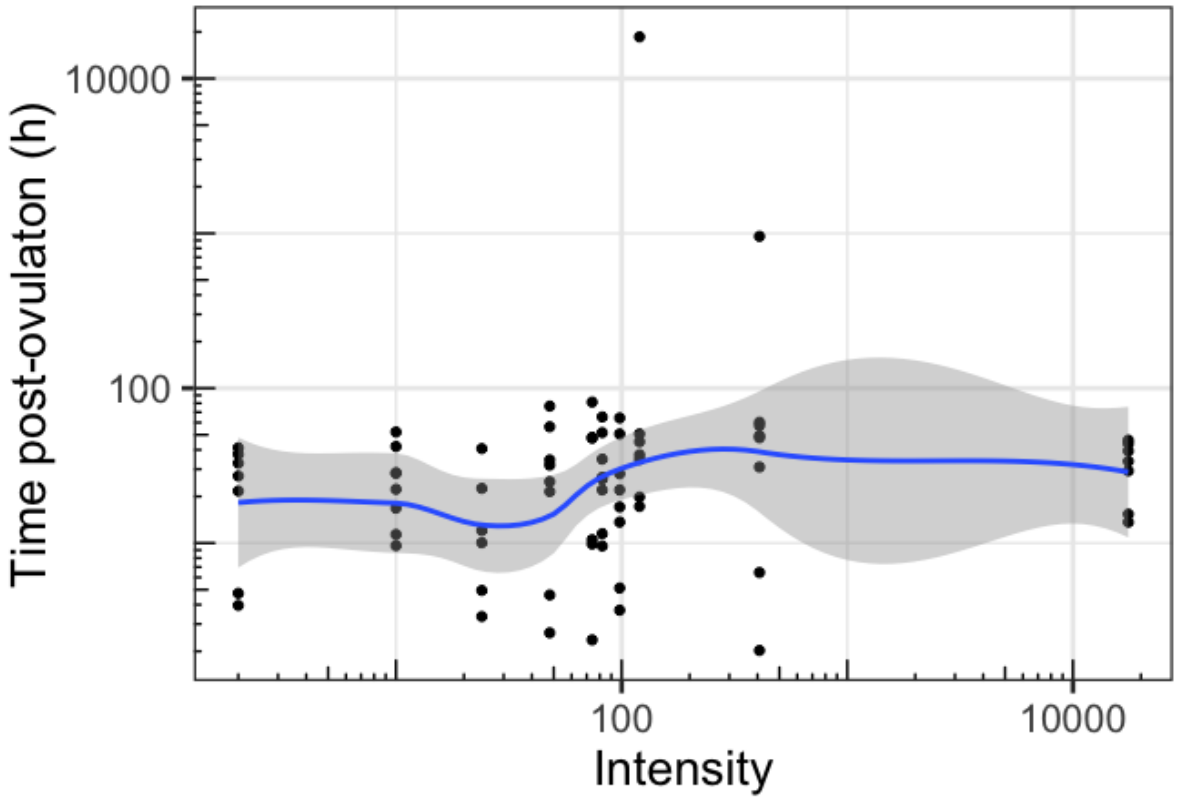


Figure 11-2 Intensity of peak (*m/z* = 1145.591560) with a non-zero loading in the second sparse principal component with high mean intensity at 120 h post-ovulation identified through sparse principal component analysis of log-normalised peak intensity matrix obtaining following the negative ion analysis of non-polar *Daphnia magna* embryo extracts collected between 2 h and 17,592 h (2 y) post-ovulation. Individual points represent biological replicates, blue trend line fitted with locally weighted scatterplot smoothing (LOESS; span = 0.5), grey shading indicates model 95% confidence interval.

11.3. Sparse principal component analysis of nESI-DIMS data from negative ion analysis of polar extracts of *Daphnia magna* embryos collected between 2 h and 17,592 h (2 y) post-ovulation

Table 11-3 Table of polar metabolite spectral peaks with non-zero loadings for sparse principal components (sPC) 1–4 from the sparse principal component analysis of nESI-DIMS data from negative ion analysis of polar extracts of *Daphnia magna* embryos collected between 2 h and 17,592 h (2 y) post-ovulation. Observed peak *m/z* is shown in addition to sPC loadings, bold values indicate peaks not detected in subitaneous embryo sample dataset. Formula indicates putative empirical formulae assigned to each peak using BEAMS, with corresponding annotation including number of possible species attributable to the aforementioned formula, respective putative compound identity with theoretical mass. Mass error (ppm) is the difference between theoretical and experimental *m/z*, normalised to experimental *m/z.* Other putative formula without compound ID assignment are indicated, for peaks with a high number of possible formulae only number are indicated for brevity.

sPC1	sPC2	sPC3	sPC4	*m/z*	Formula		Ion form	Compound identity	Theoretical mass	Mass error
0	0	0	-0.178	131.0901926	–					
0	0	0	-0.184	143.0195567	–					
0	0.031	0	0	151.0256065	–					
0	0	0	-0.098	**225.9865001**	9					
0	0	0	-0.076	230.0521507	5					
0	0	0	-0.115	230.0615796	6					
0	0	0	-0.194	**234.1170748**	3					
0	0	0	-0.238	237.9737984	16					
-0.149	0	0	0	**240.0806983**	4					
0	0	0	-0.013	**243.0361506**	12					
0	-0.008	-0.078	0	**244.0657279**	8					
0	0	0.075	0	**247.0783061**	4					
0	0	0	-0.103	251.0788399	C10H12N4O4	2	[M-H]-	Deoxyinosine Nebularine	251.07858	-1.037
						4				
0	0	0	-0.227	251.9893512	18					
0	-0.062	-0.160	0	252.1090430	C9H19NO7		[M-H]-	Choline bitartrate	252.1088776	-0.656
					4					
0	-0.263	0.449	0	**253.0709680**						
0	0	0	-0.116	256.0006465						

0	0	0.128	0	258.0475022	11					
0	0	0	-0.118	**260.1046786**	C13H15N3O3		[M-H]-	Tryptophyl-Glycine Glycyl-Tryptophan	260.104066	-2.357
					6					
0	0	0	-0.030	261.0732799						
0.036	0	0	0	263.0389091						
0	0	0	-0.003	266.0627234						
0	0	0	-0.067	**271.0705325**						
0	-0.206	0	0	**283.0821353**	C13H16O7		[M-H]-	p-Cresol glucuronide	283.082329	0.683
					C11H12O5		[M+Hac-H]-	Sinapic acid		
					C13H16O7		[M-H]-	D-Vacciniin		
					C13H16O7		[M-H]-	2-O-Benzoyl-D-glucose		
					C11H12O5		[M+Hac-H]-	cis-Sinapic acid		
					C11H12O5		[M+Hac-H]-	Elenolide		
					5					
0	-0.358	0	0	283.0882717						
0	-0.211	0	0	284.0889469						
0	-0.028	0	0	**293.5341492**						
0.007	0	0	0	**295.0843568**	C15H12N4O3		[M-H]-	4-Carboxynevirapine	295.0836646	-2.346
					11					
0	-0.052	-0.174	0	**303.0834647**						
0.074	0.012	0	-0.815	**304.0650815**						
0	-0.208	0.343	0	306.0765990						
0	0	0	-0.001	**307.0675223**	19					
0	0	0.003	0	314.0623453	21					
0	-0.185	0	0	318.0610568	C10H13N5O5		[M+Cl]-	Guanosine	318.061072	0.047
							[M+Cl]-	8-Hydroxy-deoxyguanosine		
					C11H15N5O4		[M+K-2H]-	2-O-Methyladenosine	318.061012	-0.142
							[M+K-2H]-	1-Methyladenosine		
							[M+K-2H]-	N6-Methyladenosine		
					20					

0	-0.073	0.134	0	328.0583958	C10H17N3O6S		[M+Na-2H]-	Glutathione	328.0584776	0.249
					22					
0	0	0.162	0	**328.0777974**	19					
0	-0.023	0	0	**340.0436189**						
0	0	0	-0.289	341.9906968	C9H12N3O7P		[M+K-2H]-	Cytidine 2,3-cyclic phosphate	341.9898966	-2.340
					53					
-0.171	0	0	0	**344.1164393**	21					
0	-0.240	0	0	350.0718608	38					
0	-0.204	0	0	351.0520178						
0	-0.173	0	0	352.0444784						
0	0	0	-0.044	352.0643675						
0	0	0	0.210	**353.0454938**						
0	0	0.054	0	**362.0235105**						
0	0	0	0.175	**364.0365236**						
0	-0.238	0	0	365.0910179						
0	-0.121	0	0	380.0524944						
0	-0.235	0	0	380.0616109						
-0.180	0	0	-0.073	**380.0862417**	49					
-0.142	0	0	0	**384.1150981**	C11H19NO10		[M+Hac-H]-	N-Glycolylneuraminic acid	384.1147526	-0.899
					34					
0	0	0	0.310	**393.5304275**						
0.056	0	0	0.352	**404.0191546**						
-0.197	0	0	0	**404.1372965**	38					
0	0	0	0.112	**404.5207973**						
0	0	0	0.112	**412.0021162**	108					
0	0	0	0.005	**415.0044035**	131					
0	0	0	0.339	**415.0106636**						
-0.237	0	0	0	**415.1458562**	47					
0	-0.206	0	0	421.0750644						
-0.088	0	0	0	**423.0996209**	86					
-0.126	0	0	0	**425.0612788**						
-0.130	0	0	0	425.1159017						
-0.072	0	0.064	-0.075	**431.1407704**						

0	-0.197	0	0	432.0749058						
-0.245	0	0.007	-0.023	**439.0856555**						
0	-0.120	0	0	**440.0736973**						
-0.179	0	0	0	**440.0801275**						
-0.032	0	0	0	**445.0536809**						
-0.092	0	0	0	**445.0736903**						
0	-0.273	0	0	446.0909565						
0	-0.208	0	0	447.0941763	92					
0	0	0	-0.176	453.0866019						
-0.131	0	0	0	**456.0976033**	94					
-0.142	0	-0.015	0	**456.1726082**						
-0.158	0	0	0	**457.1929716**	54					
-0.022	0	0	0	**462.0894896**						
-0.162	0	0	0	**463.1460780**						
0.039	0	0	0	**463.9997029**	183					
0.175	0.214	-0.521	0	**465.9791721**	167					
-0.163	0	0	0	**471.0763489**	157					
-0.340	0	0	0	**472.2041376**	50					
0	-0.673	0	0	**473.5922387**						
0	-0.147	-0.623	0	**474.0939921**	115					
-0.171	0	0	0	**474.1470378**	22					
-0.212	0	0	0	**475.0632387**	146					
-0.175	0	0	0	**477.0330939**	57					
0	-0.127	-0.605	0	**484.5824166**						
-0.032	0	0	0	**485.1183558**	143					
0	0.006	-0.117	0	**485.2237240**	55					
-0.677	0.067	0	0	**488.1622761**	78					
-0.102	0	0	0	**491.0870690**	173					
-0.200	0	0	0	**496.1286863**	92					
0	0	0.058	0	496.2683466						
-0.052	0.038	-0.011	0	**499.2398356**	56					
0	0	0.050	0	**522.2847155**						
-0.130	0	0.194	0	**552.1245501**						
-0.104	0	0	0	**557.0662143**						
-0.102	0	0	0	**559.0359971**						
-0.001	0	0	0	**563.1126877**						

-0.154	0	0	0	**567.1411589**						
-0.070	0	0	0	**568.1433173**						
-0.088	0	0	0	**569.0926547**						
-0.425	0	0	0	**578.1314973**						
-0.048	0	0	0	**583.1148184**						
0.218	0	0	0	**587.0297446**	C15H24N2O17P2	2	[M+Na-2H]-	Uridine diphosphate glucose	587.029698	
								Uridine diphosphategalactose		
-0.056	0	0	0	**587.1064332**	–					
0	-0.063	0	0	596.0811040	–					
-0.058	0	0	0	**597.0663458**	–					
-0.131	0	0	0	**598.3523645**	–					
0	-0.075	0	0	**604.0799317**	–					
0.088	-0.171	0	0	610.0568957	C16H25N5O15P2	4	[M+Na-2H]-	GDP-D-Rhamnose	610.056915	0.031
								ADP-glucose		
								GDP-L-fucose		
								ADP-Mannose		
0	-0.113	0	0	610.0970277	–					
0	-0.098	0	0	611.1446609	C20H32N6O12S2		[M-H]-	Oxidized glutathione	611.1446916	0.050

11.4. Sparse principal component analysis of nESI-DIMS data from positive ion analysis of polar extracts of *Daphnia magna* embryos collected between 2 h and 17,592 h (2 y) post-ovulation

Table 11-4 Table of polar metabolite spectral peaks with non-zero loadings for sparse principal components (sPC) 1–4 from the sparse principal component analysis of nESI-DIMS data from positive ion analysis of polar extracts of *Daphnia magna* embryos collected between 2 h and 17,592 h (2 y) post-ovulation. Observed peak *m/z* is shown in addition to sPC loadings, bold values indicate peaks not detected in subitaneous embryo sample dataset. Formula indicates putative empirical formulae assigned to each peak using BEAMS, with corresponding annotation including number of possible species attributable to the aforementioned formula, respective putative compound identity with theoretical mass. Mass error (ppm) is the difference between theoretical and experimental *m/z*, normalised to experimental *m/z.* Other putative formula without compound ID assignment are indicated, for peaks with a high number of possible formulae only number are indicated for brevity.

sPC1	sPC2	sPC3	sPC4	*m/z*	Formula	Ion form	Compound identity	Theoretical mass	Mass error
0	0	0.188	0	121.0843645	–	–	–	–	–
0	0	0.159	0	122.0715291	–	–	–	–	–
0	0.133	0	0	133.0319239	C5H8O2S	[M+H]+	3-Methyl sulfolene	133.0317784	-1.094
0	0.209	0	0	134.027165	C4H7NO2S	[M+H]+	–	134.0270274	-1.027
0	0	0.090	0	137.0685982	C5H10N2O	[M+Na]+	N-Mononitrosopiperazine	137.068534	-0.465
							3-Amino-2-piperidone		
0	0	0.072	0	144.0479371	C6H9NOS	[M+H]+	Raphanusamide	144.047762	-1.213
							5-Acetyl-2,3-dihydro-1,4-thiazine		
							2-Propionyl-2-thiazoline		
							(-)-5-(2-Propenyl)-2-oxazolidinethione		
							5-(2-Hydroxyethyl)-4-methylthiazole		
-0.003	0	0	0	149.0640539	C6H14P2	[M+H]+	–	149.0643524	2.002
0	0.196	0	0	**150.0586113**	–	–	–	–	–
0	0.040	0	0	**152.0570025**	–	–	–	–	–
-0.081	0	0	0	157.5378455	–	–	–	–	–
0	0	0.443	0	167.0899143	–	–	–	–	–
0	0	0.061	0	170.1292218	–	–	–	–	–
0	0	0.515	0	172.0407199	–	–	–	–	–
0	0.063	0	0	181.0950982	C7H14N2O2	[M+Na]+	–	181.0947494	-1.926

-0.276	0	0	0	**182.0339773**	C4H4N7P	[M+H]+	–	182.0338574	-0.659
					C2H7N5O3S	[M+H]+	–	182.0342384	1.434
					C6H10NO2P	[M+Na]+	–	182.0341384	0.885
-0.071	0	0	0	183.0504096	C2H8N8	[M+K]+	–	183.0503514	-0.318
					C6H14O2S2	[M+H]+	–	183.0508004	2.135
0	0	0.016	0	183.0847566	C3H6N10	[M+H]+	–	183.0849664	1.146
0	0	0.197	0	184.1446814	C9H17N3O	[M+H]+	–	184.1444384	-1.320
0	0	0.084	0	186.1240441	C8H15N3O2	[M+H]+	–	186.1237034	-1.831
0	0	0.020	0	189.0718515	–	–	–	–	–
-0.222	-0.002	0	0	**191.5408531**	–	–	–	–	–
-0.262	0	0	0	192.0556836	–	–	–	–	–
-0.310	0	0	0	**200.5460564**	–	–	–	–	–
-0.294	0	0	0	**203.563695**	–	–	–	–	–
0	0	0	0.176	207.0717666	C7H15N2OPS	[M+H]+	–	207.0715494	-1.049
					C4H8N8O	[M+Na]+	–	207.0713284	-2.116
-0.324	0	0	0	**209.0498724**	–	–	–	–	–
0	0	0.399	0	210.0506102	C10H12NPS	[M+H]+	–	210.0500854	-2.498
					C8H7N3O4	[M+H]+	–	210.0509334	1.539
-0.359	0	0	0	**211.5525371**	–	–	–	–	–
0	0	0.071	0	212.1397759	–	–	–	–	–
0	0	0.104	0	226.0457278	–	–	–	–	–
0	0.102	0	0	231.0982303	–	–	–	–	–
-0.171	-0.027	-0.115	0	**236.0553357**	–	–	–	–	–
0	0	0.126	0	245.1138267	–	–	–	–	–
-0.150	0	0.095	0	258.1105988	C13H17NO3	[M+Na]+	Pandamarilactam 3x	258.1100654	-2.067
					C6H17N7OS	[M+Na]+	–	258.1107514	0.591
					C8H20NO6P	[M+H]+	Glycerophosphocholine	258.1101034	-1.919
					C3H17N9OP2	[M+H]+	–	258.1104084	-0.738
					C10H21NO4	[M+K]+	Miglustat	258.1102184	-1.474
0	0	0.067	0	263.1395661	–	–	–	–	–
0	0	0.193	0	**265.1123345**	–	–	–	–	–
0	0	0.029	0	266.1156474	C15H17NO2	[M+Na]+	Agomelatine	266.1151504	-1.868
					5	–	–	–	–
0	0.034	-0.242	0	267.0593078	C9H12N2O6	[M+Na]+	Uridine	267.058759	-2.053
							Pseudouridine		
					6	–	–	–	–

0	0	0.080	0	269.1614356	C12H20N4O3	[M+H]+	Isoleucyl-Histidine Histidinyl-Leucine Histidinyl-Isoleucine Leucyl-Histidine	269.160817	-2.297
					2	–	–	–	–
-0.001	0	0	0	272.065862	–	–	–	–	–
0	0	0	0.004	272.1069275	5	–	–	–	–
-0.029	0	0	0	275.1515738	7	–	–	–	–
0	0	0.166	0	276.1559415	C11H21N3O5	[M+H]+	Epsilon-(gamma-Glutamyl)-lysine L-a-glutamyl-L-Lysine	276.155398	-1.967
					2	–	–	–	–
0	0.590	0	0	284.0994293	C10H13N5O5	[M+H]+	Guanosine 8-Hydroxy-deoxyguanosine	284.098946	-1.700
					4	–	–	–	–
0	0.667	0	0	306.0812489	–	–	–	–	–
0	0.205	0	0	322.0552756	C10H13N5O5	[M+K]+	Guanosine 8-Hydroxy-deoxyguanosine	322.054829	-1.385
					17		–		
0	0.665	0	0	328.0633212	–	–	–	–	–
-0.289	0	0	0	**331.6753291**	–	–	–	–	–
-0.223	0	0	0	**332.1770203**	–	–	–	–	–
0	0.407	0	0	352.0636064	–	–	–	–	–
0	0	0	0.144	355.2248204	4	–	–	–	–
-0.238	0	0	0.123	**356.1777409**	9	–	–	–	–
-0.195	0	0	0.054	356.6794062	–	–	–	–	–
-0.325	0	0	0	**360.1505037**	16	–	–	–	–
-0.073	0	0	0	**367.110097**	24	–	–	–	–
-0.115	0	0	0	370.1933764	8	–	–	–	–
-0.197	0	0	0	**370.6950165**	–	–	–	–	–
0	0	0	0.022	371.2993797	3	–	–	–	–
0	0	0	0.126	372.2420943	6	–	–	–	–
0	0.351	0	0	374.0454871	58	–	–	–	–
-0.310	0	0	0	381.0799831	C12H22O11	[M+K]+	3-b-Galactopyranosyl glucose Allolactose Alpha-Lactose	381.079374	-1.5973

							Cellobiose		
							D-Maltose		
							DEAE-cellulose		
							Epimelibiose		
							Fagopyritol A1		
							Galabiose		
							Galactinol		
							Glucinol		
							Glucose-1,3-mannose oligosaccharide		
							Inulobiose		
							Isomaltose		
							Kojibiose		
							Lactulose		
							Maltulose		
							Mannobiose		
							Melibiose		
							Neotrehalose		
							Sakebiose		
							Sucrose		
							Trehalose		
							Trehalulose		
							Turanose		
							beta-Lactose		
					39	–	–	–	–
-0.231	0	0	0	**382.083031**	37				
-0.145	0	0	0	**384.1255341**	25	–	–	–	–
0	0	0	0.849	**396.2804908**	–	–	–	–	–
0	0	0	0.041	400.2992783	3	–	–	–	–
-0.228	0	0	0	**404.1766512**	–	–	–	–	–
-0.116	0	0	0	411.0985452	53	–	–	–	–
0	0	0	0.626	411.2506512	C24H36O4	[M+Na]+	Cavipetin C	411.2505814	-0.170
					10	–	–	–	–
0	0	-0.148	0	429.1376617	–	–	–	–	–
-0.018	0	0	0	**433.0805711**	108	–	–	–	–
0	0	0.240	0	447.3585986	–	–	–	–	–

-0.141	0	0	0	**448.7433398**	22	–	–	–	–
-0.110	0	0	0	449.0666777	–	–	–	–	–
-0.130	0	0	0	**450.0700994**	75	–	–	–	–
0	0	-0.026	0	455.1243613	77	–	–	–	–
-0.074	0	-0.014	0	**471.1020158**	–	–	–	–	–
-0.258	0	0	0	**472.8123326**	–	–	–	–	–
0	0	-0.096	0	**477.1189668**	–	–	–	–	–
0	0	-0.172	0	**478.1222644**	–	–	–	–	–
0	0	0	0.176	484.4049917	–	–	–	–	–
-0.032	0	0.041	0	**485.0641071**	–	–	–	–	–
0	0.065	0.168	0	489.2957445	–	–	–	–	–
-0.025	0	0	0.115	489.3261404	–	–	–	–	–
0	0	-0.083	0	**493.0839193**	–	–	–	–	–

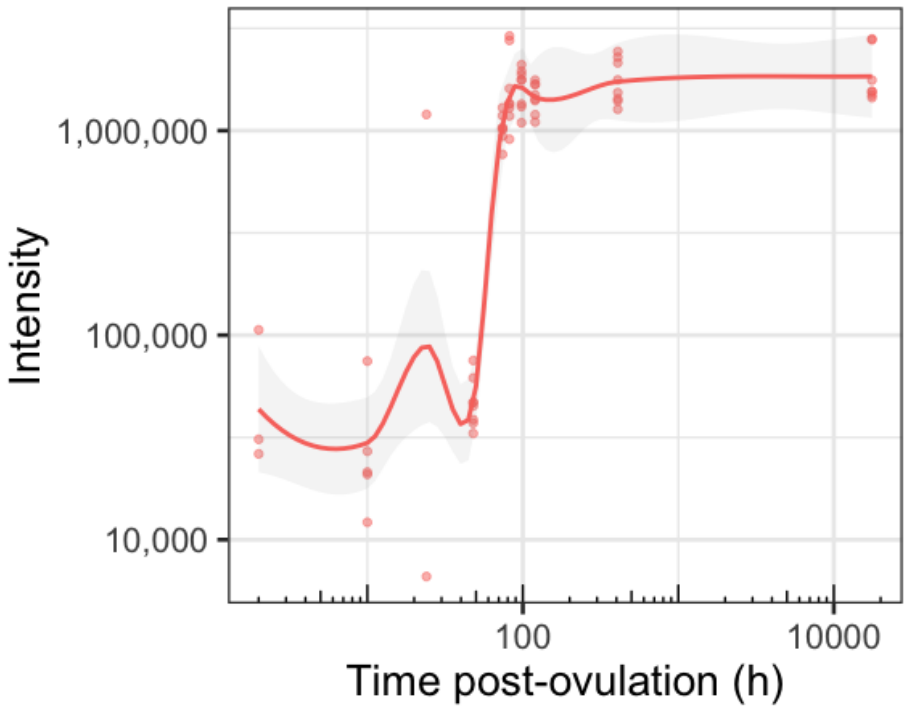


Figure 11-3 Intensity of peak (*m/z* = 370.1933764) with a none-zero loading in the first sparse principal component with high mean intensity at 24 h post-ovulation identified through sparse principal component analysis of log-normalised peak intensity matrix obtaining follow the positive ion analysis of polar *Daphnia magna* embryo extracts collected between 2 h and 17592 h (2 y) post-ovulation. Individual points represent biological replicates, lines fitted with locally weighted scatterplot smoothing (LOESS; span = 0.5), grey shading indicates model 95% confidence interval.

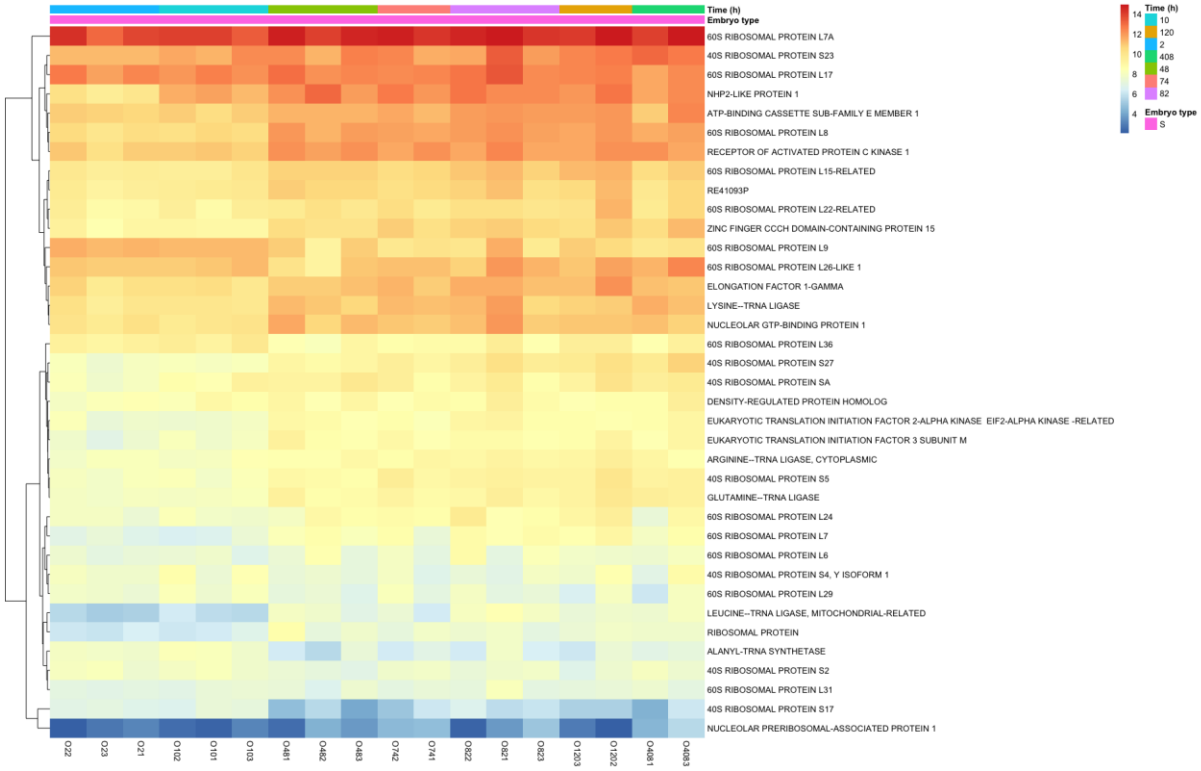


Figure 11-4 Heat map and hierarchical clustering of genes relating to protein synthesis in samples of diapause-destined *Daphnia* embryos collected between 2 h and 2 weeks (408 h) post-ovulation. Row labels indicate corresponding annotation based on OrthoDB orthologous gene family group and coloured column top panels indicate time post-ovulation (refer to legend). Colours indicate low expression (blue) and high expression (red) on log-normalised scale**.**