Environmental transcriptomics under heat stress: Can environmental RNA reveal changes in gene expression of aquatic organisms?

Robert M. Hechler*[a, b], Matthew C. Yates[c, d], Frédéric J. J. Chain[e] and Melania E. Cristescu[a]

Author affiliations:
[a] Department of Biology, McGill University, Montréal, QC, H3A 1B1, Canada

[b] Current affiliation: Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON, M5S 3B2, Canada

[c] Département des Sciences Biologiques, Université du Québec à Montréal, Montréal, QC, H2X 1Y4, Canada

[d] Current affiliation: Great Lakes Institute for Environmental Research, University of Windsor, Windsor, ON, N9C 1A2, Canada

[e] Department of Biological Sciences, University of Massachusetts Lowell, Lowell, MA, 01854, USA

*Corresponding author
Email: robert.hechler@mail.utoronto.ca

Author contributions:
RMH and MEC designed the study. RMH and MCY conducted the experiment. RMH conducted all molecular work. RMH and FJJC performed the bioinformatic analysis and produced the figures. RMH wrote the first draft of the manuscript and all authors contributed to editing the manuscript.

Keywords:
environmental transcriptomics; environmental RNA (eRNA); environmental DNA (eDNA); extra-organismal RNA; gene expression; biological monitoring; biomonitoring

Abstract: To safeguard biodiversity in a changing climate, we require taxonomic information about species turnover and insights into the health of organisms. Environmental DNA approaches are increasingly used for species identification, but cannot provide functional insights. Transcriptomic methods reveal the physiological states of macroorganisms, but are currently species specific and require tissue sampling or animal sacrifice, making community-wide assessments challenging. Here, we test if broad functional information (activity level of the transcribed genes) can be harnessed from environmental RNA (eRNA), which includes extra-organismal RNA from macroorganisms along with whole microorganisms. We exposed *Daphnia pulex* as well as phytoplankton prey and microorganism colonizers to control (20 °C) and heat stress (28 °C) conditions for seven days. We sequenced eRNA from tank water (after complete removal of *Daphnia*) as well as RNA from *Daphnia* tissue, enabling comparisons of extra-organismal and organismal RNA based gene expression profiles. Both RNA types detected similar gene expression responses. Using eRNA, we identified 42 *Daphnia* genes to be differentially expressed following heat stress. Of these, 62% were also differentially expressed and exhibited similar levels of relative expression in organismal RNA. Both RNA types recovered similar *Daphnia* heat stress responses via gene ontology terms. In addition to the extra-organismal *Daphnia* response, eRNA detected community-wide heat stress responses consisting of 199 differentially expressed genes across 9 taxa, and distinct functional profiles via KEGG orthologs. Our study demonstrates that environmental transcriptomics based on eRNA can non-invasively reveal gene expression responses of macroorganisms following environmental changes, with broad potential implications for the biomonitoring of ecological health across the trophic chain.

Introduction

In an era of ecological crises, monitoring the presence and health of organisms within an ecosystem is essential for the preservation of biodiversity. Environmental DNA is routinely used for biomonitoring, but such applications are limited to species detection [1,2]. The RNA molecule provides an additional layer of functional information, as transcribed genes reflect an organism's physiological status [3]. Thus, transcriptomics is used to infer the ecological health of organisms, but such surveys are species-specific and dependent on sampling organismal RNA (oRNA) directly, which is labor-intensive and requires tissue collection or animal sacrifice [4–9]. Non-invasive environmental RNA (eRNA) approaches would overcome these limitations by providing insights into the health of populations and communities [10–13]. However, using eRNA to inform on ecological health only been theorized (see [13]) and remains yet to be empirically tested.

We broadly define eRNA as RNA extracted from the environment, including extra-organismal RNA released by macroorganisms as well as whole microorganisms. Microbiologists have long extracted RNA from bulk samples to profile gene expression in microbial communities, referred to as metatranscriptomics [14–16]. Despite more than two decades of research, metatranscriptomics remains limited to studying the gene expression of microorganisms captured in bulk samples rich in oRNA. The general assumption has been that the eRNA released by macroorganisms into the environment is labile and degrades too rapidly to be reliably detected or quantified (but see [11]). However, recent work has demonstrated that macroorganisms do shed and/or excrete RNA into their surrounding environment and that this eRNA can be successfully extracted and quantified [17–21]. Despite this evidence supporting the

78 extractability of eRNA, environmental transcriptomics, defined here as eRNA based gene
79 expression profiling, remained untested.

80 In aquatic systems, environmental conditions modulate biodiversity and are reflected in
81 changes of metatranscriptomes [14,22–24]. Temperature is a key abiotic factor that influences
82 the physiology and fitness of ectotherms due to the relationship between body and environmental
83 temperatures [25–27]. Rapid increases in temperature, such as heatwaves, can exceed an aquatic
84 ectotherms thermal limits and exponentially increase its standard metabolic rate, resulting in
85 decreases in fitness and survival [26–29]. Transcriptomics is widely used to study the heat stress
86 response of many taxa, including fish [7,30–33], coral holobionts [34–36], molluscs [6,37], and
87 copepods [38,39]. *Daphnia pulex*, a key bioindicator species, exhibit widespread downregulation
88 of genes involved in cellular and metabolic processes under heat stress [40,41]. Clearly, gene
89 expression surveys can provide species-specific functional insights important for conserving
90 aquatic biodiversity [42–44]. However, a species by species analysis based on oRNA cannot
91 capture community-wide changes without extensive efforts. Mapping eRNA reads to the
92 reference genomes of multiple target species and using existing metatranscriptomic pipelines,
93 could enable environmental transcriptomics to revolutionize biological monitoring by revealing
94 gene expression responses across the trophic chain.

95 We constructed simple mock freshwater communities containing *Daphnia pulex,* as well
96 as three phytoplankton species and opportunistic microorganisms that colonized the artificial
97 lake water. The communities were exposed to control (20 °C) and near-lethal *Daphnia* heat
98 stress (28 °C) conditions for seven days. To enable comparisons between eRNA and oRNA
99 based gene expression profiles, we sequenced eRNA from tank water (after complete *Daphnia*
100 removal) and oRNA from *Daphnia* tissue. We conducted differential gene expression analyses

101 between the control and heat stress conditions and compared the *Daphnia* functional stress
102 response identified in eRNA and oRNA. In addition to the *Daphnia* analysis, we used eRNA to
103 investigate the community-wide heat stress response. We hypothesized that 1) eRNA contains a
104 subset of the differentially expressed *Daphnia* genes observed in oRNA after exposure to heat
105 stress and 2) eRNA captures transcriptional responses to heat stress. Based on our findings, we
106 discuss the potential of eRNA based transcriptomics for the biological monitoring of ecological
107 communities and also provide recommendations to overcome potential limitations.
108

109 **Results**

110 Daphnia pulex *gene detection and expression profiles*

111 Since we pre-filtered tank water at 60 µm to remove *Daphnia*, all *Daphnia* reads within
112 environmental RNA (eRNA) samples are of extra-organismal origin. Sequencing of eRNA from
113 tank water and organismal RNA (oRNA) from *Daphnia* tissue detected 3,941 (21%) and 17,586
114 (95%) genes (in at least 3 of 4 replicates), respectively, of the 18,449 genes that compose the
115 *Daphnia pulex* reference genome [45]. All *Daphnia* genes detected in eRNA were also detected
116 in oRNA. A principal component analysis of the regularized log normalized expression for the
117 300 most variant genes showed separation between experimental groups (Fig. 1). The samples
118 were separated across the first principal component (PC1) by RNA type and across the second
119 principal component (PC2) by temperature, clustering eRNA, oRNA, 20 °C and 28 °C within
120 their respective groups. The only exception was 20 °C eRNA sample T1, which unexpectedly
121 clustered closely with the 28 °C eRNA group. PC1 and PC2 accounted for 63% of explained
122 variance of all principal components (Fig. S1).

123

124 Daphnia pulex *differential gene expression*

125 Comparison of gene expression profiles at 20 °C and 28 °C identified 42 significantly
126 differentially expressed genes (DEG; FDR-adjusted p-value <0.1) from eRNA (8 upregulated
127 and 34 downregulated) and 3108 DEGs from oRNA samples (1470 upregulated and 1638
128 downregulated). All eRNA significant DEGs were highly up and down regulated (exceeding
129 Log_2 2 fold change), as well as 46 and 226 oRNA DEGs, respectively. Of the 93% of eRNA
130 DEGs that exhibited the same directional change (up/down regulated) as in oRNA (Table S1),
131 62% (5 upregulated and 21 downregulated) were also differentially expressed in oRNA (Fig.
132 S2). A heatmap comparing the relative expression (Z-score) of the 26 *D. pulex* DEGs common to
133 eRNA and oRNA revealed four distinct clusters based on hierarchical clustering analyses,
134 separating samples and genes by temperature and directional change, respectively (Fig. 2).
135 Across both temperature conditions, eRNA and oRNA clustered together and were not separated
136 into RNA types, indicating that eRNA and oRNA exhibit similar levels of relative gene
137 expression for these DEGs. The only exception was eRNA 20 °C T1 sample which, similar to
138 the PCA, unexpectedly clustered with the 28 °C samples.

139

140 Daphnia pulex *functional annotation*

141 Gene ontology (GO) enrichment analysis of eRNA detected genes identified 57 GO terms
142 as enriched (FDR-corrected p-value <0.05), with 33, 6, and 18 belonging to Biological Processes
143 (BP), Cellular Component (CC) and Molecular Function (MF) domains, respectively (Figs. 3 and
144 S3). Of the 42 significant *D. pulex* DEGs identified from eRNA, 27 had GO annotations from the
145 reference genome [45] (Fig. S4-5). GO enrichment analysis of eRNA DEGs identified chitin

146 metabolic process as enriched (FDR-corrected p-value = 0.016) between expected and observed
147 counts (Fig. 3).
148 GO enrichment analysis of oRNA detected genes identified 119 GO terms as enriched
149 (FDR-corrected p-value <0.05), with 100, 1, and 18 belonging to BP, CC and MF domains,
150 respectively (Figs. 3 and S3). Of the 3108 DEGs identified from oRNA, 2259 had GO
151 annotations from the reference genome [45] (Figs. S4-5). GO enrichment analysis of oRNA
152 DEGs identified 17 GO terms as enriched (FDR-corrected p-value <0.05), including chitin
153 metabolic process, proteolysis, negative regulation of hydrolase activity, negative regulation of
154 proteolysis and response to stimulus (Fig S3). Of the 26 genes found to be commonly DE in both
155 eRNA and oRNA, the GO terms chitin metabolic process and chitin binding were enriched
156 (FDR-corrected p-value <0.05; Fig. 3).
157
158 *Community composition, differential gene expression and functional annotation*
159 eRNA reads originated from a simple mock freshwater community that we constructed,
160 consisting of *Daphnia pulex* as well as three phytoplankton species and opportunistic
161 microorganisms that colonized the artificial lake water. All non-*Daphnia* eRNA reads likely
162 originate from whole microorganisms that were captured on the filter. We used the SqueezeMeta
163 pipeline [46] to align eRNA reads to the GenBank nr database using its LCA algorithm, which
164 identified 434 taxa. The community composition and relative abundances remained similar
165 between 20 °C and 28 °C conditions (Fig. S6). Across eRNA samples, an average of 38% of all
166 reads were aligned to nine species including four eukaryotes that were a priori known to persist
167 in the communities and five highly abundant taxa identified by the SqueezeMeta analysis (Table
168 S2). We detected 10,245 genes from these nine species, with the majority originating from *D.*

169 *pulex, Scenedesmus quadricauda and Ankistrodesmus falcatus,* with 3941, 3223 and 2123 genes,
170 respectively (Table S3). Differential gene expression analysis between 20 °C and 28 °C eRNA
171 samples was performed for each of the nine species and identified a total of 199 significantly
172 differentially expressed genes (Table S3). The most DEGs were identified for *A. falcatus*, *S.*
173 *quadricauda* and *D. pulex*, with 72, 47 and 42 DEGs respectively.
174 Within all eRNA reads, 12,210 KEGG IDs were identified and a non-metric
175 multidimensional scaling plot of these KEGG IDs revealed distinct functional profiles between
176 20 °C and 28 °C communities (Fig. S7). Of all KEGG IDs, 127 were identified as significantly
177 differentially expressed (DE; false discovery rate-adjusted p-value <0.1) between 20 °C and 28
178 °C (Table S4). Many of these DE KEGG IDs correspond to proteins and pathways involved in
179 stress responses, including metabolic pathway, p53 signaling pathway, ubiquitin mediated
180 proteolysis and cyclin-dependent kinase 4 (Table S4).
181
182 **Discussion**
183 The use of environmental RNA (eRNA) based transcriptomics remained previously
184 untested despite recent evidence demonstrating that extra-organismal eRNA persists in the
185 environment long enough to be extracted [17–21]. Our findings revealed that environmental
186 transcriptomics is sensitive in detecting transcriptional heat stress responses even without
187 sampling source organisms directly. From eRNA released by *Daphnia pulex* into the tank water,
188 we detected thousands of *Daphnia* genes and identified a subset of heat stress relevant genes to
189 be differentially expressed (DE) between the two temperature conditions, with levels of relative
190 expression similar to organismal RNA (oRNA). We also detected community wide changes in
191 functional profiles. Our study demonstrates the potential of environmental transcriptomics to

192 reveal the functional response of biotic communities to environmental stress, with broad
193 implications for biomonitoring ecological health across the trophic chain.
194
195 *Recovery of* Daphnia pulex *extra-organismal eRNA*
196 Several studies have recently demonstrated robust detection and persistence of extra-
197 organismal eRNA in the environment [17,19,21]. Our results are consistent with these findings
198 as we recovered substantial quantities of *Daphnia* eRNA, demonstrated by an average of 0.51%
199 of all reads mapping to the *D. pulex* reference genome (Table S2). This relatively high recovery
200 of *Daphnia* reads was particularly notable, given the general expectation that within
201 environmental samples, potentially low-abundance eRNA is 'competing' to be sequenced with
202 larger quantities of oRNA from whole microorganisms. This encouraging finding suggests that
203 most of the eRNA captured is in a protected state (e.g., cellular, vesicular, or in stable molecular
204 complexes) while free RNA forms, prone to degradation, are very short lived and potentially
205 undetectable. Our recovery of extra-organismal eRNA reads (0.51%) was similar to
206 metagenomic studies of natural aquatic ecosystems that found between 0.27% and 1.25% of
207 environmental DNA (eDNA) reads to be likely of extra-organismal origin [47–50].
208
209 Daphnia pulex *extra-organismal eRNA differential gene expression analysis*
210 Our broad environmental transcriptomics approach was able to detect 3,941 *D. pulex*
211 genes from eRNA, building upon targeted RT-PCR of eRNA which detected 6 fish genes [20].
212 Of the genes we detected, 42 were found to be significantly differentially expressed genes
213 (DEGs) under temperature stress. As expected, we identified many more (3,108) DEGs from
214 *Daphnia* oRNA. This discrepancy in DEG identification can be easily attributed to large

215 differences in the number of reads that mapped to the reference genome. On average, 1 and 200
216 million reads mapped to the *D. pulex* reference genome using eRNA and oRNA, respectively
217 (Table S2). Having less than 1% of eRNA reads matching *D. pulex* limited our power to detect
218 DEGs, but is unsurprising given that the libraries were generated from diverse environmental
219 samples which included RNA transcripts from over 400 taxa, whereas the oRNA libraries were
220 prepared from concentrated *D. pulex* tissue samples. On the other hand, we believe this diversity
221 of eRNA within water samples to be beneficial as it opens up exciting possibilities of conducting
222 community-wide surveys that encompass macro- and micro-organisms across the trophic chain.
223 For our proof-of-concept study, we performed deep sequencing (average of 158 million reads per
224 eRNA sample, Table S2) in anticipation of such diverse eRNA samples, enabling sufficient
225 recovery of *D. pulex* eRNA for gene expression analysis. Despite the lower abundance of *D.*
226 *pulex* reads from eRNA than oRNA, 93% of eRNA DEGs exhibited the same directional change
227 (up/down regulated) as in oRNA (Table S1). Additionally, the majority of eRNA DEGs (62%)
228 were also differentially expressed in oRNA and these commonly shared DEGs exhibited similar
229 levels of relative expression in both RNA types (Fig 2). These results demonstrate that similar to
230 oRNA, our environmental transcriptomics approach yielded adequate sensitivity in detecting
231 changes in gene expression of a source organism in response to heat stress.
232 Changes in gene expression are often observed before phenotypic and demographic
233 effects occur and may therefore act as an early indication of ecological stress [51–53]. We found
234 that eRNA and oRNA recovered similar *D. pulex* chronic heat stress responses which consisted
235 of extensive downregulation of genes associated with cellular and metabolic processes,
236 membrane, extracellular region, collagen and cuticle structure (Figs. 3 and S4). These results are
237 consistent with studies that also found chronically heat stressed *Daphnia* to exhibit widespread

238 downregulation of genes with similar functions [40,41]. This widespread downregulation of
239 genes has been hypothesized to be a molecular compensatory mechanism for *Daphnia* at near-
240 lethal temperatures to sacrifice their long-term fitness for immediate survival [41].

241 After seven days of heat exposure, we found few heat shock genes (4 of 67) to be
242 significantly differentially expressed but only with slight upregulation (log_2 fold change of 1.1-
243 1.3; Table S5), from *Daphnia* oRNA. This corroborates findings that indicate heat shock genes
244 to be central for *Daphnia* in mitigating cell damage following acute (hours) heat exposure, but
245 less important for the chronic (days) heat stress response [40,41]. Multiple tropical marine fish
246 species similarly do not differentially express heat shock genes during chronic heat stress [31],
247 suggesting that heat shock genes might not be optimal for assessments of communities under
248 chronic heat stress. Furthermore, we detected chitin related gene ontology (GO) terms to be
249 commonly enriched in both eRNA and oRNA DEGs (Fig. 3), corroborating *Daphnia* studies that
250 linked chitin genes to environmental stressors [40,54–56]. Hydrolytic and proteolytic related GO
251 terms were enriched in eRNA (Fig. 3), and are commonly involved in *Daphnia* and copepod heat
252 stress response by stabilizing and cleaning up misfolded proteins [38,40]. It is promising that
253 using eRNA, we were able to detect thousands of genes and identify a subset as differentially
254 expressed with similar levels of relative expression as oRNA, and detect the widespread
255 downregulation of genes associated with chronic heat stress. We also found that many GO terms
256 were commonly shared in both RNA types, and that stress related terms were enriched in eRNA.
257 Collectively, our results indicate that environmental transcriptomics based on eRNA is sensitive
258 in detecting heat stress responses of source organisms through changes in gene expression and
259 functional GO terms. Future empirical work is necessary to determine how our results extend to
260 more complex natural ecosystems.

261

262 *Community composition and function*

263 Since we pre-filtered the tank water at 60 μm to remove *Daphnia*, all non-*Daphnia* RNA
264 recovered in our eRNA samples likely originates from whole microorganisms that passed the
265 filter. After mapping all eRNA reads to GenBank nr, we found the community composition to
266 remain largely unchanged for 20 °C and 28 °C conditions (Fig. S6). Within our eRNA samples,
267 gene expression analysis of nine species (across five phyla) detected 10,245 genes with 199 of
268 these being significantly differentially expressed (DE) between 20 °C and 28 °C (Table S3).
269 Although community compositions remained similar, analysis of KEGG orthologs revealed
270 distinct functional profiles for communities in the two temperature treatments (Fig. S7). We
271 found many stress-related KEGG orthologs to be DE, including proteins involved in binding,
272 kinase oxidoreductase, and cellular senescence, a pathway indicative of cellular arrest and DNA
273 damage (Table S4). Our community wide results reflect similar patterns of differential gene
274 expression and shifts in functional terms observed in conventional aquatic metatranscriptomic
275 studies based on bulk microorganism samples [14–16,22,57–60]. For example, a freshwater
276 metatranscriptome likely under heat stress exhibited similar stress responses which included
277 oxidoreductase, binding and kinase proteins [61]. We build upon this rich metatranscriptomics
278 literature by demonstrating that in addition to yielding microorganism insights, environmental
279 transcriptomics can also non-invasively detect the gene expression response of macroorganisms
280 to environmental stress.

281

282 *Recommendations and future directions*

283 Interest in non-lethal transcriptomics is growing, as such an approach increases animal
284 welfare, allows for repeated survey of individuals, and linking gene expression to fitness
285 following exposure to stress [8,62,63]. However, current non-lethal methods are species specific
286 and involve animal handling, and sampling for blood or tissue, which may still result in organism
287 stress and reduced survival upon release into the environment [64–66]. Inferring the
288 physiological state of organisms from non-invasive eRNA samples would represent a substantial
289 improvement for animal welfare as we could circumvent these stressful sampling procedures.
290 We demonstrated that environmental transcriptomics based on eRNA can non-invasively detect
291 the heat stress response of progenitor organisms, with similar sensitivity to that of the
292 conventional and invasive oRNA approach.

293 Perhaps the most exciting prospect of environmental transcriptomics is the potential to
294 monitor the ecological health of entire communities and thereby gain insights into ecosystem
295 level conditions. Biomonitoring traditionally relies upon identifying bioindicator taxa that are
296 correlated with environmental conditions, but transcriptomic surveys can provide valuable early
297 warnings since gene expression changes are likely to occur before compositional shifts could be
298 detected [12]. Conventional metatranscriptomics has demonstrated that gene expression profiling
299 of microorganisms reflects environmental conditions [22,58,67,68]. Environmental
300 transcriptomics based on eRNA extends beyond microorganisms by revealing expression signals
301 from extra-organismal eRNA, thus providing functional information of species across the trophic
302 chain. Such an approach could enable researchers to correlate functional pathways expressed by
303 community assemblages to ecosystem level conditions [13]. Gaining insights into the ecological

304 health of populations and communities, as well as ecosystem level conditions, will enable
305 conservation initiatives to be proactive in their efforts and respond before local extinction events
306 occur.

307 We anticipate multiple challenges for environmental transcriptomic biomonitoring of
308 natural ecosystems, such as achieving adequate sequencing depth of eRNA in a cost-efficient
309 manner (Table 1). However, sequencing technologies are rapidly advancing and the costs
310 continue to decrease [69]. The continuation of such rapid technological advancement will enable
311 future studies to generate more reads less expensively, thereby increasing the sensitivity of
312 environmental transcriptomics to detect extra-organismal eRNA transcripts. Additionally, the
313 continued development of low input mRNA-seq library preparation kits will enable researchers
314 to prepare libraries for environmental transcriptomics with minimal template eRNA. Molecular
315 technological advancements that enhance extraction efficiency, enrich eukaryotic RNA and
316 remove ribosomal RNA from diverse samples will increase the yield of coding eRNA and thus
317 the power of environmental transcriptomics.

318 While sequencing and molecular advancements will resolve challenges associated with
319 eRNA depth, perhaps the largest remaining challenge for environmental transcriptomics to
320 overcome is the lack of a baseline to compare against. Differential gene expression analyses are
321 conducted by comparing a treatment (i.e. stressed community) against a control 'baseline' (i.e.
322 healthy community), to identify genes involved with the stress response. However, in natural
323 ecosystems, it is difficult to obtain a control baseline due to population demographics and
324 multiple stressors acting simultaneously on a system. This is similarly shared by
325 metatranscriptomics, and multiple statistical and mathematical modelling solutions have been
326 developed to overcome this problem [70–72]. Additionally, we propose supervised machine

327 learning (SML) models as a baseline-free solution for assessing community health using
328 environmental transcriptomics. SML models could be trained with existing datasets
329 (transcriptomics, metatranscriptomics and metagenomic) that have identified genes, functional
330 pathways and taxa associated with varying degrees of ecological quality [73]. Once trained,
331 environmental transcriptomic data could be used as input and the SML model would predict the
332 stressor most likely at play and the ecological health status of the community.

333 By providing functional insights in a non-invasive manner, eRNA based approaches
334 could push beyond the limitations of eDNA species detection and oRNA invasive transcriptomic
335 surveys. We demonstrated, for the first time, that eRNA can be used to detect thousands of
336 genes, and that environmental transcriptomics is sensitive enough to identify a subset of
337 functionally relevant genes as differentially expressed under heat stress, with relative expression
338 levels similar to the traditional oRNA approach. Additionally, within all eRNA reads, we were
339 able to reveal changes in functional profiles for the community in response to heat stress.
340 Collectively, these findings demonstrate the potential of environmental transcriptomics to
341 revolutionize biological monitoring by providing functional information across the trophic chain
342 and insights into the health of ecological communities. Our study represents a significant
343 advancement for the non-invasive biomonitoring of ecological community health, but future
344 empirical work is necessary to determine if environmental transcriptomics can be extended to
345 natural ecosystems.
346

347 **Materials and Methods**

348 *Experimental design*

349 We constructed simple mock freshwater communities composed of *Daphnia pulex,* three
350 algae species (*Ankistrodesmus falcatus, Scenedesmus quadricauda,* and *Pseudokirchneriella*
351 *subcapitata*) and microorganisms that colonized the artificial lake water. The communities were
352 exposed to 20 °C (control) and 28 °C (neat-lethal stress for Daphnia) temperatures for seven days
353 to mimic natural heatwave conditions and allow for eRNA accumulation in tank water.
354 Communities were reared in 8 L tanks containing artificial lake water [74] and were seeded with
355 a starting population of 50 parthenogenetic females of *Daphnia pulex* (clone from Illinois, USA;
356 40.24, -87.78). To generate dense *Daphnia* populations, tanks were kept in controlled chambers
357 for 45 days at optimal growing conditions: 20 °C, relative humidity of 50% and a photoperiod of
358 16:8 hour light:dark. *Daphnia* were fed 1.5 mL algae (1:1:1 ratio of *A. falcatus, S. quadricauda,*
359 and *P. subcapitata*) twice per week. After 45 days, the *Daphnia* populations in all tanks reached
360 a minimum of 500 individuals and were transferred to controlled experimental chambers pre-set
361 to the respective temperatures (20 °C and 28 °C).
362
363 *Sample collection*
364 Each temperature treatment had four tanks (biological replicates), and we collected four
365 eRNA and oRNA technical replicates per tank. We stirred the tanks for 60 seconds and filtered
366 500 mL of water through a 60 µm mesh to remove all *Daphnia* and ephippia. To capture eRNA,
367 the *Daphnia* free water was then filtered through a 0.7 µm glass microfiber filter. Each filter was
368 cut in half and placed in individual 1.5 mL microcentrifuge tubes containing 370 µL RLT buffer
369 (Qiagen) and 3.7 µL β-mercaptoethanol. Tubes were immediately stored at -80 °C until eRNA
370 extraction. Negative eRNA filtration control samples (500 mL of distilled water) were collected
371 at the beginning, middle and end of the sampling day, using the same methods. For each oRNA

372 sample, six live organisms (3 adults and 3 juveniles) per tank were haphazardly collected using a
373 sterile pipette and transferred to a 1.5 mL microcentrifuge tube containing 370 µL RLT buffer
374 (Qiagen) and 3.7 µL β-mercaptoethanol. To maintain the *Daphnia* heat stress transcriptional
375 signal, tubes were flash frozen in liquid nitrogen for 2 seconds, and immediately stored at -80 °C
376 until RNA extraction. The following system was used for sample identification:
377 RNAtype_Temperature&TankID (e.g. sample "eRNA_20T1" refers to eRNA collected at 20 °C
378 from tank 1).
379
380 *Extractions of eRNA and oRNA samples*
381 Samples of eRNA and oRNA were extracted using the RNeasy Mini Kit (Qiagen),
382 following the manufacturer's protocol with the following modifications. eRNA samples were
383 thawed on ice, gently vortexed for 5 seconds and then centrifuged for 3 minutes at 13,300 rpm to
384 separate the liquid from filter. To increase the eRNA yield, liquid from two technical replicates
385 was transferred into a 5 mL microcentrifuge tube containing 800 µL 70% EtOh and pipette
386 mixed. This mixture was transferred to an RNeasy spin column (700 µL at a time) and
387 centrifuged at 10,000 rpm for 15 seconds. We repeated this four times, until all remaining liquid
388 from the 5 mL microcentrifuge tube was transferred to the spin column. For oRNA, *Daphnia*
389 samples were thawed on ice and homogenized via sterile pestle and mortar mixer for 60 seconds.
390 We then added 400 µL of 70% EtOh to the homogenized tissue. The resulting mixture was
391 pipette mixed and transferred to an RNeasy spin column, where it was centrifuged at 10,000 rpm
392 for 15 seconds. The remainder of the eRNA and oRNA extraction procedures followed the
393 manufacturer's instructions, with RNA being eluted in 40 µL of DNase and RNase free
394 molecular grade water.

395

396 *DNA digestion*

397 Immediately following RNA extraction, an 18 µL aliquot of both RNA types underwent
398 two rounds of DNA digestion using the DNA-*free*™ DNA Removal Kit (Invitrogen) which
399 provided rDNase. We completed the first round of DNA digestion following the manufacturer's
400 instructions, but skipped the DNase inaction step. To ensure no DNA carryover, we completed a
401 second round of DNA digestion as follows: 2.7 µL of DNase I Buffer and 1 µL rDNase were
402 added and pipette mixed into each sample and the plate was gently vortexed, centrifuged, and
403 then incubated at 37 °C for 20 minutes. Following incubation, 2 µL of DNase Inactivation
404 Reagent was added to each well. The plate was incubated for 2 minutes at room temperature, and
405 centrifuged at 3700 rpm for 5 minutes to pellet the DNase Inactivation Reagent. The supernatant
406 containing RNA was transferred to individual sterile microcentrifuge tubes and stored at -80 °C
407 until library preparation.

408

409 *RNA-seq library preparation and sequencing*

410 We prepared libraries for whole transcriptome shotgun sequencing (RNA-seq) using the
411 Illumina Stranded Total RNA Prep with Ribo-Zero Plus kit and IDT® for Illumina® RNA UD
412 Indexes Set A, following the manufacturer's instructions. Equal volumes of RNA (post 2 rounds
413 of DNA digestion) from each technical replicate were pooled and used as input material (250 ng
414 and 26.4-65.56 ng for oRNA and eRNA, respectively) to prepare one biological replicate library
415 per tank. During library amplification, 12 and 17 PCR cycles were conducted for oRNA and
416 eRNA, respectively, to account for differences in starting material. Library quantification,

417 quality control and equimolar pool sequencing of the 16 libraries was conducted at the McGill
418 Genome Centre on one Illumina NovaSeq 6000 S4 lane using paired-end 100 bp reads.
419
420 *Contamination prevention*
421 We submerged all experimental and sampling equipment in a 20% bleach solution for 15
422 minutes and rinsed thoroughly five times with distilled water prior to use. All eRNA samples
423 were processed in a pre-PCR clean laboratory dedicated to low quality environmental and
424 ancient DNA/RNA samples. Prior to entering the pre-PCR sterile laboratory, researchers entered
425 a decontamination room to change into the dedicated clean lab coats and put on face masks,
426 hairnets and lab clogs with shoe covers. The laboratory workbench was soaked in a 20% bleach
427 solution for 10 minutes and RNase WiPER™ (iNtRON Biotechnology) and 20% bleach was
428 used to thoroughly wipe pipettes, vortex mixers and centrifuges. Aerosol filter barrier pipette tips
429 were used to prevent cross-contamination of samples. Filtration negative control samples and
430 molecular negative control samples (consisting of reagents or DNA/RNA free molecular grade
431 water) were processed at each step alongside eRNA samples, and were free of contaminating
432 nucleic acids as identified by failed PCR amplification using universal COI primers known to
433 amplify *Daphnia* DNA [75]. Successful DNA digestion of each eRNA and oRNA sample was
434 verified by failed PCR amplification of RNA post-DNA digestion x2 using universal COI
435 primers known to amplify *Daphnia* DNA [75]. RNA-seq libraries were prepared from filtration
436 and molecular negative control samples and were free of contamination as indicated by quality
437 control checks and failed amplification conducted at the McGill Genome Centre.
438
439 *Bioinformatic pipeline and statistical approaches*

440 Raw FASTQ files underwent initial sequencing quality inspection using FastQC [76].
441 Low quality sequences and adapters were removed using Trimmomatic [77] (ILLUMINACLIP:
442 lluminaStrandedTotalRNA_adapter.fa:2:30:15 TRAILING:30 HEADCROP:1 MINLEN:90). All
443 statistical analyses were conducted using R (R Core Team, 2021, R version 4.1.2). With default
444 parameters, STAR [78] was used to map all eRNA and oRNA reads that passed quality control
445 and trimming to the *Daphnia pulex* reference genome [45]. Using Kallisto [79], we also mapped
446 eRNA reads to the reference genomes and transcriptomes of species known a priori to persist in
447 the communities and highly abundant species identified by alignment of eRNA reads to NCBI nr
448 database [80–87]. FeatureCounts [88] was used on the STAR resulting BAM file to quantify
449 gene counts. Only those genes with at least one count in three out of four eRNA 20 or 28 °C
450 samples were kept and considered in the subsequent differential expression (DE) analysis.
451 DE analysis was conducted with DESeq2 [89], comparing all 20 °C samples versus all 28
452 °C samples, for both eRNA and oRNA samples. A gene was determined to be significantly
453 differentially expressed if it had a false discovery rate (FDR) adjusted p-value < 0.1. WEGO 2.0
454 [90] was used to classify GO terms (acquired from [45]) to all annotated *D. pulex* DEGs. Gene
455 ontology enrichment analysis was conducted using topGO [91] by comparing GO terms of the
456 relevant genes to the genomic background. A FDR corrected p-value < 0.05 indicated significant
457 enrichment. A principal component analysis (PCA) of the regularized log (rlog)-normalized
458 eRNA and oRNA gene expression profiles (300 most variant genes) at both optimal and heat
459 stressed conditions was conducted using DESeq2's plotPCA function [89]. Heatmaps showing
460 the relative expression (Z-score calculated for each gene) of all significantly differentially
461 expressed genes between 20 and 28 °C, in both *D. pulex* oRNA and eRNA, and all other species
462 in eRNA, were generated using pheatmap [92].

463 We used the SqueezeMeta metatranscriptomics pipeline for the eRNA community-wide
464 analysis [46]. Briefly, metatranscriptomes were assembled in sequential mode and contigs were
465 assembled using Megahit [93]. Diamond was used to align all contigs to the GenBank nr and
466 KEGG databases for taxonomic assignment and KEGG ID annotation, respectively [94,95]. The
467 SQMtools R package was used to further analyze the SqueezeMeta data [96]. Stacked-bar plots
468 were created to represent the most relatively abundant taxa. We conducted non-metric
469 multidimensional scaling analysis of all KEGG IDs to visualize the community functional
470 profiles.
471
472 **Data Availability**
473 Sequences for all RNA-seq libraries have been uploaded to the NCBI SRA database and will be
474 made public upon acceptance of this manuscript.
475
476 **Acknowledgements**
477 We thank Elif Irmak Bektas, Cathy Shen and Gabrielle Pellegrin for assisting with animal
478 husbandry and sample collection. We are grateful to Lars Iversen and Caren Helbing for
479 providing comments on an early version of this manuscript and Takahiro Maruki for input
480 regarding the bioinformatic analysis. This study was supported by the Natural Sciences and
481 Engineering Research Council (NSERC) of Canada through the Canada Research Chair program
482 and a Discovery Grant to Melania Cristescu.
483
484
485

References

1. Deiner K, Bik HM, Mächler E, Seymour M, Lacoursière-Roussel A, Altermatt F, et al. Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. Mol Ecol. 2017;26: 5872–5895. doi:10.1111/mec.14350

2. Cristescu ME, Hebert PDN. Uses and misuses of environmental dna in biodiversity science and conservation. Annu Rev Ecol Evol Syst. 2018;49: 209–230. doi:10.1146/annurev-ecolsys-110617-062306

3. Huang W, Sher Y-P, Peck K, Fung YCB. Matching gene activity with physiological functions. Proc Natl Acad Sci. 2002;99: 2603–2608. doi:10.1073/pnas.042684399

4. Alvarez M, Schrey AW, Richards CL. Ten years of transcriptomics in wild populations: what have we learned about their ecology and evolution? Mol Ecol. 2015;24: 710–725. doi:10.1111/mec.13055

5. Baillon L, Pierron F, Coudret R, Normendeau E, Caron A, Peluhet L, et al. Transcriptome profile analysis reveals specific signatures of pollutants in Atlantic eels. Ecotoxicology. 2015;24: 71–84. doi:10.1007/s10646-014-1356-x

6. Gleason LU, Burton RS. RNA-seq reveals regional differences in transcriptome response to heat stress in the marine snail Chlorostoma funebralis. Mol Ecol. 2015;24: 610–627. doi:10.1111/mec.13047

7. Houde ALS, Akbarzadeh A, Günther OP, Li S, Patterson DA, Farrell AP, et al. Salmonid gene expression biomarkers indicative of physiological responses to changes in salinity and temperature, but not dissolved oxygen. J Exp Biol. 2019;222. doi:10.1242/jeb.198036

8. Jeffries KM, Teffer A, Michaleski S, Bernier NJ, Heath DD, Miller KM. The use of non-lethal sampling for transcriptomics to assess the physiological status of wild fishes. Comp Biochem Physiol B Biochem Mol Biol. 2021;256: 110629. doi:10.1016/j.cbpb.2021.110629

9. Miller KM, Günther OP, Li S, Kaukinen KH, Ming TJ. Molecular indices of viral disease development in wild migrating salmon. Conserv Physiol. 2017;5: cox036. doi:10.1093/conphys/cox036

10. Amarasiri M, Furukawa T, Nakajima F, Sei K. Pathogens and disease vectors/hosts monitoring in aquatic environments: Potential of using eDNA/eRNA based approach. Sci Total Environ. 2021;796: 148810. doi:10.1016/j.scitotenv.2021.148810

11. Cristescu ME. Can environmental rna revolutionize biodiversity science? Trends Ecol Evol. 2019;34: 694–697. doi:10.1016/j.tree.2019.05.003

12. Veilleux HD, Misutka MD, Glover CN. Environmental DNA and environmental RNA: Current and prospective applications for biological monitoring. Sci Total Environ. 2021;782: 146891. doi:10.1016/j.scitotenv.2021.146891

522 13. Yates MC, Derry AM, Cristescu ME. Environmental rna: a revolution in ecological
523 resolution? Trends Ecol Evol. 2021;36: 601–609. doi:10.1016/j.tree.2021.03.001

524 14. Frias-Lopez J, Shi Y, Tyson GW, Coleman ML, Schuster SC, Chisholm SW, et al. Microbial
525 community gene expression in ocean surface waters. Proc Natl Acad Sci U S A. 2008;105:
526 3805–3810. doi:10.1073/pnas.0708897105

527 15. Gilbert JA, Field D, Huang Y, Edwards R, Li W, Gilna P, et al. Detection of large numbers
528 of novel sequences in the metatranscriptomes of complex marine microbial communities.
529 PLoS ONE. 2008;3. doi:10.1371/journal.pone.0003042

530 16. Poretsky RS, Bano N, Buchan A, LeCleir G, Kleikemper J, Pickering M, et al. Analysis of
531 microbial gene transcripts in environmental samples. Appl Environ Microbiol. 2005;71:
532 4121–4126. doi:10.1128/AEM.71.7.4121-4126.2005

533 17. Kagzi K, Hechler RM, Fussmann GF, Cristescu ME. Environmental RNA degrades more
534 rapidly than environmental DNA across a broad range of pH conditions. Mol Ecol Resour.
535 2022;22: 2640–2650. doi:10.1111/1755-0998.13655

536 18. Littlefair JE, Rennie MD, Cristescu ME. Environmental nucleic acids: A field-based
537 comparison for monitoring freshwater habitats using eDNA and eRNA. Mol Ecol Resour.
538 2022;0: 1–13. doi:10.1111/1755-0998.13671

539 19. Marshall NT, Vanderploeg HA, Chaganti SR. Environmental (e)RNA advances the
540 reliability of eDNA by predicting its age. Sci Rep. 2021;11: 2769. doi:10.1038/s41598-
541 021-82205-4

542 20. Tsuri K, Ikeda S, Hirohara T, Shimada Y, Minamoto T, Yamanaka H. Messenger RNA
543 typing of environmental RNA (eRNA): A case study on zebrafish tank water with
544 perspectives for the future development of eRNA analysis on aquatic vertebrates. Environ
545 DNA. 2021;3: 14–21. doi:10.1002/edn3.169

546 21. Wood SA, Biessy L, Latchford JL, Zaiko A, von Ammon U, Audrezet F, et al. Release and
547 degradation of environmental DNA and RNA in a marine system. Sci Total Environ.
548 2020;704: 135314. doi:10.1016/j.scitotenv.2019.135314

549 22. Salazar G, Paoli L, Alberti A, Huerta-Cepas J, Ruscheweyh H-J, Cuenca M, et al. Gene
550 expression changes and community turnover differentially shape the global ocean
551 metatranscriptome. Cell. 2019;179: 1068-1083.e21. doi:10.1016/j.cell.2019.10.014

552 23. Sunagawa S, Coelho LP, Chaffron S, Kultima JR, Labadie K, Salazar G, et al. Structure and
553 function of the global ocean microbiome. Science. 2015;348: 1261359.
554 doi:10.1126/science.1261359

555 24. Sunday JM, Bates AE, Dulvy NK. Global analysis of thermal tolerance and latitude in
556 ectotherms. Proc R Soc B Biol Sci. 2011;278: 1823–1830. doi:10.1098/rspb.2010.1295

557 25. Huey RB, Berrigan D. Temperature, demography, and ectotherm fitness. Am Nat. 2001;158:
558 204–210. doi:10.1086/321314

559 26. Schulte PM. The effects of temperature on aerobic metabolism: towards a mechanistic
560 understanding of the responses of ectotherms to a changing environment. Podrabsky JE,
561 Stillman JH, Tomanek L, editors. J Exp Biol. 2015;218: 1856–1866.
562 doi:10.1242/jeb.118851

563 27. Rodgers EM. Adding climate change to the mix: responses of aquatic ectotherms to the
564 combined effects of eutrophication and warming. Biol Lett. 2021;17: 20210442.
565 doi:10.1098/rsbl.2021.0442

566 28. Harley CDG, Randall Hughes A, Hultgren KM, Miner BG, Sorte CJB, Thornber CS, et al.
567 The impacts of climate change in coastal marine systems. Ecol Lett. 2006;9: 228–241.
568 doi:10.1111/j.1461-0248.2005.00871.x

569 29. Reid AJ, Carlson AK, Creed IF, Eliason EJ, Gell PA, Johnson PTJ, et al. Emerging threats
570 and persistent conservation challenges for freshwater biodiversity. Biol Rev Camb Philos
571 Soc. 2019;94: 849–873. doi:10.1111/brv.12480

572 30. Narum SR, Campbell NR. Transcriptomic response to heat stress among ecologically
573 divergent populations of redband trout. BMC Genomics. 2015;16: 103.
574 doi:10.1186/s12864-015-1246-5

575 31. Veilleux HD, Ryu T, Donelson JM, van Herwerden L, Seridi L, Ghosheh Y, et al. Molecular
576 processes of transgenerational acclimation to a warming ocean. Nat Clim Change. 2015;5:
577 1074–1078. doi:10.1038/nclimate2724

578 32. Veilleux HD, Ryu T, Donelson JM, Ravasi T, Munday PL. Molecular response to extreme
579 summer temperatures differs between two genetically differentiated populations of a coral
580 reef fish. Front Mar Sci. 2018;5. doi:doi.org/10.3389/fmars.2018.00349

581 33. Akbarzadeh A, Günther OP, Houde AL, Li S, Ming TJ, Jeffries KM, et al. Developing
582 specific molecular biomarkers for thermal stress in salmonids. BMC Genomics. 2018;19:
583 749. doi:10.1186/s12864-018-5108-9

584 34. Savary R, Barshis DJ, Voolstra CR, Cárdenas A, Evensen NR, Banc-Prandi G, et al. Fast and
585 pervasive transcriptomic resilience and acclimation of extremely heat-tolerant coral
586 holobionts from the northern Red Sea. Proc Natl Acad Sci. 2021;118.
587 doi:10.1073/pnas.2023298118

588 35. Voolstra CR, Valenzuela JJ, Turkarslan S, Cárdenas A, Hume BCC, Perna G, et al.
589 Contrasting heat stress response patterns of coral holobionts across the Red Sea suggest
590 distinct mechanisms of thermal tolerance. Mol Ecol. 2021;30: 4466–4480.
591 doi:10.1111/mec.16064

592 36. Ziegler M, Seneca FO, Yum LK, Palumbi SR, Voolstra CR. Bacterial community dynamics
593 are linked to patterns of coral heat tolerance. Nat Commun. 2017;8: 14213.
594 doi:10.1038/ncomms14213

595 37. Chen N, Huang Z, Lu C, Shen Y, Luo X, Ke C, et al. Different transcriptomic responses to
596 thermal stress in heat-tolerant and heat-sensitive pacific abalones indicated by cardiac
597 performance. Front Physiol. 2019;9. doi:10.3389/fphys.2018.01895

598 38. Kelly MW, Pankey MS, DeBiasse MB, Plachetzki DC. Adaptation to heat stress reduces
599 phenotypic and transcriptional plasticity in a marine copepod. Funct Ecol. 2017;31: 398–
600 406. doi:10.1111/1365-2435.12725

601 39. Semmouri I, Asselman J, Van Nieuwerburgh F, Deforce D, Janssen CR, De Schamphelaere
602 KAC. The transcriptome of the marine calanoid copepod Temora longicornis under heat
603 stress and recovery. Mar Environ Res. 2019;143: 10–23.
604 doi:10.1016/j.marenvres.2018.10.017

605 40. Becker D, Reydelet Y, Lopez JA, Jackson C, Colbourne JK, Hawat S, et al. The
606 transcriptomic and proteomic responses of Daphnia pulex to changes in temperature and
607 food supply comprise environment-specific and clone-specific elements. BMC Genomics.
608 2018;19. doi:10.1186/s12864-018-4742-6

609 41. Yampolsky LY, Zeng E, Lopez J, Williams PJ, Dick KB, Colbourne JK, et al. Functional
610 genomics of acclimation and adaptation in response to thermal stress in Daphnia. BMC
611 Genomics. 2014;15: 859. doi:10.1186/1471-2164-15-859

612 42. Bozinovic F, Pörtner H-O. Physiological ecology meets climate change. Ecol Evol. 2015;5:
613 1025–1030. doi:10.1002/ece3.1403

614 43. Evans TG, Hofmann GE. Defining the limits of physiological plasticity: how gene
615 expression can assess and predict the consequences of ocean change. Philos Trans R Soc
616 B Biol Sci. 2012;367: 1733–1745. doi:10.1098/rstb.2012.0019

617 44. Wikelski M, Cooke SJ. Conservation physiology. Trends Ecol Evol. 2006;21: 38–46.
618 doi:10.1016/j.tree.2005.10.018

619 45. Ye Z, Xu S, Spitze K, Asselman J, Jiang X, Ackerman MS, et al. A new reference genome
620 assembly for the microcrustacean daphnia pulex. G3 GenesGenomesGenetics. 2017;7:
621 1405–1416. doi:10.1534/g3.116.038638

622 46. Tamames J, Puente-Sánchez F. Squeezemeta, a highly portable, fully automatic
623 metagenomic analysis pipeline. Front Microbiol. 2019;9.
624 doi:doi.org/10.3389/fmicb.2018.03349

625 47. Cowart DA, Murphy KR, Cheng C-HC. Metagenomic sequencing of environmental DNA
626 reveals marine faunal assemblages from the West Antarctic Peninsula. Mar Genomics.
627 2018;37: 148–160. doi:10.1016/j.margen.2017.11.003

628 48. Monchamp M-È, Walsh DA, Garner RE, Kraemer SA, Beisner BE, Cristescu ME, et al.
629 Comparative analysis of zooplankton diversity in freshwaters: What can we gain from
630 metagenomic analysis? Environ DNA. 2022;0: 1–15. doi:10.1002/edn3.335

631 49. Stat M, Huggett MJ, Bernasconi R, DiBattista JD, Berry TE, Newman SJ, et al. Ecosystem
632 biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine
633 environment. Sci Rep. 2017;7: 12240. doi:10.1038/s41598-017-12501-5

634 50. Székely D, Corfixen NL, Mørch LL, Knudsen SW, McCarthy ML, Teilmann J, et al.
635 Environmental DNA captures the genetic diversity of bowhead whales (Balaena
636 mysticetus) in West Greenland. Environ DNA. 2021;3: 248–260. doi:10.1002/edn3.176

637 51. Fedorenkova A, Vonk JA, Lenders HJR, Ouborg NJ, Breure AM, Hendriks AJ.
638 Ecotoxicogenomics: bridging the gap between genes and populations. Environ Sci
639 Technol. 2010;44: 4328–4333. doi:10.1021/es9037287

640 52. Jovic K, Sterken MG, Grilli J, Bevers RPJ, Rodriguez M, Riksen JAG, et al. Temporal
641 dynamics of gene expression in heat-stressed Caenorhabditis elegans. PLOS ONE.
642 2017;12: e0189445. doi:10.1371/journal.pone.0189445

643 53. Snell TW, Brogdon SE, Morgan MB. Gene expression profiling in ecotoxicology.
644 Ecotoxicology. 2003;12: 475–483. doi:10.1023/B:ECTX.0000003033.09923.a8

645 54. Chain FJJ, Finlayson S, Crease T, Cristescu M. Variation in transcriptional responses to
646 copper exposure across Daphnia pulex lineages. Aquat Toxicol. 2019;210: 85–97.
647 doi:10.1016/j.aquatox.2019.02.016

648 55. Connon R, Hooper HL, Sibly RM, Lim F-L, Heckmann L-H, Moore DJ, et al. Linking
649 molecular and population stress responses in daphnia magna exposed to cadmium.
650 Environ Sci Technol. 2008;42: 2181–2188. doi:10.1021/es702469b

651 56. Shaw JR, Colbourne JK, Davey JC, Glaholt SP, Hampton TH, Chen CY, et al. Gene
652 response profiles for Daphnia pulex exposed to the environmental stressor cadmium
653 reveals novel crustacean metallothioneins. BMC Genomics. 2007;8: 477.
654 doi:10.1186/1471-2164-8-477

655 57. Aylward FO, Eppley JM, Smith JM, Chavez FP, Scholin CA, DeLong EF. Microbial
656 community transcriptional networks are conserved in three domains at ocean basin scales.
657 Proc Natl Acad Sci. 2015;112: 5443–5448. doi:10.1073/pnas.1502883112

658 58. Marchetti A, Schruth DM, Durkin CA, Parker MS, Kodner RB, Berthiaume CT, et al.
659 Comparative metatranscriptomics identifies molecular bases for the physiological
660 responses of phytoplankton to varying iron availability. Proc Natl Acad Sci. 2012;109:
661 E317–E325. doi:10.1073/pnas.1118408109

662 59. Moniruzzaman M, Wurch LL, Alexander H, Dyhrman ST, Gobler CJ, Wilhelm SW. Virus-
663 host relationships of marine single-celled eukaryotes resolved from metatranscriptomics.
664 Nat Commun. 2017;8: 16054. doi:10.1038/ncomms16054

60. Vorobev A, Dupouy M, Carradec Q, Delmont TO, Annamalé A, Wincker P, et al. Transcriptome reconstruction and functional analysis of eukaryotic marine plankton communities via high-throughput metagenomics and metatranscriptomics. Genome Res. 2020;30: 647–659. doi:10.1101/gr.253070.119

61. Trench-Fiol S, Fink P. Metatranscriptomics from a small aquatic system: microeukaryotic community functions through the diurnal cycle. Front Microbiol. 2020;11. doi:doi.org/10.3389/fmicb.2020.01006

62. Czypionka T, Krugman T, Altmüller J, Blaustein L, Steinfartz S, Templeton AR, et al. Ecological transcriptomics – a non-lethal sampling approach for endangered fire salamanders. Methods Ecol Evol. 2015;6: 1417–1425. doi:10.1111/2041-210X.12431

63. Veldhoen N, Beckerton JE, Mackenzie-Grieve J, Stevenson MR, Truelson RL, Helbing CC. Development of a non-lethal method for evaluating transcriptomic endpoints in Arctic grayling (Thymallus arcticus). Ecotoxicol Environ Saf. 2014;105: 43–50. doi:10.1016/j.ecoenv.2014.03.030

64. Martins CL, Walker TI, Reina RD. Stress-related physiological changes and post-release survival of elephant fish (Callorhinchus milii) after longlining, gillnetting, angling and handling in a controlled setting. Fish Res. 2018;204: 116–124. doi:10.1016/j.fishres.2018.01.016

65. Portz DE, Woodley CM, Cech JJ. Stress-associated impacts of short-term holding on fishes. Rev Fish Biol Fish. 2006;16: 125–170. doi:10.1007/s11160-006-9012-z

66. Young T, Walker SP, Alfaro AC, Fletcher LM, Murray JS, Lulijwa R, et al. Impact of acute handling stress, anaesthesia, and euthanasia on fish plasma biochemistry: implications for veterinary screening and metabolomic sampling. Fish Physiol Biochem. 2019;45: 1485–1494. doi:10.1007/s10695-019-00669-8

67. Moran MA, Satinsky B, Gifford SM, Luo H, Rivers A, Chan L-K, et al. Sizing up metatranscriptomics. ISME J. 2013;7: 237–243. doi:10.1038/ismej.2012.94

68. Shi Y, Tyson GW, Eppley JM, DeLong EF. Integrated metatranscriptomic and metagenomic analyses of stratified microbial assemblages in the open ocean. ISME J. 2011;5: 999–1013. doi:10.1038/ismej.2010.189

69. Wetterstrand K. DNA Sequencing Costs: Data. In: Genome.gov [Internet]. 2021. Available: https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data

70. Klingenberg H, Meinicke P. How to normalize metatranscriptomic count data for differential expression analysis. PeerJ. 2017;5: e3859. doi:10.7717/peerj.3859

71. Mallick H, Rahnavard A, McIver LJ, Ma S, Zhang Y, Nguyen LH, et al. Multivariable association discovery in population-scale meta-omics studies. PLOS Comput Biol. 2021;17: e1009442. doi:10.1371/journal.pcbi.1009442

701 72. Zhang Y, Thompson KN, Huttenhower C, Franzosa EA. Statistical approaches for
702 differential expression analysis in metatranscriptomics. Bioinformatics. 2021;37: i34–i41.
703 doi:10.1093/bioinformatics/btab327

704 73. Cordier T, Lanzén A, Apothéloz-Perret-Gentil L, Stoeck T, Pawlowski J. Embracing
705 environmental genomics and machine learning for routine biomonitoring. Trends
706 Microbiol. 2019;27: 387–397. doi:10.1016/j.tim.2018.10.012

707 74. Celis-Salgado MP, Cairns A, Kim N, Yan ND. The FLAMES medium: a new, soft-water
708 culture and bioassay medium for Cladocera. SIL Proc 1922-2010. 2008;30: 265–271.
709 doi:10.1080/03680770.2008.11902123

710 75. Leray M, Yang JY, Meyer CP, Mills SC, Agudelo N, Ranwez V, et al. A new versatile
711 primer set targeting a short fragment of the mitochondrial COI region for metabarcoding
712 metazoan diversity: application for characterizing coral reef fish gut contents. Front Zool.
713 2013;10: 34. doi:10.1186/1742-9994-10-34

714 76. Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010.
715 Available: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

716 77. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence
717 data. Bioinforma Oxf Engl. 2014;30: 2114–2120. doi:10.1093/bioinformatics/btu170

718 78. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast
719 universal RNA-seq aligner. Bioinformatics. 2013;29: 15–21.
720 doi:10.1093/bioinformatics/bts635

721 79. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq
722 quantification. Nat Biotechnol. 2016;34: 525–527. doi:10.1038/nbt.3519

723 80. Aeschlimann SH, Jönsson F, Postberg J, Stover NA, Petera RL, Lipps H-J, et al. The draft
724 assembly of the radically organized stylonychia lemnae macronuclear genome. Genome
725 Biol Evol. 2014;6: 1707–1723. doi:10.1093/gbe/evu139

726 81. Han J, Park JC, Choi B-S, Kim M-S, Kim H-S, Hagiwara A, et al. The genome of the marine
727 monogonont rotifer Brachionus plicatilis: Genome-wide expression profiles of 28
728 cytochrome P450 genes in response to chlorpyrifos and 2-ethyl-phenanthrene. Aquat
729 Toxicol. 2019;214: 105230. doi:10.1016/j.aquatox.2019.105230

730 82. Hirooka S, Hirose Y, Kanesaki Y, Higuchi S, Fujiwara T, Onuma R, et al. Acidophilic green
731 algal genome provides insights into adaptation to an acidic environment. Proc Natl Acad
732 Sci. 2017;114: E8304–E8313. doi:10.1073/pnas.1707072114

733 83. Nag Dasgupta C, Nayaka S, Toppo K, Singh AK, Deshpande U, Mohapatra A. Draft genome
734 sequence and detailed characterization of biofuel production by oleaginous microalga
735 Scenedesmus quadricauda LWG002611. Biotechnol Biofuels. 2018;11: 308.
736 doi:10.1186/s13068-018-1308-4

737 84. Prochnik SE, Umen J, Nedelcu AM, Hallmann A, Miller SM, Nishii I, et al. Genomic
738 analysis of organismal complexity in the multicellular green alga volvox carteri. Science.
739 2010. doi:10.1126/science.1188800

740 85. Rast P, Glöckner I, Boedeker C, Jeske O, Wiegand S, Reinhardt R, et al. Three novel species
741 with peptidoglycan cell walls form the new genus lacunisphaera gen. nov. in the family
742 opitutaceae of the verrucomicrobial subdivision 4. Front Microbiol. 2017;8.
743 doi:doi.org/10.3389/fmicb.2017.00202

744 86. Schomaker RA, Dudycha JL. De novo transcriptome assembly of the green alga
745 Ankistrodesmus falcatus. PLOS ONE. 2021;16: e0251668.
746 doi:10.1371/journal.pone.0251668

747 87. Suzuki S, Yamaguchi H, Nakajima N, Kawachi M. Raphidocelis subcapitata
748 (=Pseudokirchneriella subcapitata) provides an insight into genome evolution and
749 environmental adaptations in the Sphaeropleales. Sci Rep. 2018;8: 8058.
750 doi:10.1038/s41598-018-26331-6

751 88. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for
752 assigning sequence reads to genomic features. Bioinformatics. 2014;30: 923–930.
753 doi:10.1093/bioinformatics/btt656

754 89. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-
755 seq data with DESeq2. Genome Biol. 2014;15: 550. doi:10.1186/s13059-014-0550-8

756 90. Ye J, Zhang Y, Cui H, Liu J, Wu Y, Cheng Y, et al. WEGO 2.0: a web tool for analyzing and
757 plotting GO annotations, 2018 update. Nucleic Acids Res. 2018;46: W71–W75.
758 doi:10.1093/nar/gky400

759 91. Alexa A, Rahnenfuhrer J. Topgo: enrichment analysis for gene ontology. Bioconductor
760 version: Release (3.14); 2022. doi:10.18129/B9.bioc.topGO

761 92. Kolde R. pheatmap: Pretty Heatmaps. 2019. Available: https://CRAN.R-
762 project.org/package=pheatmap

763 93. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. MEGAHIT: an ultra-fast single-node
764 solution for large and complex metagenomics assembly via succinct de Bruijn graph.
765 Bioinformatics. 2015;31: 1674–1676. doi:10.1093/bioinformatics/btv033

766 94. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat
767 Methods. 2015;12: 59–60. doi:10.1038/nmeth.3176

768 95. Kanehisa M, Goto S. Kegg: kyoto encyclopedia of genes and genomes. Nucleic Acids Res.
769 2000;28: 27–30. doi:10.1093/nar/28.1.27

770 96. Puente-Sánchez F, García-García N, Tamames J. SQMtools: automated processing and
771 visual analysis of ’omics data with R and anvi’o. BMC Bioinformatics. 2020;21: 358.
772 doi:10.1186/s12859-020-03703-2

773 97. He S, Wurtzel O, Singh K, Froula JL, Yilmaz S, Tringe SG, et al. Validation of two
774 ribosomal RNA removal methods for microbial metatranscriptomics. Nat Methods.
775 2010;7: 807–812. doi:10.1038/nmeth.1507

776 98. Zhao S, Zhang Y, Gamini R, Zhang B, von Schack D. Evaluation of two main RNA-seq
777 approaches for gene quantification in clinical RNA sequencing: polyA+ selection versus
778 rRNA depletion. Sci Rep. 2018;8: 4781. doi:10.1038/s41598-018-23226-4

779 99. Zhao W, He X, Hoadley KA, Parker JS, Hayes DN, Perou CM. Comparison of RNA-Seq by
780 poly (A) capture, ribosomal RNA depletion, and DNA microarray for expression profiling.
781 BMC Genomics. 2014;15: 419. doi:10.1186/1471-2164-15-419

782 100. Hempel CA, Wright N, Harvie J, Hleap JS, Adamowicz SJ, Steinke D. Metagenomics
783 versus total RNA sequencing: most accurate data-processing tools, microbial identification
784 accuracy and perspectives for ecological assessments. Nucleic Acids Res. 2022;50: 9279–
785 9293. doi:10.1093/nar/gkac689

786 101. Lewin HA, Robinson GE, Kress WJ, Baker WJ, Coddington J, Crandall KA, et al. Earth
787 BioGenome Project: Sequencing life for the future of life. Proc Natl Acad Sci. 2018;115:
788 4325–4333. doi:10.1073/pnas.1720115115

789 102. Shakya M, Lo C-C, Chain PSG. Advances and challenges in metatranscriptomic analysis.
790 Front Genet. 2019;10. doi:10.3389/fgene.2019.00904

791 103. Fulda S, Gorman AM, Hori O, Samali A. Cellular stress responses: cell survival and cell
792 death. Int J Cell Biol. 2010;2010: 214074. doi:10.1155/2010/214074

793

794

795

796

797

798

799

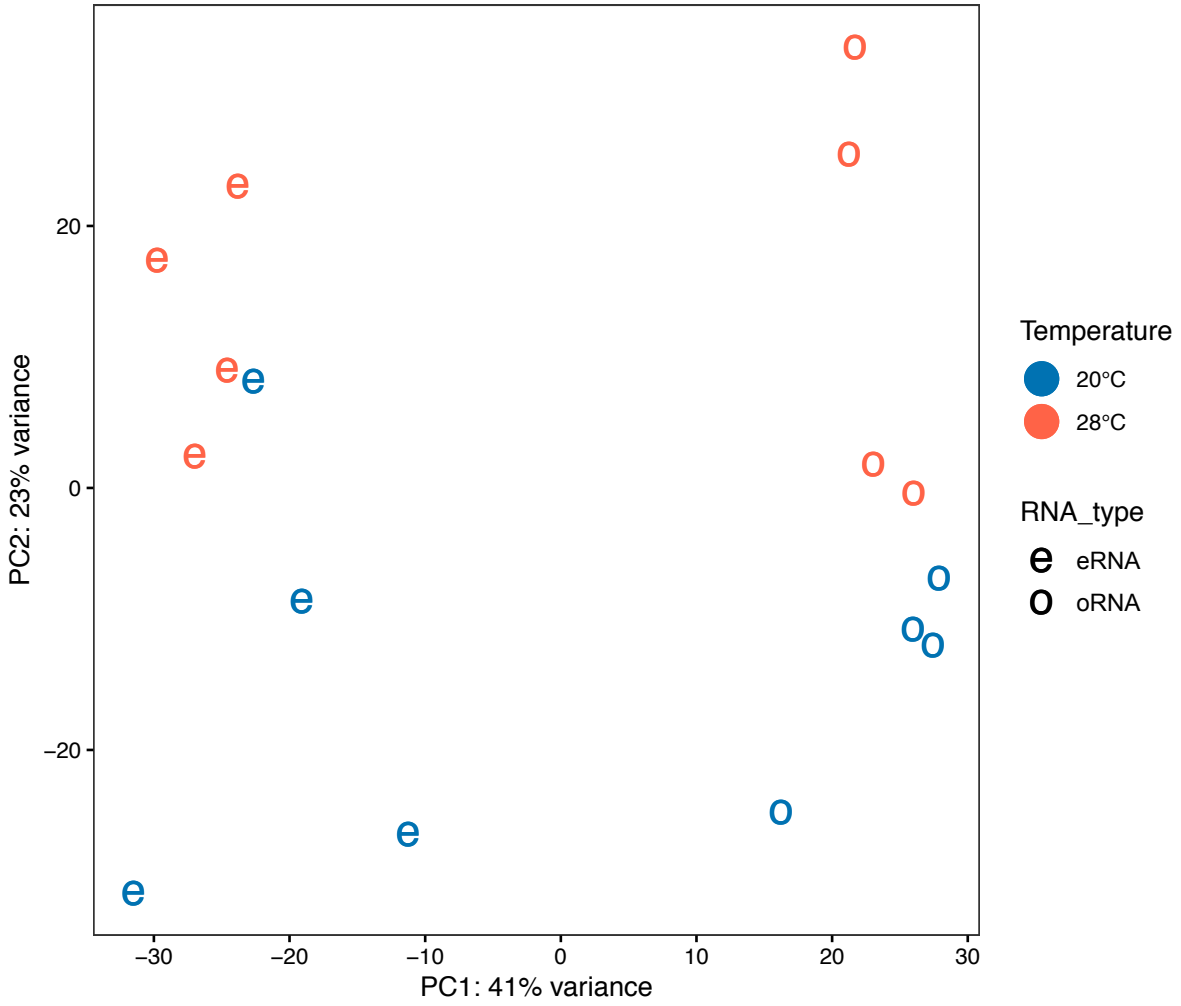
800

801

802

803 **Figures and Tables**

804



805

806 **Fig. 1.** Principal component analysis (PCA) of the top 300 *Daphnia pulex* genes, as ranked by
807 variance. Prior to the PCA, genes were filtered to include only those genes that had ≥ 1 count in
808 either 3/4 20 °C or 28 °C samples. Sample eRNA_20T1 clusters more closely with 28 °C
809 samples than 20 °C samples. PCA revealed separation by RNA type and temperature across the
810 PC1 and PC2, respectively. eRNA refers to extra-organismal RNA released by *D. pulex* into the
811 environment, whereas oRNA refers to organismal RNA extracted directly from *D. pulex* tissue.
812

813
814
815
816
817
818

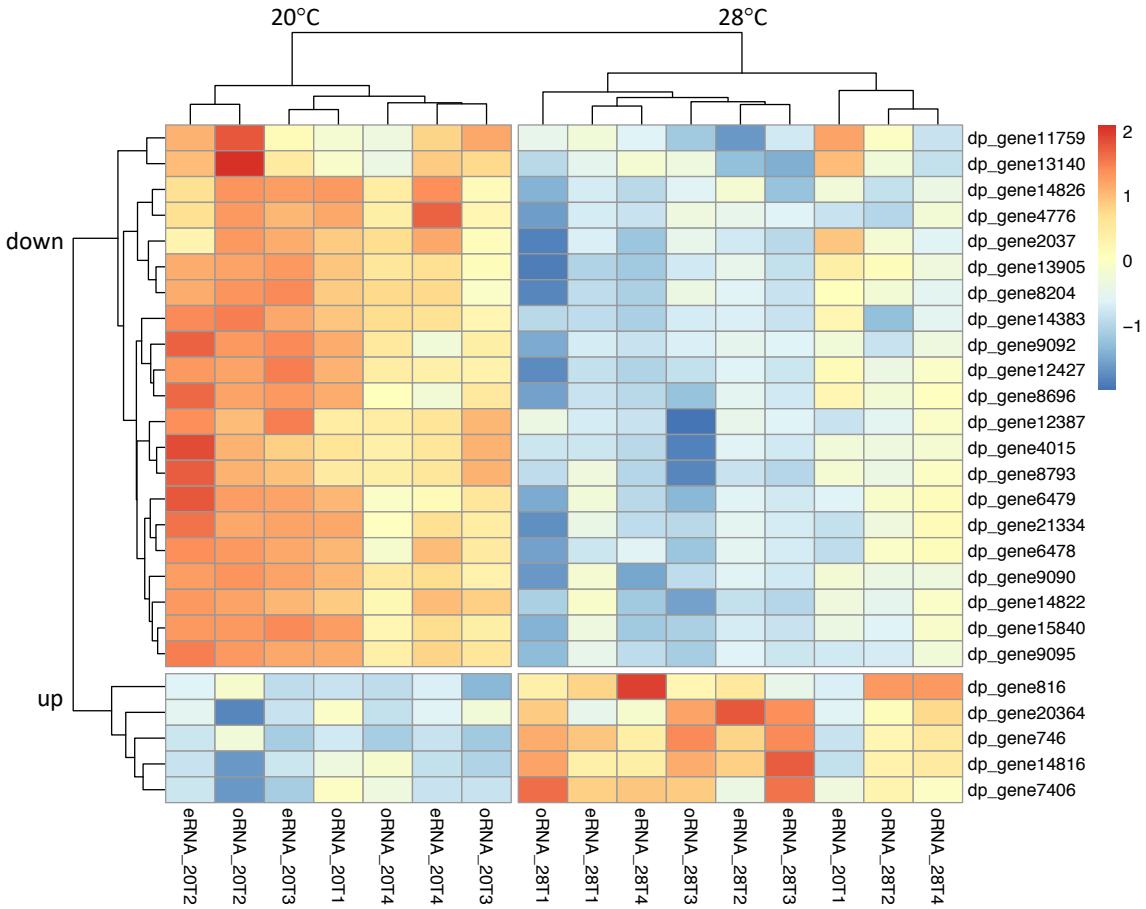


Fig. 2. Heatmap showing the relative expression (Z-score calculated for each gene) of all commonly significantly differentially expressed (DE) *Daphnia pulex* genes (false discovery rate adjusted p-value <0.1) between 20 °C and 28 °C, in both environmental RNA (eRNA) and organismal RNA (oRNA). DE analysis was conducted only for those genes that had ≥ 1 count in either 3/4 20 °C or 3/4 28 °C samples. Hierarchical clustering analysis revealed four groups: 20 °C and 28 °C samples, and up and down regulated genes. Sample eRNA_20T1 clusters with 28 °C samples. Colors indicate levels of relative expression, with blue and red indicating low and high, respectively. DE statistics and functional description of the genes are provided in Table S1.

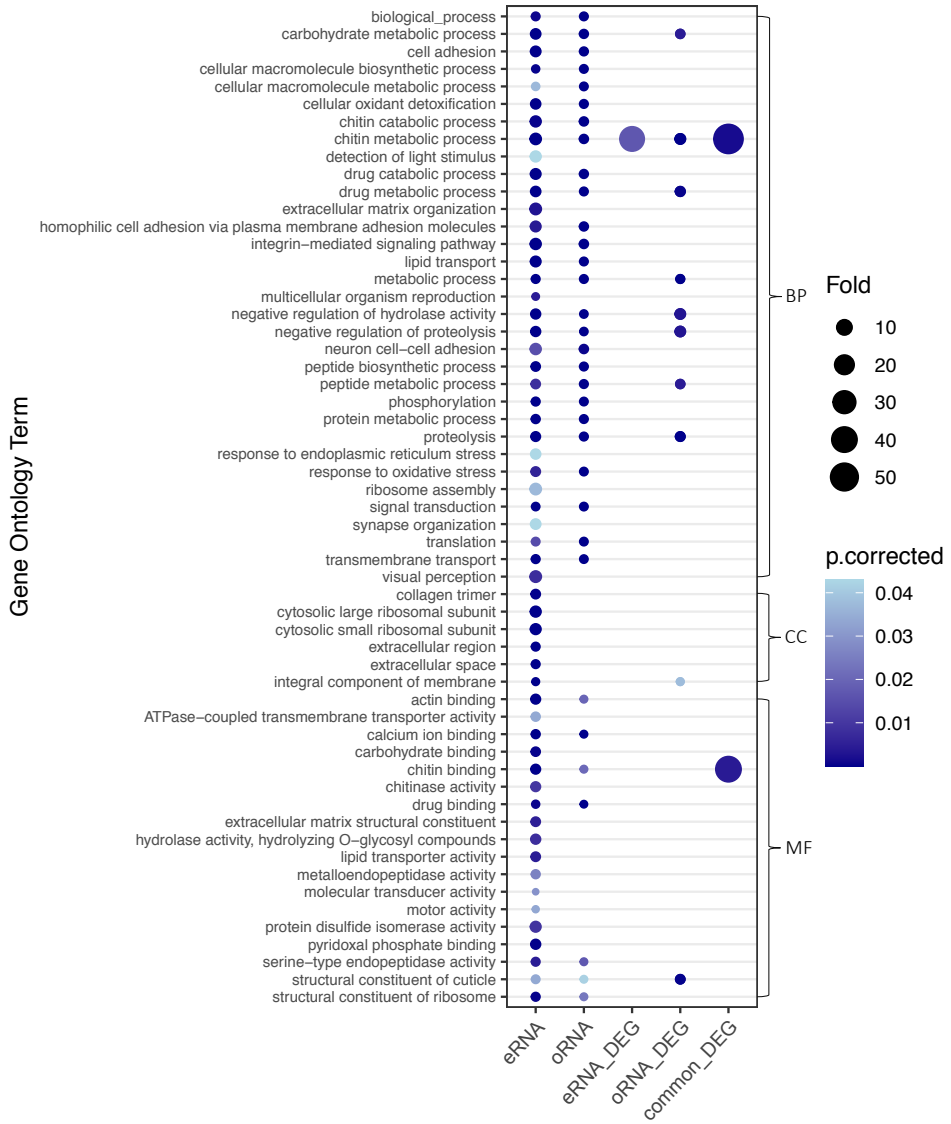


Fig. 3. Bubble plot of the Gene Ontology (GO) enriched terms associated with *Daphnia pulex* eRNA and oRNA detected genes, and eRNA, oRNA and commonly identified differentially expressed genes (DEG). The eRNA genes originate from extra-organismal *Daphnia* RNA. We used a false discovery rate (FDR) corrected p-value <0.05 in topGO to identify GO terms as significantly enriched compared to the genomic background. The bubble color indicates the FDR corrected p-value for the weight test statistic, while the size is proportional to the fold change between expected and observed genes annotated with GO term. The biological processes, cellular components and molecular functions GO domains are represented by BP, CC, and MF, respectively. Only those GO terms enriched for eRNA genes are shown. Full list of GO terms enriched for oRNA genes is provided in Fig. S3.

Table 1. Problems associated with using environmental transcriptomics based on eRNA to monitor the health of ecological communities and potential experimental (E), molecular (M), sequencing (S) and bioinformatic (B) solutions.

Problems	Description of problems	Possible solutions	References
Poor understanding of the ecology of eRNA	Little is known about the origin, state, persistence and degradation of eRNA in the environment.	E: Perform experiments that address fundamental questions of the ecology of eRNA (i.e. What are major sources of eRNA? What state is eRNA in? What factors influence eRNA persistence? How do abiotic factors influence eRNA decay rates?)	[11]
No understanding of environmental transcriptomics in natural ecosystems	No information available about how our results using environmental transcriptomics will translate to natural ecosystems	E: Perform mesocosm experiments using more complex, multi-trophic communities E: Perform studies using ecosystem wide manipulations E: Perform studies in natural ecosystems	
Rapid degradation of eRNA	eRNA is thought to degrade too rapidly to extract and quantify, but studies indicate that eRNA is detectable for up to 57 hours post organism release	E: Filter water as soon as possible (ideally < 4 hours) E: Freeze samples at -80 °C directly after filtration M: Keep samples on ice during extraction M: Minimize free-thaw cycles	[17,19,21]
Low concentration of eRNA	eRNA samples have much lower RNA concentrations than samples collected from an organisms tissue (i.e. 1 ng/µL vs 200 ng/µL)	E: Increase RNA yield by filtering large volumes of water (> 500 mL); EM: Collect additional technical replicates and pool them during the extraction step; M: Use RNA-seq library prep kits designed for low input starting material	
Low proportion of mRNA	mRNAs are of interest for gene expression analyses as they encode for proteins, but they account for < 5% of total RNA. Non-coding ribosomal and small RNAs account for > 85%	M: Perform rRNA depletion to remove rRNA. However, many commercial kits are optimized to remove human, mouse, rat and bacteria rRNA, and may not deplete rRNA from all species; M: Perform poly(A) enrichment to retain only eukaryotic mRNA. This will eliminate prokaryotic RNA, ribosomal and small RNAs, which make up the majority of total RNA B: If rRNA persists, use it to perform taxonomic identification analysis	[97–100]
Low eRNA recovery	eRNA samples are composed of extra-organismal RNA from macro-organisms and whole microorganism. Thus, eRNA reads will be in the minority (estimated <1%), resulting in low power of bioinformatic analyses to detect eRNA based changes in gene expression	E: Pre-filter water with fine mesh to size-selectively remove organisms; E: Use large filter pore size to reduce capture of microorganisms (i.e. > 0.2 µM); E: Sample eRNA from ecosystems that are densely populated with target species M: Perform rRNA depletion on total RNA samples; M: Perform poly(A) enrichment to eliminate prokaryotic RNA; S: Perform deep sequencing to increase eRNA reads (i.e. > 150 million reads per sample)	
Low eRNA read depth	eRNA samples are extremely diverse, resulting in reads from taxa across the trophic chain. Transcript abundance in a sample will influence read depth, and if many micro-organisms are present, eukaryotic genes will have low read depth. Current molecular methods for enrichment and depletion are optimized for organismal, not diverse environmental samples.	M: Create low input poly(A) enrichment methods. Current enrichment methods are designed for organismal RNA from tissue samples and require high concentration of input RNA; M: Increase eukaryotic gene read depth through poly(A) enrichment; MB: Create rRNA depletion methods that remove rRNA from all taxa as current depletion methods are optimized to remove single taxa and bacterial rRNA. If unable to remove all rRNA from eRNA samples, researchers could use rRNA reads for taxonomic information; M: Perform rRNA depletion; M: Amplify biomarker genes via targeted (qPCR, ddPCR) approaches to analyze specific biomarker genes	[13]
Difficulty characterizing read function due to lack of eukaryotic reference genomes	Reference genomes are used to functionally and taxonomically characterize reads. eRNA samples will capture eRNA from eukaryotes across the trophic chain, but genomes are available for only 0.4% of eukaryotic species.	ES: Sequence all eukaryotic species (e.g. Earth BioGenome Project) B: *De novo* assembly of reads into longer scaffolds, which enables the identification of homologs.	[101,102]

Lack of baseline in natural ecosystems to compare against	Differential gene expression analyses are conducted by comparing treatment samples (i.e. stressed community) against a control 'baseline' (i.e. health community). In natural ecosystems it is difficult or often not possible to obtain a baseline control as many stressors act simultaneously on a system	E: Conduct time series eRNA sampling of the same environment and compare time points; E: Compare environments with similar abiotic conditions. If healthy, one could serve as a baseline to compare against; EM: Couple eRNA with targeted PCR approaches (qRT-PCR, ddPCR) and biomarker genes to infer physiological states of specific species; EB: Reconstruct mock communities under neutral conditions as a baseline; B: The cellular stress response is triggered by general stress and as such genes and pathways involved could be broadly indicative of community stress; B: Use housekeeping genes with consistent expression levels as baselines to compare against; B: Supervised machine learning models could be trained with existing metatranscriptomic, transcriptomic and metagenomic data that have identified functional terms, metabolic pathways and taxa associated with stressors. Environmental transcriptomic data could be used as input and the model would predict the stressors involved and health of the community	[13,71–73,103]
False detection of differential transcript expression due to changes in population densities	Between environments, species across the trophic chain exhibit differences in abundances. RNA transcript abundance varies with DNA abundance. Thus, changes in RNA transcript abundance, independent of differential expression, could be recovered.	EB: Collect and sequence eDNA in parallel. eDNA can acts as a species abundance estimator, independent of transcript abundance, that we can normalize eRNA reads to. Allows us to track whether the proportion of eRNA reads follows species abundance, or is up or down regulated; B: 'Taxon-specific scaling' accounts for variation in abundances by separating and then normalizing metatranscriptomic data into organism-specific bins. A differential expression analysis could then be performed after the normalized matrices are recombined into community wide data	[13,22,70,72]