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

Short title: Using eRNA transcriptomics to assess biotic heat stress responses

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

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 (expression 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 heat stress responses of *Daphnia.* Using eRNA, we identified 32 *Daphnia* genes to be differentially expressed following heat stress. Of these, 17 were also differentially expressed and exhibited similar levels of relative expression in organismal RNA. In addition to the extra-organismal *Daphnia* response, eRNA detected community-wide heat stress responses consisting of distinct functional profiles and 121 differentially expressed genes across 8 taxa. 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.

50 **Introduction**

51 In an era of ecological crises, monitoring the presence and health of organisms within an
52 ecosystem is essential for the preservation of biodiversity. Environmental DNA is routinely used
53 for biomonitoring, but such applications are limited to species detection (Cristescu & Hebert,
54 2018; Deiner et al., 2017). The RNA molecule provides an additional layer of functional
55 information, as transcribed genes reflect an organism's physiological status (Huang et al., 2002).
56 Thus, transcriptomics is used to infer the ecological health of organisms, but such surveys are
57 species-specific and dependent on sampling organismal RNA (oRNA) directly, which is labor-
58 intensive and requires tissue collection or animal sacrifice (Alvarez et al., 2015; Baillon et al.,
59 2015; Gleason & Burton, 2015; Houde et al., 2019; Jeffries et al., 2021; Miller et al., 2017).
60 Non-invasive environmental RNA (eRNA) approaches could overcome these limitations by
61 providing insights into the health of populations and communities (Amarasiri et al., 2021;
62 Cristescu, 2019; Veilleux et al., 2021; Yates et al., 2021). However, using eRNA to inform on
63 ecological health has only been theorized (see Yates et al., 2021) and remains yet to be
64 empirically tested.

65 We broadly define eRNA as RNA extracted from the environment, including extra-
66 organismal RNA released by macroorganisms as well as whole microorganisms. Microbiologists
67 have long extracted RNA from bulk samples to profile gene expression in microbial
68 communities, referred to as metatranscriptomics (Frias-Lopez et al., 2008; Gilbert et al., 2008;
69 Poretsky et al., 2005). Despite more than two decades of research, metatranscriptomics remains
70 limited to studying the gene expression of microorganisms captured in bulk samples rich in
71 oRNA. The general assumption has been that the eRNA released by macroorganisms into the
72 environment is labile and degrades too rapidly to be reliably detected or quantified (but see

73 Cristescu, 2019). However, the RNA shed by macroorganisms into their surrounding
74 environment can be successfully extracted and quantified (T. Jo et al., 2022; Kagzi et al., 2022;
75 Littlefair et al., 2022; Marshall et al., 2021; Miyata et al., 2021; Tsuri et al., 2021; von Ammon et
76 al., 2019; Wood et al., 2020). Despite this evidence supporting the extractability of eRNA,
77 environmental transcriptomics, defined here as eRNA based gene expression profiling, remained
78 untested.

79 In aquatic systems, environmental conditions modulate biodiversity and are reflected in
80 changes of metatranscriptomes (Frias-Lopez et al., 2008; Salazar et al., 2019; Sunagawa et al.,
81 2015; Sunday et al., 2011). Temperature is a key abiotic factor that influences the physiology and
82 fitness of ectotherms due to the relationship between body and environmental temperatures
83 (Huey & Berrigan, 2001; Schulte, 2015; Rodgers, 2021). Rapid increases in temperature, such as
84 heatwaves, can exceed the thermal limits of aquatic ectotherms and exponentially increase
85 standard metabolic rates, resulting in decreases in fitness and survival (Harley et al., 2006; Reid
86 et al., 2019; Rodgers, 2021; Schulte, 2015). Transcriptomics is widely used to study the heat
87 stress response of many taxa, including fish (Akbarzadeh et al., 2018; Houde et al., 2019; Narum
88 & Campbell, 2015; Veilleux et al., 2015, 2018), coral holobionts (Savary et al., 2021; Voolstra et
89 al., 2021), molluscs (Chen et al., 2019; Gleason & Burton, 2015), and copepods (Kelly et al.,
90 2017; Semmouri et al., 2019). *Daphnia pulex*, a key bioindicator species, exhibits widespread
91 downregulation of genes involved in metabolic processes under heat stress (Becker et al., 2018;
92 Yampolsky et al., 2014). Clearly, gene expression surveys can provide species-specific
93 functional insights important for conserving aquatic biodiversity (Bozinovic & Pörtner, 2015;
94 Evans & Hofmann, 2012; Wikelski & Cooke, 2006). However, a species by species analysis
95 based on oRNA cannot capture community-wide changes without extensive efforts. Mapping

96 eRNA reads to the reference genomes of multiple target species and using existing
97 metatranscriptomic pipelines, could potentially enable environmental transcriptomics to reveal
98 gene expression responses across the trophic chain.

99 We constructed simple mock freshwater communities containing *Daphnia pulex,* as well
100 as three phytoplankton species and opportunistic microorganisms that colonized the artificial
101 lake water. The communities were exposed to control (20 °C) and near lethal *Daphnia* heat stress
102 (28 °C) conditions for seven days. To enable comparisons between eRNA and oRNA based gene
103 expression profiles, we sequenced eRNA from tank water (after complete *Daphnia* removal) and
104 oRNA from *Daphnia* tissue. We conducted differential gene expression analyses between the
105 control and heat stress conditions and compared the *Daphnia* functional stress response
106 identified in eRNA and oRNA. In addition to the *Daphnia* analysis, we used eRNA to investigate
107 the community-wide heat stress response. We hypothesized that 1) eRNA captures
108 transcriptional responses to heat stress and 2) eRNA contains a subset of the differentially
109 expressed *Daphnia* genes observed in oRNA after exposure to heat stress. Based on our findings,
110 we discuss the potential of eRNA based transcriptomics for the biological monitoring of
111 ecological communities and also provide recommendations to overcome potential limitations.
112

113 **Materials and Methods**

114 *Experimental design*

115 We constructed simple mock freshwater communities composed of *Daphnia pulex,* three
116 algae species (*Ankistrodesmus falcatus, Scenedesmus quadricauda,* and *Pseudokirchneriella*
117 *subcapitata*) and microorganisms that colonized the artificial lake water. The communities were
118 exposed to 20 °C (control) and 28 °C (neat-lethal stress for Daphnia) temperatures for seven days

119 to mimic natural heatwave conditions and allow for eRNA accumulation in tank water.
120 Communities were reared in 8 L tanks containing artificial lake water (Celis-Salgado et al.,
121 2008) and were seeded with a starting population of 50 parthenogenetic females of *Daphnia*
122 *pulex* (clone from Illinois, USA; 40.24, -87.78). *Daphnia* populations were kept in controlled
123 chambers for 45 days at control growing conditions: 20 °C, relative humidity of 50% and a
124 photoperiod of 16:8 hour light:dark. *Daphnia* were fed 1.5 mL algae (1:1:1 ratio of *A. falcatus, S.*
125 *quadricauda,* and *P. subcapitata*) twice per week. After 45 days, the *Daphnia* populations in all
126 tanks reached a minimum of 500 individuals and we began the experiment by transferring tanks
127 to controlled experimental chambers pre-set to the respective temperatures (20 °C and 28 °C).
128
129 *Sample collection*
130 Each temperature treatment had four tanks (biological replicates), and we collected four
131 eRNA and oRNA technical replicates per tank. We stirred the tanks for 60 seconds and filtered
132 500 mL of water through a 60 µm mesh to remove all *Daphnia* and ephippia. To capture eRNA,
133 the *Daphnia* free water was then filtered through a 0.7 µm glass microfiber filter. Each filter was
134 cut in half and placed in individual 1.5 mL microcentrifuge tubes containing 370 µL RLT buffer
135 (Qiagen) and 3.7 µL β-mercaptoethanol. Tubes were immediately stored at -80 °C until eRNA
136 extraction. Negative eRNA filtration control samples (500 mL of distilled water) were collected
137 at the beginning, middle and end of the sampling day, using the same methods. For each oRNA
138 sample, six live organisms (3 adults and 3 juveniles) per tank were haphazardly collected using a
139 sterile pipette and transferred to a 1.5 mL microcentrifuge tube containing 370 µL RLT buffer
140 (Qiagen) and 3.7 µL β-mercaptoethanol. To maintain the *Daphnia* heat stress transcriptional
141 signal, tubes were flash frozen in liquid nitrogen for 2 seconds, and immediately stored at -80 °C

142 until RNA extraction. The following system was used for sample identification:
143 RNAtype_Temperature&TankID (e.g. sample “eRNA_20T1" refers to eRNA collected at 20 °C
144 from tank 1).
145
146 *Extractions of eRNA and oRNA samples*
147 Samples of eRNA and oRNA were extracted using the RNeasy Mini Kit (Qiagen),
148 following the manufacturer's protocol with the following modifications. eRNA samples were
149 thawed on ice, gently vortexed for 5 seconds and then centrifuged for 3 minutes at 13,300 rpm to
150 separate the liquid from filter. To increase the eRNA yield, liquid from two technical replicates
151 was transferred into a 5 mL microcentrifuge tube containing 800 µL 70% EtOh and pipette
152 mixed. This mixture was transferred to an RNeasy spin column (700 µL at a time) and
153 centrifuged at 10,000 rpm for 15 seconds. We repeated this four times, until all remaining liquid
154 from the 5 mL microcentrifuge tube was transferred to the spin column. For oRNA, *Daphnia*
155 samples were thawed on ice and homogenized via sterile pestle and mortar mixer for 60 seconds.
156 We then added 400 µL of 70% EtOh to the homogenized tissue. The resulting mixture was
157 pipette mixed and transferred to an RNeasy spin column, where it was centrifuged at 10,000 rpm
158 for 15 seconds. The remainder of the eRNA and oRNA extraction procedures followed the
159 manufacturer’s instructions, with RNA being eluted in 40 µL of DNase and RNase free
160 molecular grade water.
161
162 *DNA digestion*
163 Immediately following RNA extraction, an 18 µL aliquot of both RNA types underwent
164 two rounds of DNA digestion using the DNA-*free*™ DNA Removal Kit (Invitrogen) which

165 provided rDNase. We completed the first round of DNA digestion following the manufacturer's
166 instructions, but skipped the DNase inaction step. To ensure no DNA carryover, we completed a
167 second round of DNA digestion as follows: 2.7 µL of DNase I Buffer and 1 µL rDNase were
168 added and pipette mixed into each sample and the plate was gently vortexed, centrifuged, and
169 then incubated at 37 °C for 20 minutes. Following incubation, 2 µL of DNase Inactivation
170 Reagent was added to each well. The plate was incubated for 2 minutes at room temperature, and
171 centrifuged at 3700 rpm for 5 minutes to pellet the DNase Inactivation Reagent. The supernatant
172 containing RNA was transferred to individual sterile microcentrifuge tubes and stored at -80 °C
173 until library preparation.
174
175 *RNA-seq library preparation and sequencing*
176 We prepared libraries for whole transcriptome shotgun sequencing (RNA-seq) using the
177 Illumina Stranded Total RNA Prep with Ribo-Zero Plus kit (rRNA depletion) and IDT® for
178 Illumina® RNA UD Indexes Set A, following the manufacturer's instructions. Equal volumes of
179 RNA (post 2 rounds of DNA digestion) from each technical replicate were pooled and used as
180 input material (250 ng and 26.4-65.56 ng for oRNA and eRNA, respectively) to prepare one
181 biological replicate library per tank. During library amplification, we used 12 and 17 PCR cycles
182 for oRNA and eRNA, respectively, which followed the instructions outlined by the manufacturer
183 for the respective RNA input amounts. Library quantification, quality control and equimolar pool
184 sequencing of the 16 libraries was conducted at the McGill Genome Centre on one Illumina
185 NovaSeq 6000 S4 lane using paired-end 100 bp reads.
186
187 *Contamination prevention*

188 We submerged all experimental and sampling equipment in a 20% bleach solution for 15
189 minutes and rinsed thoroughly five times with distilled water prior to use. All eRNA samples
190 were processed in a pre-PCR clean laboratory dedicated to low quality environmental and
191 ancient DNA/RNA samples. Prior to entering the pre-PCR sterile laboratory, researchers entered
192 a decontamination room to change into the dedicated clean lab coats and put on face masks,
193 hairnets and lab clogs with shoe covers. The laboratory workbench was soaked in a 20% bleach
194 solution for 10 minutes and RNase WiPER™ (iNtRON Biotechnology) and 20% bleach was
195 used to thoroughly wipe pipettes, vortex mixers and centrifuges. Aerosol filter barrier pipette tips
196 were used to prevent cross-contamination of samples. Filtration negative control samples and
197 molecular negative control samples (consisting of reagents or DNA/RNA free molecular grade
198 water) were processed at each step alongside eRNA samples, and were free of contaminating
199 nucleic acids as identified by failed PCR amplification using universal COI primers known to
200 amplify *Daphnia* DNA (Leray et al., 2013). Successful DNA digestion of each eRNA and oRNA
201 sample was verified by failed PCR amplification of RNA post-DNA digestion x2 using universal
202 COI primers known to amplify *Daphnia* DNA (Leray et al., 2013). RNA-seq libraries were
203 prepared from filtration and molecular negative control samples and were free of contamination
204 as indicated by quality control checks and failed amplification conducted at the McGill Genome
205 Centre.
206
207 *Bioinformatic pipeline and statistical approaches*
208 Raw FASTQ files underwent initial sequencing quality inspection using FastQC
209 (Andrews, 2010). Low quality sequences and adapters were removed using Trimmomatic
210 (Bolger et al., 2014) (ILLUMINACLIP: lluminaStrandedTotalRNA_adapter.fa:2:30:15

211 TRAILING:30 HEADCROP:1 MINLEN:90). All statistical analyses were conducted using R (R
212 Core Team, 2021, R version 4.1.2). With default parameters, STAR (Dobin et al., 2013) was
213 used to map all eRNA and oRNA reads that passed quality control and trimming to the *Daphnia*
214 *pulex* reference genome (Z. Ye et al., 2017). FeatureCounts (Liao et al., 2014) was used on the
215 BAM file output from STAR to quantify gene counts. Only those genes with a sum of ten or
216 more counts in either 20 or 28 °C were considered detected and used in the subsequent
217 differential expression (DE) analyses, for both eRNA and oRNA.

218 DE analysis was conducted with DESeq2 (Love et al., 2014), comparing all 20 °C
219 samples versus all 28 °C samples, for both eRNA and oRNA samples. To determine genes as
220 significantly differentially expressed, we used a false discovery rate (FDR) adjusted p-value $<$
221 0.05. A Pearson's chi-squared test with Yates' continuity correction was conducted to test for an
222 association between the significantly differentially expressed genes identified in eRNA and
223 oRNA. WEGO 2.0 (J. Ye et al., 2018) was used to classify GO terms (acquired from the
224 reference genome) to all annotated *D. pulex* DEGs. Gene ontology enrichment analysis was
225 conducted using topGO (Alexa & Rahnenfuhrer, 2022) by comparing GO terms of the relevant
226 genes to the genomic background. A FDR corrected p-value < 0.05 indicated significant
227 enrichment. A principal component analysis (PCA) of the regularized log (rlog)-normalized
228 eRNA and oRNA gene expression profiles (300 most variant genes) at both optimal and heat
229 stressed conditions was conducted using DESeq2's plotPCA function (Love et al., 2014).
230 Heatmaps showing the relative expression (Z-score calculated for each gene) of all significantly
231 differentially expressed genes between 20 and 28 °C, in both *D. pulex* oRNA and eRNA were
232 generated using pheatmap (Kolde, 2019).

233 We used the SqueezeMeta metatranscriptomics pipeline for the eRNA community-wide
234 analysis (Tamames & Puente-Sánchez, 2019). Briefly, metatranscriptomes were assembled in
235 sequential mode and contigs were assembled using Megahit (Li et al., 2015). Diamond was used
236 to align all contigs to the GenBank nr and KEGG databases for taxonomic assignment and
237 KEGG ID annotation, respectively (Buchfink et al., 2015; Kanehisa & Goto, 2000). We
238 performed differential gene expression analyses as described above for *Daphnia*, after we
239 mapped, using Kallisto (Bray et al., 2016), eRNA reads to the reference genomes and
240 transcriptomes of species known a priori to persist in the communities (*A. falcatus, S.*
241 *quadricauda,* and *P. subcapitata*) and highly abundant species identified by SqueezeMeta
242 (*Brachionus plicatilis*, *Volvox cateri*, *Stylonychia lemnae*, *Chlamydomonas eustigma*)
243 (Aeschlimann et al., 2014; Han et al., 2019; Hirooka et al., 2017; Nag Dasgupta et al., 2018;
244 Prochnik et al., 2010; Schomaker & Dudycha, 2021; Suzuki et al., 2018). We used the SQMtools
245 R package to further analyze the SqueezeMeta data and created stacked-bar plots to represent the
246 most relatively abundant taxa (Puente-Sánchez et al., 2020). We conducted non-metric
247 multidimensional scaling analysis of all KEGG IDs to visualize the community-wide functional
248 profiles.
249
250 **Results**
251 Daphnia pulex *gene detection and expression profiles*
252 All *Daphnia* reads within environmental RNA (eRNA) samples are of extra-organismal
253 origin as we pre-filtered tank water at 60 μm to remove *Daphnia*. Sequencing eRNA from tank
254 water and organismal RNA (oRNA) from *Daphnia* tissue detected 3,919 (21%) and 17,244
255 (93%) genes, respectively, of the 18,449 genes that compose the *Daphnia pulex* reference

256 genome (Z. Ye et al., 2017). All *Daphnia* genes detected in eRNA were also detected in oRNA.
257 A principal component analysis of the regularized log normalized expression for the 300 most
258 variable genes showed separation between experimental groups (Fig. 1). The samples were
259 separated across the first principal component (PC1) by RNA type and across the second
260 principal component (PC2) by temperature, clustering eRNA, oRNA, 20 °C and 28 °C within
261 their respective groups. The only exception was 20 °C eRNA sample T1, which unexpectedly
262 clustered closely with the 28 °C eRNA group. PC1 and PC2 accounted for 64% of explained
263 variance of all principal components (Fig. S1).
264
265 Daphnia pulex *differential gene expression*
266 Comparison of gene expression profiles at 20 °C and 28 °C identified 32 significantly
267 differentially expressed genes (DEG; false discovery rate adjusted p-value <0.05) from eRNA (4
268 upregulated and 28 downregulated) and 2351 DEGs from oRNA samples (1105 upregulated and
269 1246 downregulated). All eRNA significant DEGs were substantially up and down regulated
270 (exceeding Log_2 2 fold change), as well as 41 and 202 oRNA DEGs, respectively. The same
271 directional change (up/down regulated) was observed in 31/32 eRNA DEGs as in oRNA (Table
272 S1). Of all eRNA DEGs, 17 (2 upregulated and 15 downregulated) were also significantly
273 differentially expressed in oRNA (Fig. S2). There was a significant association between DEGs
274 identified in oRNA and DEGs identified in eRNA ($\chi 2$ = 18.26, df = 1, p-value = 1.92e-05). A
275 heatmap comparing the relative expression (Z-score) of the 17 *D. pulex* DEGs common to eRNA
276 and oRNA revealed four distinct clusters based on hierarchical clustering analyses, separating
277 samples and genes by temperature and directional change, respectively (Fig. 2). Across both
278 temperature conditions, eRNA and oRNA clustered together and were not separated into RNA

279 types, indicating that eRNA and oRNA exhibit similar levels of relative gene expression for
280 these DEGs. The only exception was eRNA 20 °C T1 sample which, similar to the PCA,
281 unexpectedly clustered with the 28 °C samples.
282
283 Daphnia pulex *functional annotation*
284 Gene ontology (GO) enrichment analysis of eRNA detected genes identified 50 GO terms
285 as enriched (FDR-corrected p-value <0.05), with 26, 6, and 18 belonging to Biological Processes
286 (BP), Cellular Component (CC) and Molecular Function (MF) domains, respectively (Figs. 3 and
287 S3). Of the 32 significant *D. pulex* DEGs identified from eRNA, 21 had GO annotations from the
288 reference genome (Z. Ye et al., 2017) (Fig. S4-5). GO enrichment analysis of highly DE
289 (exceeding Log_2 2 fold change) eRNA genes, and DEGs common to both eRNA and oRNA,
290 identified structural constituent of cuticle and chitin metabolic process, as enriched, respectively
291 (FDR-corrected p-value <0.05; Fig. 3).
292 GO enrichment analysis of oRNA detected genes identified 114 GO terms as enriched
293 (FDR-corrected p-value <0.05), with 100, 1, and 13 belonging to BP, CC and MF domains,
294 respectively (Figs. 3 and S3). Of the 2351 DEGs identified from oRNA, 1686 had GO
295 annotations from the reference genome (Z. Ye et al., 2017) (Figs. S4-5). GO enrichment analysis
296 of highly DE (exceeding Log_2 2 fold change) oRNA genes identified 16 GO terms as enriched
297 (FDR-corrected p-value <0.05), including structural constituent of cuticle, chitin metabolic
298 process, proteolysis, negative regulation of hydrolase activity, negative regulation of proteolysis
299 and response to stimulus (Fig S3).
300

301 *Community composition, differential gene expression and functional annotation*

302 eRNA reads originated from a simple mock freshwater community that we constructed,
303 consisting of *Daphnia pulex* as well as three phytoplankton species and opportunistic
304 microorganisms that colonized the artificial lake water. All non-*Daphnia* eRNA reads likely
305 originate from whole microorganisms that were captured on the filter. We used the SqueezeMeta
306 pipeline (Tamames & Puente-Sánchez, 2019) to align eRNA reads to the GenBank nr database
307 using its LCA algorithm, which identified 434 taxa. The community composition and relative
308 abundances remained similar between 20 °C and 28 °C conditions (Fig. S6). Across eRNA
309 samples, an average of 36% of all reads were aligned to eight species including four eukaryotes
310 that were a priori known to persist in the communities and four highly abundant taxa identified
311 by the SqueezeMeta analysis (Table S2). We detected 10,026 genes from these eight species,
312 with the majority originating from *D. pulex, Scenedesmus quadricauda and Ankistrodesmus*
313 *falcatus,* with 3919, 3180 and 1982 genes, respectively (Table S3). Differential gene expression
314 analysis between 20 °C and 28 °C eRNA samples was performed for each of the eight species
315 and identified a total of 121 significantly differentially expressed genes (FDR-adjusted p-value <
316 0.05). The highest numbers of DEGs were identified for *A. falcatus*, *D. pulex* and *S. quadricauda*
317 with 44, 32 and 20 DEGs respectively.

318 Within all eRNA reads, 12,211 KEGG IDs were identified and a non-metric
319 multidimensional scaling plot of these KEGG IDs revealed distinct functional profiles between
320 20 °C and 28 °C communities (Fig. S7). Of all KEGG IDs, 93 were identified as significantly
321 differentially expressed (DE; false discovery rate-adjusted p-value <0.05) between 20 °C and 28
322 °C (Table S4). This included KEGG IDs corresponding to proteins and pathways involved in

323 stress responses, such as the p53 signaling pathway, ubiquitin mediated proteolysis, metabolic
324 pathway and cyclin-dependent kinase 4 (Table S4).
325
326 **Discussion**
327 The use of environmental transcriptomics based on eRNA remained previously untested
328 despite recent evidence demonstrating that extra-organismal eRNA persists in the environment
329 long enough to be extracted (Kagzi et al., 2022; Littlefair et al., 2022; Marshall et al., 2021; Tsuri
330 et al., 2021; Wood et al., 2020). Our findings revealed that environmental transcriptomics is
331 sensitive in detecting transcriptional heat stress responses even without sampling source
332 organisms directly. From eRNA released by *Daphnia pulex* into the tank water, we detected
333 thousands of *D. pulex* genes and identified a subset of heat stress relevant genes to be
334 differentially expressed between the two temperature conditions, with levels of relative
335 expression similar to organismal RNA (oRNA). We also detected community wide changes in
336 functional profiles. Our study demonstrates the ability of environmental transcriptomics to reveal
337 gene expression responses of macroorganisms, and potentially complex biotic communities,
338 following environmental changes.
339
340 *Recovery of* Daphnia pulex *extra-organismal eRNA*
341 Several studies have recently demonstrated robust detection and persistence of extra-
342 organismal eRNA in the environment (Kagzi et al., 2022; Marshall et al., 2021; Wood et al.,
343 2020). We similarly were able to recover *Daphnia* eRNA in all samples, and had an average of
344 0.51% of all eRNA reads mapping to the *D. pulex* reference genome (Table S2). Our recovery of
345 extra-organismal eRNA reads was similar to metagenomic studies of natural aquatic ecosystems

346 that found between 0.27% and 1.25% of environmental DNA (eDNA) reads to be likely of extra-
347 organismal origin (Cowart et al., 2018; Monchamp et al., 2022; Stat et al., 2017; Székely et al.,
348 2021). We pre-processed our libraries with rRNA depletion, but could have potentially increased
349 our recovery of *Daphnia* eRNA reads by targeting eukaryotic RNA through poly(A) enrichment.
350 Nevertheless, our recovery of eRNA was particularly notable, given the general expectation that
351 within environmental samples, potentially low-abundance eRNA is 'competing' to be sequenced
352 with larger quantities of oRNA from whole microorganisms. This encouraging finding suggests
353 that most of the eRNA captured is in a protected state (e.g., cellular, vesicular, or in stable
354 molecular complexes) while free RNA forms, prone to degradation, are very short lived and
355 potentially undetectable.
356
357 Daphnia pulex *extra-organismal eRNA differential gene expression analysis*
358 Our broad environmental transcriptomics approach was able to detect 3,919 *D. pulex*
359 genes from eRNA, building upon targeted RT-PCR of eRNA which detected 6 fish genes (Tsuri
360 et al., 2021). As expected, we identified many more (2,351) differentially expressed genes
361 (DEGs) under temperature stress from *Daphnia* oRNA than eRNA (32). This discrepancy in
362 DEG identification can be easily attributed to large differences in the amount of RNA captured
363 that is specific to *Daphnia* among other organisms in the community. This is reflected in the
364 number of reads, 1 and 200 million, that mapped to the *D. pulex* reference genome using eRNA
365 and oRNA, respectively (Table S2). Having less than 1% of eRNA reads matching *D. pulex*
366 limited our power to detect DEGs, but is expected given that the libraries were generated from
367 diverse environmental samples which included RNA transcripts from over 400 taxa, whereas the
368 oRNA libraries were prepared from concentrated *D. pulex* tissue samples. On the other hand, we

369 believe this diversity of eRNA within water samples could be beneficial as it opens up exciting
370 possibilities of conducting community-wide surveys. For our proof-of-concept study, we
371 performed rRNA depletion and deep sequencing (average of 158 million reads per eRNA
372 sample, Table S2) in anticipation of such diverse eRNA samples to enable sufficient recovery of
373 *D. pulex* eRNA for gene expression analysis. Despite the lower abundance of *D. pulex* reads
374 from eRNA than oRNA, 97% of eRNA DEGs exhibited the same directional change (up/down
375 regulated) as in oRNA (Table S1). We found a significant association between the DEGs
376 identified in oRNA and those identified in eRNA, and similar levels of relative expression in
377 commonly shared DEGs (Fig 2). These results demonstrate that similar to conventional oRNA
378 tissue samples, our environmental transcriptomics approach yielded adequate sensitivity in
379 detecting changes in gene expression of a source organism in response to heat stress.
380 Changes in gene expression are often observed before phenotypic and demographic
381 effects occur and may therefore act as an early indication of ecological stress (Fedorenkova et
382 al., 2010; Jovic et al., 2017; Snell et al., 2003). We found that eRNA and oRNA recovered
383 similar *D. pulex* chronic heat stress responses which consisted of extensive downregulation of
384 genes associated with metabolic and cellular processes, membrane, binding, extracellular region,
385 collagen and cuticle structure (Figs. 3 and S4). These results are consistent with studies that also
386 found chronically heat stressed *Daphnia* to exhibit widespread downregulation of genes with
387 similar functions (Becker et al., 2018; Yampolsky et al., 2014). This widespread downregulation
388 of metabolic genes has been hypothesized to be a molecular compensatory mechanism for
389 *Daphnia* at near-lethal temperatures to sacrifice their long-term fitness for immediate survival
390 (Yampolsky et al., 2014).

391 After seven days of heat exposure, only few (4/62) heat shock genes were slightly
392 upregulated (log_2 fold change =1.1-1.3; Table S5) in *Daphnia* oRNA. This aligns with findings
393 suggesting heat shock genes to be less important for the chronic (days) *Daphnia* heat stress
394 response, but central for mitigating cell damage following acute (hours) heat exposure (Becker et
395 al., 2018; Yampolsky et al., 2014). Multiple tropical marine fish species similarly do not
396 differentially express heat shock genes during chronic heat stress (Veilleux et al., 2015),
397 suggesting that heat shock genes might not be optimal for assessments of communities under
398 chronic heat stress. Furthermore, we detected chitin related gene ontology (GO) terms to be
399 commonly enriched in both eRNA and oRNA DEGs (Fig. 3), corroborating *Daphnia* studies that
400 linked chitin genes to environmental stressors (Becker et al., 2018; Chain et al., 2019; Connon et
401 al., 2008; Shaw et al., 2007). Hydrolytic and proteolytic related GO terms were enriched in
402 eRNA (Fig. 3), and are commonly involved in *Daphnia* and copepod heat stress response by
403 stabilizing and cleaning up misfolded proteins (Becker et al., 2018; Kelly et al., 2017). It is
404 promising that using eRNA, we were able to detect the widespread downregulation of genes
405 associated with chronic heat stress with similar levels of relative expression as oRNA. We also
406 found that many GO terms were commonly shared in both RNA types, and that stress related
407 terms were enriched in eRNA. Collectively, our results indicate that environmental
408 transcriptomics based on eRNA is sensitive in detecting heat stress responses of source
409 organisms through changes in gene expression and functional GO terms. Future empirical work
410 is necessary to determine how our results extend to more complex natural ecosystems.

411

412 *Community composition and function*

413 We found the community composition to remain largely unchanged for 20 °C and 28 °C
414 conditions, after mapping all eRNA reads to GenBank nr (Fig. S6). Since we pre-filtered the tank
415 water at 60 μm to remove *Daphnia*, all non-*Daphnia* RNA recovered in our eRNA samples
416 likely originates from whole microorganisms that passed the filter. Within our eRNA samples,
417 gene expression analysis of eight species detected 10,026 genes with 121 of these being
418 significantly differentially expressed (DE) between 20 °C and 28 °C (Table S3). Although
419 community compositions remained similar, analysis of KEGG orthologs revealed distinct
420 functional profiles for communities in the two temperature treatments (Fig. S7). We found stress-
421 related KEGG orthologs to be DE, including proteins involved in binding, kinase and
422 oxidoreductase (Table S4). Our community wide results reflect similar patterns of differential
423 gene expression and shifts in functional terms observed in conventional aquatic
424 metatranscriptomic studies based on bulk microorganism samples (Aylward et al., 2015; Frias-
425 Lopez et al., 2008; Gilbert et al., 2008; Marchetti et al., 2012; Moniruzzaman et al., 2017;
426 Poretsky et al., 2005; Salazar et al., 2019; Vorobev et al., 2020). For example, a freshwater
427 metatranscriptome likely under heat stress exhibited similar stress responses which included
428 oxidoreductase, binding and kinase proteins (Trench-Fiol & Fink, 2020). We build upon this rich
429 metatranscriptomics literature by demonstrating that in addition to yielding microorganism
430 insights, environmental transcriptomics can also non-invasively detect the gene expression
431 response of macroorganisms to environmental stress.

432

433 *Recommendations and future directions*

434 Interest in non-lethal transcriptomics is growing, as such an approach increases animal
435 welfare, allows for repeated survey of individuals, and linking gene expression to fitness
436 following exposure to stress (Czypionka et al., 2015; Jeffries et al., 2021; Veldhoen et al., 2014).
437 However, conventional non-lethal methods are species specific and involve animal handling and
438 sampling for blood or tissue, which may still result in organism stress and reduced survival upon
439 release into the environment (Martins et al., 2018; Portz et al., 2006; Young et al., 2019).
440 Inferring the physiological state of organisms from non-invasive eRNA samples would
441 circumvent these stressful sampling procedures and represent a substantial improvement for
442 animal welfare. We demonstrated that similar to the conventional oRNA approach,
443 environmental transcriptomics based on eRNA can non-invasively detect heat stress responses of
444 progenitor organisms.

445 An exciting prospect of environmental transcriptomics is the potential to monitor the
446 ecological health of complex communities, including multiple macroorganisms. Biomonitoring
447 traditionally relies upon identifying bioindicator taxa that are correlated with environmental
448 conditions, but transcriptomic surveys can provide valuable early warnings since gene expression
449 changes are likely to occur before compositional shifts could be detected (Veilleux et al., 2021).
450 Conventional metatranscriptomics has demonstrated that gene expression profiling of
451 microorganisms reflects environmental conditions (Marchetti et al., 2012; Moran et al., 2013;
452 Salazar et al., 2019; Shi et al., 2011). Environmental transcriptomics extends beyond
453 microorganisms and has the potential to provide functional information of species across the
454 trophic chain as eRNA includes extra-organismal RNA from a diversity of macroorganisms.
455 However, further empirical work is required before such community-wide surveys should be

456 applied to biomonitoring efforts; population-level assessments that couple eRNA with targeted
457 approaches (e.g. qPCR, ddPCR) and stress biomarkers are more realistically applicable in the
458 short-term. Nevertheless, gaining insights into the ecological health of populations and
459 communities would enable conservation initiatives to be proactive in their efforts and respond
460 before local extinction events occur. Environmental transcriptomics is in its infancy and
461 refinement through mesocosm experiments and natural environment studies must first be
462 conducted before this approach can be applied further.

463 We anticipate multiple challenges for environmental transcriptomic biomonitoring in
464 natural environments, such as achieving adequate sequencing depth of eRNA in a cost-efficient
465 manner (Table 1). However, sequencing technologies are rapidly advancing and the costs
466 continue to decrease (Wetterstrand, 2021). The continuation of such technological advancement
467 will enable future studies to generate more reads less expensively, thereby increasing the
468 sensitivity of environmental transcriptomics to detect extra-organismal eRNA transcripts.
469 Additionally, the continued development of low input mRNA-seq library preparation kits will
470 enable researchers to prepare libraries for environmental transcriptomics with minimal template
471 eRNA. Molecular technological advancements that enhance extraction efficiency, enrich
472 eukaryotic RNA and remove ribosomal RNA from diverse samples may increase the yield of
473 coding eRNA and thus the power of environmental transcriptomics.

474 While sequencing and molecular advancements will resolve challenges associated with
475 eRNA depth, perhaps the largest remaining challenge for environmental transcriptomics to
476 overcome is the lack of a baseline to compare against. Differential gene expression analyses are
477 conducted by comparing a treatment against a control 'baseline' to identify genes involved with
478 the stress response. However, in natural ecosystems, it is difficult to obtain a control baseline due

479 to population demographics and multiple stressors acting simultaneously on a system. This is
480 similarly shared by metatranscriptomics, and multiple statistical and mathematical modelling
481 solutions have been developed to overcome this problem (Klingenberg & Meinicke, 2017;
482 Mallick et al., 2021; Zhang et al., 2021).

483 By providing functional insights in a non-invasive manner, eRNA based approaches
484 could push beyond the limitations of eDNA species detection and conventional tissue-based
485 transcriptomic surveys. We demonstrated that environmental transcriptomics can detect
486 thousands of genes, and that a subset of functionally relevant genes can be identified as
487 differentially expressed under heat stress, with levels of relative expression similar to the
488 conventional oRNA approach. Within all eRNA reads, we were able to reveal changes in
489 functional profiles for the community in response to heat stress. Collectively, these findings
490 demonstrate the ability of eRNA to detect changes in gene expression following environmental
491 stress and the potential of environmental transcriptomics to provide functional information
492 across the trophic chain. Future empirical work is necessary to determine if this approach can be
493 extended to natural ecosystems.
494

495

496 **Data accessibility:**

497 Sequences for all RNA-seq libraries have been uploaded to the NCBI SRA database and will be
498 made public upon acceptance of this manuscript. Reviewers can preview the uploaded files:
499 dataview.ncbi.nlm.nih.gov/object/PRJNA830892?reviewer=mcvpem4l5cupljoaj9jn53kg1e

500

501 **Author contributions:**

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

506

507 **Acknowledgements**

508 We thank Elif Irmak Bektas, Cathy Shen and Gabrielle Pellegrin for assisting with animal
509 husbandry and sample collection. We are grateful to Lars Iversen and Caren Helbing for
510 providing comments on an early version of this manuscript and Takahiro Maruki for input
511 regarding the bioinformatic analysis. This study was supported by the Natural Sciences and
512 Engineering Research Council (NSERC) of Canada through the Canada Research Chair program
513 and a Discovery Grant to Melania Cristescu.

514

515 **References**

516 Aeschlimann, S. H., Jönsson, F., Postberg, J., Stover, N. A., Petera, R. L., Lipps, H.-J., Nowacki,
517 M., & Swart, E. C. (2014). The draft assembly of the radically organized stylonychia
518 lemnae macronuclear genome. *Genome Biology and Evolution*, *6*(7), 1707–1723.
519 https://doi.org/10.1093/gbe/evu139
520 Akbarzadeh, A., Günther, O. P., Houde, A. L., Li, S., Ming, T. J., Jeffries, K. M., Hinch, S. G.,
521 & Miller, K. M. (2018). Developing specific molecular biomarkers for thermal stress in
522 salmonids. *BMC Genomics*, *19*(1), 749. https://doi.org/10.1186/s12864-018-5108-9
523 Alexa, A., & Rahnenfuhrer, J. (2022). *Topgo: Enrichment analysis for gene ontology* (2.46.0).
524 Bioconductor version: Release (3.14). https://doi.org/10.18129/B9.bioc.topGO
525 Alvarez, M., Schrey, A. W., & Richards, C. L. (2015). Ten years of transcriptomics in wild
526 populations: What have we learned about their ecology and evolution? *Molecular*
527 *Ecology*, *24*(4), 710–725. https://doi.org/10.1111/mec.13055
528 Amarasiri, M., Furukawa, T., Nakajima, F., & Sei, K. (2021). Pathogens and disease
529 vectors/hosts monitoring in aquatic environments: Potential of using eDNA/eRNA based
530 approach. *Science of The Total Environment*, *796*, 148810.
531 https://doi.org/10.1016/j.scitotenv.2021.148810
532 Andrews, S. (2010). *FastQC: a quality control tool for high throughput sequence data*.
533 https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
534 Aylward, F. O., Eppley, J. M., Smith, J. M., Chavez, F. P., Scholin, C. A., & DeLong, E. F.
535 (2015). Microbial community transcriptional networks are conserved in three domains at
536 ocean basin scales. *Proceedings of the National Academy of Sciences*, *112*(17), 5443–
537 5448. https://doi.org/10.1073/pnas.1502883112
538 Baillon, L., Pierron, F., Coudret, R., Normendeau, E., Caron, A., Peluhet, L., Labadie, P.,
539 Budzinski, H., Durrieu, G., Sarraco, J., Elie, P., Couture, P., Baudrimont, M., &
540 Bernatchez, L. (2015). Transcriptome profile analysis reveals specific signatures of
541 pollutants in Atlantic eels. *Ecotoxicology*, *24*(1), 71–84. https://doi.org/10.1007/s10646-
542 014-1356-x
543 Becker, D., Reydelet, Y., Lopez, J. A., Jackson, C., Colbourne, J. K., Hawat, S., Hippler, M.,
544 Zeis, B., & Paul, R. J. (2018). The transcriptomic and proteomic responses of Daphnia
545 pulex to changes in temperature and food supply comprise environment-specific and
546 clone-specific elements. *BMC Genomics*, *19*. https://doi.org/10.1186/s12864-018-4742-6
547 Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina
548 sequence data. *Bioinformatics (Oxford, England)*, *30*(15), 2114–2120.
549 https://doi.org/10.1093/bioinformatics/btu170
550 Bozinovic, F., & Pörtner, H.-O. (2015). Physiological ecology meets climate change. *Ecology*
551 *and Evolution*, *5*(5), 1025–1030. https://doi.org/10.1002/ece3.1403
552 Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq
553 quantification. *Nature Biotechnology*, *34*(5), Article 5. https://doi.org/10.1038/nbt.3519
554 Buchfink, B., Xie, C., & Huson, D. H. (2015). Fast and sensitive protein alignment using
555 DIAMOND. *Nature Methods*, *12*(1), Article 1. https://doi.org/10.1038/nmeth.3176
556 Celis-Salgado, M. P., Cairns, A., Kim, N., & Yan, N. D. (2008). The FLAMES medium: A new,
557 soft-water culture and bioassay medium for Cladocera. *SIL Proceedings, 1922-2010*,
558 *30*(2), 265–271. https://doi.org/10.1080/03680770.2008.11902123

Chain, F. J. J., Finlayson, S., Crease, T., & Cristescu, M. (2019). Variation in transcriptional responses to copper exposure across Daphnia pulex lineages. *Aquatic Toxicology*, *210*, 85–97. https://doi.org/10.1016/j.aquatox.2019.02.016

Chen, N., Huang, Z., Lu, C., Shen, Y., Luo, X., Ke, C., & You, W. (2019). Different transcriptomic responses to thermal stress in heat-tolerant and heat-sensitive pacific abalones indicated by cardiac performance. *Frontiers in Physiology*, *9*. https://doi.org/10.3389/fphys.2018.01895

Connon, R., Hooper, H. L., Sibly, R. M., Lim, F.-L., Heckmann, L.-H., Moore, D. J., Watanabe, H., Soetaert, A., Cook, K., Maund, S. J., Hutchinson, T. H., Moggs, J., Coen, W. D., Iguchi, T., & Callaghan, A. (2008). Linking molecular and population stress responses in daphnia magna exposed to cadmium. *Environmental Science & Technology*, *42*(6), 2181–2188. https://doi.org/10.1021/es702469b

Cordier, T., Lanzén, A., Apothéloz-Perret-Gentil, L., Stoeck, T., & Pawlowski, J. (2019). Embracing environmental genomics and machine learning for routine biomonitoring. *Trends in Microbiology*, *27*(5), 387–397. https://doi.org/10.1016/j.tim.2018.10.012

Cowart, D. A., Murphy, K. R., & Cheng, C.-H. C. (2018). Metagenomic sequencing of environmental DNA reveals marine faunal assemblages from the West Antarctic Peninsula. *Marine Genomics*, *37*, 148–160. https://doi.org/10.1016/j.margen.2017.11.003

Cristescu, M. E. (2019). Can environmental rna revolutionize biodiversity science? *Trends in Ecology & Evolution*, *34*(8), 694–697. https://doi.org/10.1016/j.tree.2019.05.003

Cristescu, M. E., & Hebert, P. D. N. (2018). Uses and misuses of environmental dna in biodiversity science and conservation. *Annual Review of Ecology, Evolution, and Systematics*, *49*(1), 209–230. https://doi.org/10.1146/annurev-ecolsys-110617-062306

Czypionka, T., Krugman, T., Altmüller, J., Blaustein, L., Steinfartz, S., Templeton, A. R., & Nolte, A. W. (2015). Ecological transcriptomics – a non-lethal sampling approach for endangered fire salamanders. *Methods in Ecology and Evolution*, *6*(12), 1417–1425. https://doi.org/10.1111/2041-210X.12431

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D. M., Vere, N. de, Pfrender, M. E., & Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, *26*(21), 5872–5895. https://doi.org/10.1111/mec.14350

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

Evans, T. G., & Hofmann, G. E. (2012). Defining the limits of physiological plasticity: How gene expression can assess and predict the consequences of ocean change. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1596), 1733–1745. https://doi.org/10.1098/rstb.2012.0019

Fedorenkova, A., Vonk, J. A., Lenders, H. J. R., Ouborg, N. J., Breure, A. M., & Hendriks, A. J. (2010). Ecotoxicogenomics: Bridging the gap between genes and populations. *Environmental Science & Technology*, *44*(11), 4328–4333. https://doi.org/10.1021/es9037287

Frias-Lopez, J., Shi, Y., Tyson, G. W., Coleman, M. L., Schuster, S. C., Chisholm, S. W., & DeLong, E. F. (2008). Microbial community gene expression in ocean surface waters. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(10), 3805–3810. https://doi.org/10.1073/pnas.0708897105

605 Fulda, S., Gorman, A. M., Hori, O., & Samali, A. (2010). Cellular stress responses: Cell survival
606 and cell death. *International Journal of Cell Biology*, *2010*, 214074.
607 https://doi.org/10.1155/2010/214074
608 Gilbert, J. A., Field, D., Huang, Y., Edwards, R., Li, W., Gilna, P., & Joint, I. (2008). Detection
609 of large numbers of novel sequences in the metatranscriptomes of complex marine
610 microbial communities. *PLoS ONE*, *3*(8). https://doi.org/10.1371/journal.pone.0003042
611 Gleason, L. U., & Burton, R. S. (2015). RNA-seq reveals regional differences in transcriptome
612 response to heat stress in the marine snail Chlorostoma funebralis. *Molecular Ecology*,
613 *24*(3), 610–627. https://doi.org/10.1111/mec.13047
614 Han, J., Park, J. C., Choi, B.-S., Kim, M.-S., Kim, H.-S., Hagiwara, A., Park, H. G., Lee, B.-Y.,
615 & Lee, J.-S. (2019). The genome of the marine monogonont rotifer Brachionus plicatilis:
616 Genome-wide expression profiles of 28 cytochrome P450 genes in response to
617 chlorpyrifos and 2-ethyl-phenanthrene. *Aquatic Toxicology*, *214*, 105230.
618 https://doi.org/10.1016/j.aquatox.2019.105230
619 Harley, C. D. G., Randall Hughes, A., Hultgren, K. M., Miner, B. G., Sorte, C. J. B., Thornber,
620 C. S., Rodriguez, L. F., Tomanek, L., & Williams, S. L. (2006). The impacts of climate
621 change in coastal marine systems. *Ecology Letters*, *9*(2), 228–241.
622 https://doi.org/10.1111/j.1461-0248.2005.00871.x
623 He, S., Wurtzel, O., Singh, K., Froula, J. L., Yilmaz, S., Tringe, S. G., Wang, Z., Chen, F.,
624 Lindquist, E. A., Sorek, R., & Hugenholtz, P. (2010). Validation of two ribosomal RNA
625 removal methods for microbial metatranscriptomics. *Nature Methods*, *7*(10), Article 10.
626 https://doi.org/10.1038/nmeth.1507
627 Hempel, C. A., Wright, N., Harvie, J., Hleap, J. S., Adamowicz, S. J., & Steinke, D. (2022).
628 Metagenomics versus total RNA sequencing: Most accurate data-processing tools,
629 microbial identification accuracy and perspectives for ecological assessments. *Nucleic*
630 *Acids Research*, *50*(16), 9279–9293. https://doi.org/10.1093/nar/gkac689
631 Hirooka, S., Hirose, Y., Kanesaki, Y., Higuchi, S., Fujiwara, T., Onuma, R., Era, A., Ohbayashi,
632 R., Uzuka, A., Nozaki, H., Yoshikawa, H., & Miyagishima, S. (2017). Acidophilic green
633 algal genome provides insights into adaptation to an acidic environment. *Proceedings of*
634 *the National Academy of Sciences*, *114*(39), E8304–E8313.
635 https://doi.org/10.1073/pnas.1707072114
636 Houde, A. L. S., Akbarzadeh, A., Günther, O. P., Li, S., Patterson, D. A., Farrell, A. P., Hinch, S.
637 G., & Miller, K. M. (2019). Salmonid gene expression biomarkers indicative of
638 physiological responses to changes in salinity and temperature, but not dissolved oxygen.
639 *Journal of Experimental Biology*, *222*(13). https://doi.org/10.1242/jeb.198036
640 Huang, W., Sher, Y.-P., Peck, K., & Fung, Y. C. B. (2002). Matching gene activity with
641 physiological functions. *Proceedings of the National Academy of Sciences*, *99*(5), 2603–
642 2608. https://doi.org/10.1073/pnas.042684399
643 Huey, R. B., & Berrigan, D. (2001). Temperature, demography, and ectotherm fitness. *The*
644 *American Naturalist*, *158*(2), 204–210. https://doi.org/10.1086/321314
645 Jeffries, K. M., Teffer, A., Michaleski, S., Bernier, N. J., Heath, D. D., & Miller, K. M. (2021).
646 The use of non-lethal sampling for transcriptomics to assess the physiological status of
647 wild fishes. *Comparative Biochemistry and Physiology Part B: Biochemistry and*
648 *Molecular Biology*, *256*, 110629. https://doi.org/10.1016/j.cbpb.2021.110629

Jo, T. S., Matsuda, N., Hirohara, T., & Yamanaka, H. (2022). *Simple and efficient preservation of fish environmental RNA in filtered water samples via RNAlater* [Preprint]. Research Square. https://doi.org/doi.org/10.21203/rs.3.rs-2170577/v1

Jo, T., Tsuri, K., Hirohara, T., & Yamanaka, H. (2022). Warm temperature and alkaline conditions accelerate environmental RNA degradation. *Environmental DNA*, *00*, 1–13. https://doi.org/10.1002/edn3.334

Jovic, K., Sterken, M. G., Grilli, J., Bevers, R. P. J., Rodriguez, M., Riksen, J. A. G., Allesina, S., Kammenga, J. E., & Snoek, L. B. (2017). Temporal dynamics of gene expression in heat-stressed Caenorhabditis elegans. *PLOS ONE*, *12*(12), e0189445. https://doi.org/10.1371/journal.pone.0189445

Kagzi, K., Hechler, R. M., Fussmann, G. F., & Cristescu, M. E. (2022). Environmental RNA degrades more rapidly than environmental DNA across a broad range of pH conditions. *Molecular Ecology Resources*, *22*(7), 2640–2650. https://doi.org/10.1111/1755-0998.13655

Kanehisa, M., & Goto, S. (2000). Kegg: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, *28*(1), 27–30. https://doi.org/10.1093/nar/28.1.27

Kelly, M. W., Pankey, M. S., DeBiasse, M. B., & Plachetzki, D. C. (2017). Adaptation to heat stress reduces phenotypic and transcriptional plasticity in a marine copepod. *Functional Ecology*, *31*(2), 398–406. https://doi.org/10.1111/1365-2435.12725

Klingenberg, H., & Meinicke, P. (2017). How to normalize metatranscriptomic count data for differential expression analysis. *PeerJ*, *5*, e3859. https://doi.org/10.7717/peerj.3859

Kolde, R. (2019). *pheatmap: Pretty Heatmaps* (1.0.12). https://CRAN.R-project.org/package=pheatmap

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., Boehm, J. T., & Machida, R. J. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. *Frontiers in Zoology*, *10*(1), 34. https://doi.org/10.1186/1742-9994-10-34

Lewin, H. A., Robinson, G. E., Kress, W. J., Baker, W. J., Coddington, J., Crandall, K. A., Durbin, R., Edwards, S. V., Forest, F., Gilbert, M. T. P., Goldstein, M. M., Grigoriev, I. V., Hackett, K. J., Haussler, D., Jarvis, E. D., Johnson, W. E., Patrinos, A., Richards, S., Castilla-Rubio, J. C., … Zhang, G. (2018). Earth BioGenome Project: Sequencing life for the future of life. *Proceedings of the National Academy of Sciences*, *115*(17), 4325–4333. https://doi.org/10.1073/pnas.1720115115

Li, D., Liu, C.-M., Luo, R., Sadakane, K., & Lam, T.-W. (2015). MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, *31*(10), 1674–1676. https://doi.org/10.1093/bioinformatics/btv033

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*(7), 923–930. https://doi.org/10.1093/bioinformatics/btt656

Littlefair, J. E., Rennie, M. D., & Cristescu, M. E. (2022). Environmental nucleic acids: A field-based comparison for monitoring freshwater habitats using eDNA and eRNA. *Molecular Ecology Resources*, *0*, 1–13. https://doi.org/10.1111/1755-0998.13671

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Mallick, H., Rahnavard, A., McIver, L. J., Ma, S., Zhang, Y., Nguyen, L. H., Tickle, T. L., Weingart, G., Ren, B., Schwager, E. H., Chatterjee, S., Thompson, K. N., Wilkinson, J. E., Subramanian, A., Lu, Y., Waldron, L., Paulson, J. N., Franzosa, E. A., Bravo, H. C., & Huttenhower, C. (2021). Multivariable association discovery in population-scale meta-omics studies. *PLOS Computational Biology*, *17*(11), e1009442. https://doi.org/10.1371/journal.pcbi.1009442

Marchetti, A., Schruth, D. M., Durkin, C. A., Parker, M. S., Kodner, R. B., Berthiaume, C. T., Morales, R., Allen, A. E., & Armbrust, E. V. (2012). Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability. *Proceedings of the National Academy of Sciences*, *109*(6), E317–E325. https://doi.org/10.1073/pnas.1118408109

Marshall, N. T., Vanderploeg, H. A., & Chaganti, S. R. (2021). Environmental (e)RNA advances the reliability of eDNA by predicting its age. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-82205-4

Martins, C. L., Walker, T. I., & Reina, R. D. (2018). Stress-related physiological changes and post-release survival of elephant fish (Callorhinchus milii) after longlining, gillnetting, angling and handling in a controlled setting. *Fisheries Research*, *204*, 116–124. https://doi.org/10.1016/j.fishres.2018.01.016

Miller, K. M., Günther, O. P., Li, S., Kaukinen, K. H., & Ming, T. J. (2017). Molecular indices of viral disease development in wild migrating salmon. *Conservation Physiology*, *5*(1), cox036. https://doi.org/10.1093/conphys/cox036

Miyata, K., Inoue, Y., Amano, Y., Nishioka, T., Yamane, M., Kawaguchi, T., Morita, O., & Honda, H. (2021). Fish environmental RNA enables precise ecological surveys with high positive predictivity. *Ecological Indicators*, *128*, 107796. https://doi.org/10.1016/j.ecolind.2021.107796

Monchamp, M.-È., Walsh, D. A., Garner, R. E., Kraemer, S. A., Beisner, B. E., Cristescu, M. E., & Gregory-Eaves, I. (2022). Comparative analysis of zooplankton diversity in freshwaters: What can we gain from metagenomic analysis? *Environmental DNA*, *0*, 1–15. https://doi.org/10.1002/edn3.335

Moniruzzaman, M., Wurch, L. L., Alexander, H., Dyhrman, S. T., Gobler, C. J., & Wilhelm, S. W. (2017). Virus-host relationships of marine single-celled eukaryotes resolved from metatranscriptomics. *Nature Communications*, *8*(1), 16054. https://doi.org/10.1038/ncomms16054

Moran, M. A., Satinsky, B., Gifford, S. M., Luo, H., Rivers, A., Chan, L.-K., Meng, J., Durham, B. P., Shen, C., Varaljay, V. A., Smith, C. B., Yager, P. L., & Hopkinson, B. M. (2013). Sizing up metatranscriptomics. *The ISME Journal*, *7*(2), Article 2. https://doi.org/10.1038/ismej.2012.94

Nag Dasgupta, C., Nayaka, S., Toppo, K., Singh, A. K., Deshpande, U., & Mohapatra, A. (2018). Draft genome sequence and detailed characterization of biofuel production by oleaginous microalga Scenedesmus quadricauda LWG002611. *Biotechnology for Biofuels*, *11*(1), 308. https://doi.org/10.1186/s13068-018-1308-4

Narum, S. R., & Campbell, N. R. (2015). Transcriptomic response to heat stress among ecologically divergent populations of redband trout. *BMC Genomics*, *16*(1), 103. https://doi.org/10.1186/s12864-015-1246-5

Poretsky, R. S., Bano, N., Buchan, A., LeCleir, G., Kleikemper, J., Pickering, M., Pate, W. M., Moran, M. A., & Hollibaugh, J. T. (2005). Analysis of microbial gene transcripts in environmental samples. *Applied and Environmental Microbiology*, *71*(7), 4121–4126. https://doi.org/10.1128/AEM.71.7.4121-4126.2005

Portz, D. E., Woodley, C. M., & Cech, J. J. (2006). Stress-associated impacts of short-term holding on fishes. *Reviews in Fish Biology and Fisheries*, *16*(2), 125–170. https://doi.org/10.1007/s11160-006-9012-z

Prochnik, S. E., Umen, J., Nedelcu, A. M., Hallmann, A., Miller, S. M., Nishii, I., Ferris, P., Kuo, A., Mitros, T., Fritz-Laylin, L. K., Hellsten, U., Chapman, J., Simakov, O., Rensing, S. A., Terry, A., Pangilinan, J., Kapitonov, V., Jurka, J., Salamov, A., … Rokhsar, D. S. (2010). Genomic analysis of organismal complexity in the multicellular green alga volvox carteri. *Science*. https://doi.org/10.1126/science.1188800

Puente-Sánchez, F., García-García, N., & Tamames, J. (2020). SQMtools: Automated processing and visual analysis of ’omics data with R and anvi’o. *BMC Bioinformatics*, *21*(1), 358. https://doi.org/10.1186/s12859-020-03703-2

Reid, A. J., Carlson, A. K., Creed, I. F., Eliason, E. J., Gell, P. A., Johnson, P. T. J., Kidd, K. A., MacCormack, T. J., Olden, J. D., Ormerod, S. J., Smol, J. P., Taylor, W. W., Tockner, K., Vermaire, J. C., Dudgeon, D., & Cooke, S. J. (2019). Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biological Reviews of the Cambridge Philosophical Society*, *94*(3), 849–873. https://doi.org/10.1111/brv.12480

Rodgers, E. M. (2021). Adding climate change to the mix: Responses of aquatic ectotherms to the combined effects of eutrophication and warming. *Biology Letters*, *17*(10), 20210442. https://doi.org/10.1098/rsbl.2021.0442

Salazar, G., Paoli, L., Alberti, A., Huerta-Cepas, J., Ruscheweyh, H.-J., Cuenca, M., Field, C. M., Coelho, L. P., Cruaud, C., Engelen, S., Gregory, A. C., Labadie, K., Marec, C., Pelletier, E., Royo-Llonch, M., Roux, S., Sánchez, P., Uehara, H., Zayed, A. A., … Sunagawa, S. (2019). Gene expression changes and community turnover differentially shape the global ocean metatranscriptome. *Cell*, *179*(5), 1068-1083.e21. https://doi.org/10.1016/j.cell.2019.10.014

Savary, R., Barshis, D. J., Voolstra, C. R., Cárdenas, A., Evensen, N. R., Banc-Prandi, G., Fine, M., & Meibom, A. (2021). Fast and pervasive transcriptomic resilience and acclimation of extremely heat-tolerant coral holobionts from the northern Red Sea. *Proceedings of the National Academy of Sciences*, *118*(19). https://doi.org/10.1073/pnas.2023298118

Schomaker, R. A., & Dudycha, J. L. (2021). De novo transcriptome assembly of the green alga Ankistrodesmus falcatus. *PLOS ONE*, *16*(5), e0251668. https://doi.org/10.1371/journal.pone.0251668

Schulte, P. M. (2015). The effects of temperature on aerobic metabolism: Towards a mechanistic understanding of the responses of ectotherms to a changing environment. *Journal of Experimental Biology*, *218*(12), 1856–1866. https://doi.org/10.1242/jeb.118851

Semmouri, I., Asselman, J., Van Nieuwerburgh, F., Deforce, D., Janssen, C. R., & De Schamphelaere, K. A. C. (2019). The transcriptome of the marine calanoid copepod Temora longicornis under heat stress and recovery. *Marine Environmental Research*, *143*, 10–23. https://doi.org/10.1016/j.marenvres.2018.10.017

Shakya, M., Lo, C.-C., & Chain, P. S. G. (2019). Advances and challenges in metatranscriptomic analysis. *Frontiers in Genetics*, *10*. https://doi.org/10.3389/fgene.2019.00904

Shaw, J. R., Colbourne, J. K., Davey, J. C., Glaholt, S. P., Hampton, T. H., Chen, C. Y., Folt, C. L., & Hamilton, J. W. (2007). Gene response profiles for Daphnia pulex exposed to the environmental stressor cadmium reveals novel crustacean metallothioneins. *BMC Genomics*, *8*, 477. https://doi.org/10.1186/1471-2164-8-477

Shi, Y., Tyson, G. W., Eppley, J. M., & DeLong, E. F. (2011). Integrated metatranscriptomic and metagenomic analyses of stratified microbial assemblages in the open ocean. *The ISME Journal*, *5*(6), Article 6. https://doi.org/10.1038/ismej.2010.189

Snell, T. W., Brogdon, S. E., & Morgan, M. B. (2003). Gene expression profiling in ecotoxicology. *Ecotoxicology*, *12*(6), 475–483. https://doi.org/10.1023/B:ECTX.0000003033.09923.a8

Stat, M., Huggett, M. J., Bernasconi, R., DiBattista, J. D., Berry, T. E., Newman, S. J., Harvey, E. S., & Bunce, M. (2017). Ecosystem biomonitoring with eDNA: Metabarcoding across the tree of life in a tropical marine environment. *Scientific Reports*, *7*(1), 12240. https://doi.org/10.1038/s41598-017-12501-5

Sunagawa, S., Coelho, L. P., Chaffron, S., Kultima, J. R., Labadie, K., Salazar, G., Djahanschiri, B., Zeller, G., Mende, D. R., Alberti, A., Cornejo-Castillo, F. M., Costea, P. I., Cruaud, C., d'Ovidio, F., Engelen, S., Ferrera, I., Gasol, J. M., Guidi, L., Hildebrand, F., … Bork, P. (2015). Structure and function of the global ocean microbiome. *Science*, *348*(6237), 1261359. https://doi.org/10.1126/science.1261359

Sunday, J. M., Bates, A. E., & Dulvy, N. K. (2011). Global analysis of thermal tolerance and latitude in ectotherms. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1713), 1823–1830. https://doi.org/10.1098/rspb.2010.1295

Suzuki, S., Yamaguchi, H., Nakajima, N., & Kawachi, M. (2018). Raphidocelis subcapitata (=Pseudokirchneriella subcapitata) provides an insight into genome evolution and environmental adaptations in the Sphaeropleales. *Scientific Reports*, *8*(1), 8058. https://doi.org/10.1038/s41598-018-26331-6

Székely, D., Corfixen, N. L., Mørch, L. L., Knudsen, S. W., McCarthy, M. L., Teilmann, J., Heide-Jørgensen, M. P., & Olsen, M. T. (2021). Environmental DNA captures the genetic diversity of bowhead whales (Balaena mysticetus) in West Greenland. *Environmental DNA*, *3*(1), 248–260. https://doi.org/10.1002/edn3.176

Tamames, J., & Puente-Sánchez, F. (2019). Squeezemeta, a highly portable, fully automatic metagenomic analysis pipeline. *Frontiers in Microbiology*, *9*. https://doi.org/doi.org/10.3389/fmicb.2018.03349

Trench-Fiol, S., & Fink, P. (2020). Metatranscriptomics from a small aquatic system: Microeukaryotic community functions through the diurnal cycle. *Frontiers in Microbiology*, *11*. https://doi.org/doi.org/10.3389/fmicb.2020.01006

Tsuri, K., Ikeda, S., Hirohara, T., Shimada, Y., Minamoto, T., & Yamanaka, H. (2021). Messenger RNA typing of environmental RNA (eRNA): A case study on zebrafish tank water with perspectives for the future development of eRNA analysis on aquatic vertebrates. *Environmental DNA*, *3*(1), 14–21. https://doi.org/10.1002/edn3.169

Veilleux, H. D., Misutka, M. D., & Glover, C. N. (2021). Environmental DNA and environmental RNA: Current and prospective applications for biological monitoring. *Science of The Total Environment*, *782*, 146891. https://doi.org/10.1016/j.scitotenv.2021.146891

Veilleux, H. D., Ryu, T., Donelson, J. M., Ravasi, T., & Munday, P. L. (2018). Molecular response to extreme summer temperatures differs between two genetically differentiated populations of a coral reef fish. *Frontiers in Marine Science*, *5*. https://doi.org/doi.org/10.3389/fmars.2018.00349

Veilleux, H. D., Ryu, T., Donelson, J. M., van Herwerden, L., Seridi, L., Ghosheh, Y., Berumen, M. L., Leggat, W., Ravasi, T., & Munday, P. L. (2015). Molecular processes of transgenerational acclimation to a warming ocean. *Nature Climate Change*, *5*(12), Article 12. https://doi.org/10.1038/nclimate2724

Veldhoen, N., Beckerton, J. E., Mackenzie-Grieve, J., Stevenson, M. R., Truelson, R. L., & Helbing, C. C. (2014). Development of a non-lethal method for evaluating transcriptomic endpoints in Arctic grayling (Thymallus arcticus). *Ecotoxicology and Environmental Safety*, *105*, 43–50. https://doi.org/10.1016/j.ecoenv.2014.03.030

von Ammon, U., Wood, S. A., Laroche, O., Zaiko, A., Lavery, S. D., Inglis, G. J., & Pochon, X. (2019). Linking Environmental DNA and RNA for Improved Detection of the Marine Invasive Fanworm Sabella spallanzanii. *Frontiers in Marine Science*, *6*. https://doi.org/10.3389/fmars.2019.00621

Voolstra, C. R., Valenzuela, J. J., Turkarslan, S., Cárdenas, A., Hume, B. C. C., Perna, G., Buitrago-López, C., Rowe, K., Orellana, M. V., Baliga, N. S., Paranjape, S., Banc-Prandi, G., Bellworthy, J., Fine, M., Frias-Torres, S., & Barshis, D. J. (2021). Contrasting heat stress response patterns of coral holobionts across the Red Sea suggest distinct mechanisms of thermal tolerance. *Molecular Ecology*, *30*(18), 4466–4480. https://doi.org/10.1111/mec.16064

Vorobev, A., Dupouy, M., Carradec, Q., Delmont, T. O., Annamalé, A., Wincker, P., & Pelletier, E. (2020). Transcriptome reconstruction and functional analysis of eukaryotic marine plankton communities via high-throughput metagenomics and metatranscriptomics. *Genome Research*, *30*(4), 647–659. https://doi.org/10.1101/gr.253070.119

Wetterstrand, K. (2021). *DNA Sequencing Costs: Data*. Genome.Gov. https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data

Wikelski, M., & Cooke, S. J. (2006). Conservation physiology. *Trends in Ecology & Evolution*, *21*(1), 38–46. https://doi.org/10.1016/j.tree.2005.10.018

Wood, S. A., Biessy, L., Latchford, J. L., Zaiko, A., von Ammon, U., Audrezet, F., Cristescu, M. E., & Pochon, X. (2020). Release and degradation of environmental DNA and RNA in a marine system. *Science of The Total Environment*, *704*, 135314. https://doi.org/10.1016/j.scitotenv.2019.135314

Yampolsky, L. Y., Zeng, E., Lopez, J., Williams, P. J., Dick, K. B., Colbourne, J. K., & Pfrender, M. E. (2014). Functional genomics of acclimation and adaptation in response to thermal stress in Daphnia. *BMC Genomics*, *15*(1), 859. https://doi.org/10.1186/1471-2164-15-859

Yates, M. C., Derry, A. M., & Cristescu, M. E. (2021). Environmental rna: A revolution in ecological resolution? *Trends in Ecology & Evolution*, *36*(7), 601–609. https://doi.org/10.1016/j.tree.2021.03.001

Ye, J., Zhang, Y., Cui, H., Liu, J., Wu, Y., Cheng, Y., Xu, H., Huang, X., Li, S., Zhou, A., Zhang, X., Bolund, L., Chen, Q., Wang, J., Yang, H., Fang, L., & Shi, C. (2018). WEGO 2.0: A web tool for analyzing and plotting GO annotations, 2018 update. *Nucleic Acids Research*, *46*(W1), W71–W75. https://doi.org/10.1093/nar/gky400

Ye, Z., Xu, S., Spitze, K., Asselman, J., Jiang, X., Ackerman, M. S., Lopez, J., Harker, B., Raborn, R. T., Thomas, W. K., Ramsdell, J., Pfrender, M. E., & Lynch, M. (2017). A new

reference genome assembly for the microcrustacean daphnia pulex. *G3: Genes|Genomes|Genetics*, *7*(5), 1405–1416. https://doi.org/10.1534/g3.116.038638

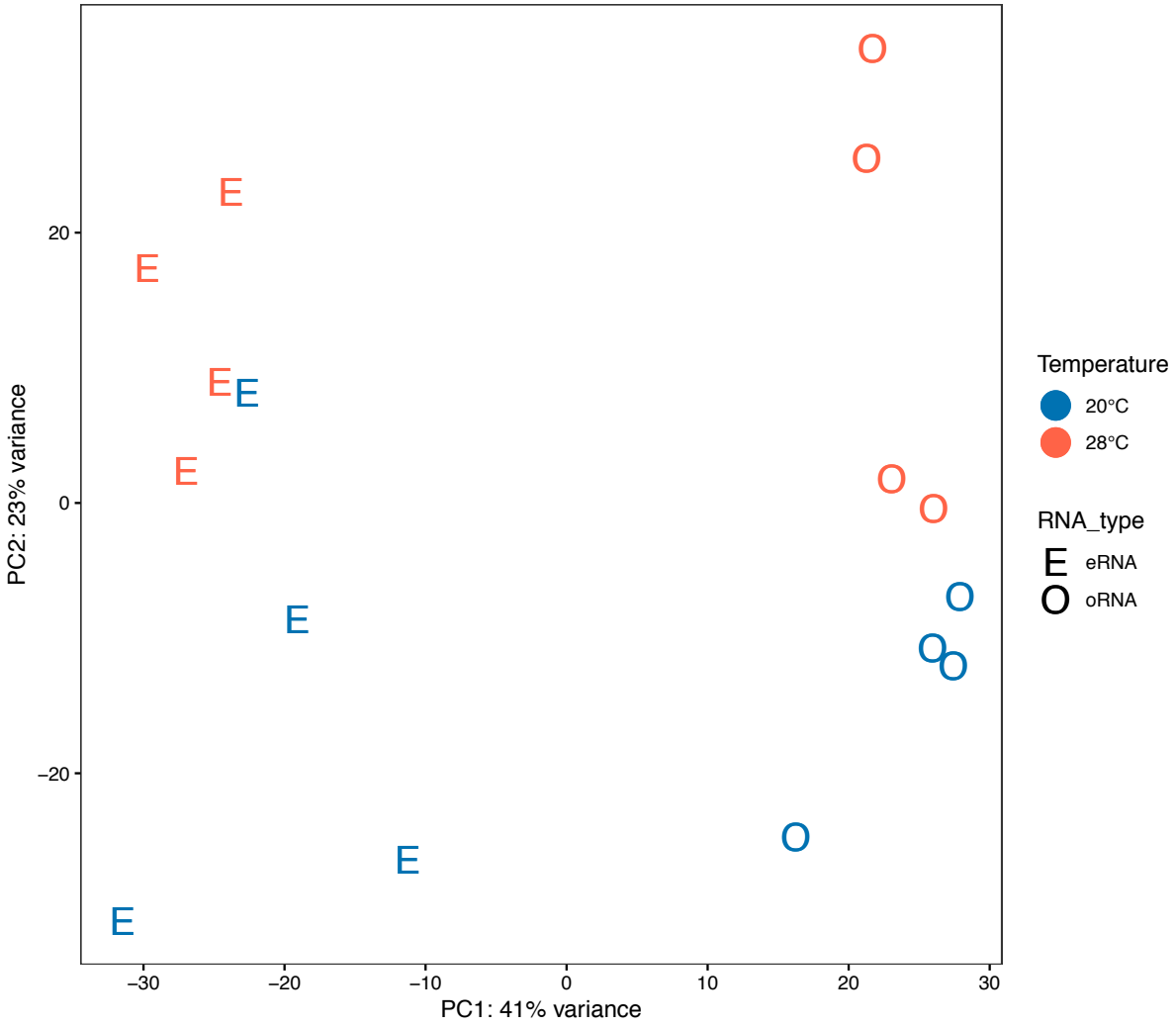
Young, T., Walker, S. P., Alfaro, A. C., Fletcher, L. M., Murray, J. S., Lulijwa, R., & Symonds, J. (2019). Impact of acute handling stress, anaesthesia, and euthanasia on fish plasma biochemistry: Implications for veterinary screening and metabolomic sampling. *Fish Physiology and Biochemistry*, *45*(4), 1485–1494. https://doi.org/10.1007/s10695-019-00669-8

Zhang, Y., Thompson, K. N., Huttenhower, C., & Franzosa, E. A. (2021). Statistical approaches for differential expression analysis in metatranscriptomics. *Bioinformatics*, *37*, i34–i41. https://doi.org/10.1093/bioinformatics/btab327

Zhao, S., Zhang, Y., Gamini, R., Zhang, B., & von Schack, D. (2018). Evaluation of two main RNA-seq approaches for gene quantification in clinical RNA sequencing: PolyA+ selection versus rRNA depletion. *Scientific Reports*, *8*(1), Article 1. https://doi.org/10.1038/s41598-018-23226-4

Zhao, W., He, X., Hoadley, K. A., Parker, J. S., Hayes, D. N., & Perou, C. M. (2014). Comparison of RNA-Seq by poly (A) capture, ribosomal RNA depletion, and DNA microarray for expression profiling. *BMC Genomics*, *15*(1), 419. https://doi.org/10.1186/1471-2164-15-419

895 **Figures and Tables**



896

897 **Fig. 1.** Principal component analysis (PCA) of the top 300 *Daphnia pulex* genes, as ranked by
898 variance. eRNA refers to extra-organismal RNA released by *D. pulex* into the environment,
899 whereas oRNA refers to organismal RNA extracted directly from *D. pulex* tissue. Prior to the
900 PCA, genes were filtered to include only those that had a sum ≥ 10 counts in either 20 °C or 28
901 °C. PCA revealed separation by RNA type and temperature across the PC1 and PC2,
902 respectively (except eRNA_20T1 clusters more closely with 28 °C samples than with other 20
903 °C samples).
904
905
906

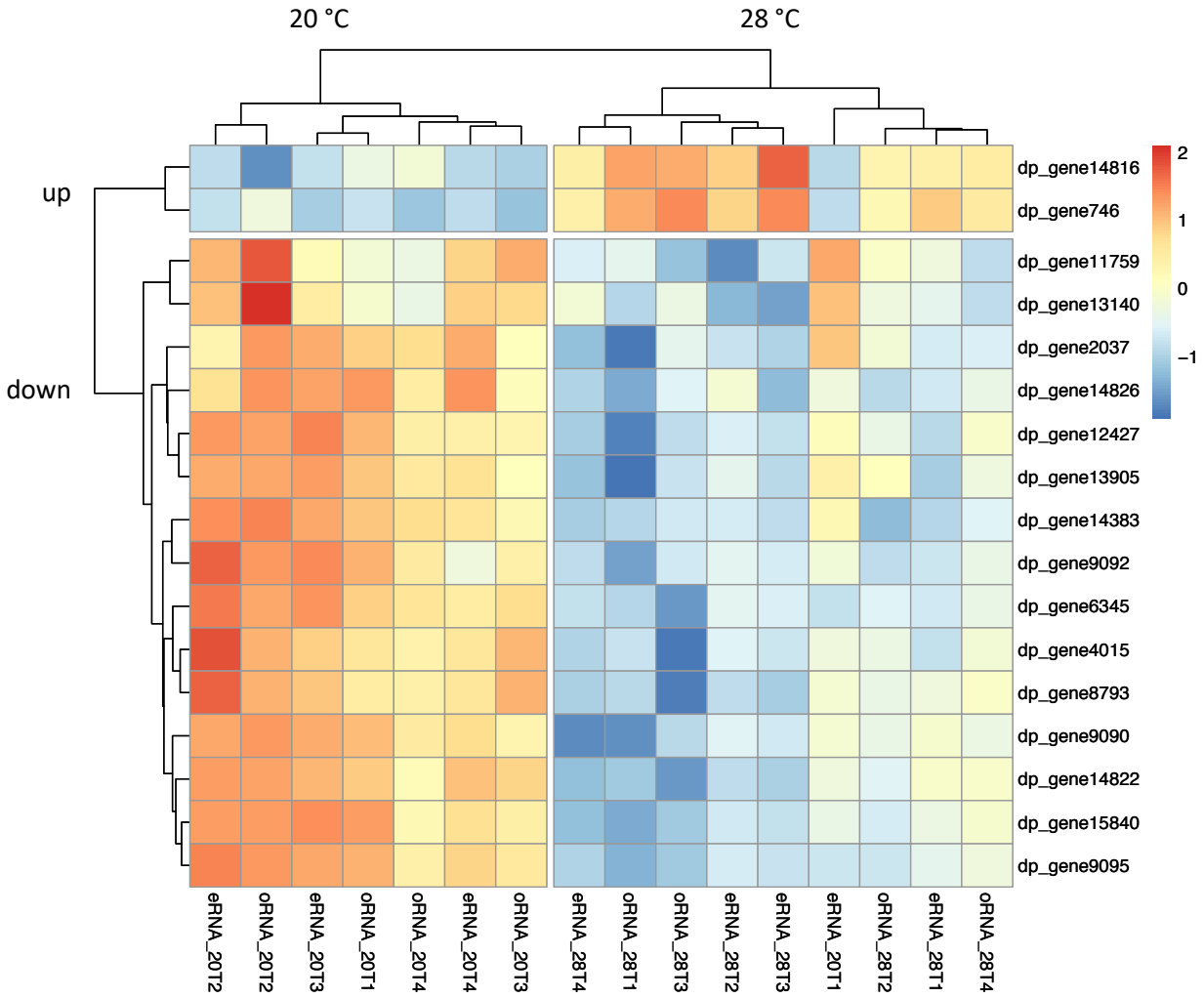


Fig. 2. Heatmap showing the relative expression (Z-score calculated for each gene) of all commonly significantly differentially expressed (DE) *Daphnia pulex* genes (false discovery rate adjusted p-value <0.05) between 20 °C and 28 °C, in both environmental RNA (eRNA) and organismal RNA (oRNA). DE analysis was conducted only for those genes that had ≥ 10 counts in either 20 °C or 28 °C samples. Hierarchical clustering analysis revealed four groups: 20 °C and 28 °C samples, and up and down regulated genes (except 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.

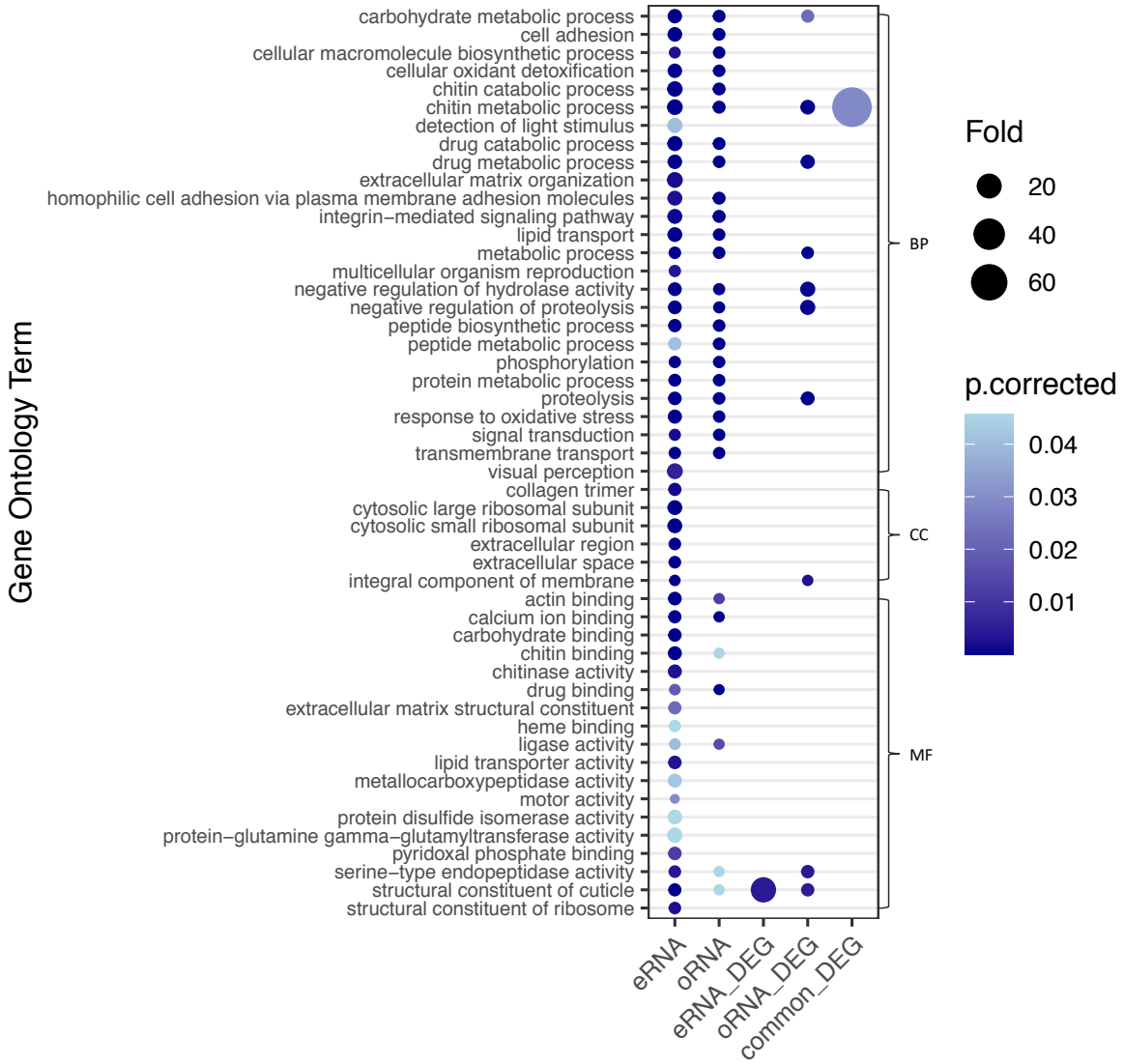


Fig. 3. Bubble plot of the Gene Ontology (GO) enriched terms associated with *Daphnia pulex* eRNA and oRNA detected genes, and eRNA, oRNA and commonly identified differentially expressed genes (DEG). The eRNA genes originate from extra-organismal *Daphnia* RNA. We used a false discovery rate (FDR) corrected p-value <0.05 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 for biomonitoring 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?)	(Cristescu, 2019)
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. Alternatively could potentially use preservative solution if freezing immediately is not possible. M: Keep samples on ice during extraction M: Minimize free-thaw cycles	(T. S. Jo et al., 2022; Kagzi et al., 2022; Marshall et al., 2021; Wood et al., 2020)
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	(He et al., 2010; Hempel et al., 2022; S. Zhao et al., 2018; W. Zhao et al., 2014)
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	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;	(Yates et al., 2021)

	optimized for organismal, not diverse environmental samples.	M: Amplify biomarker genes via targeted (qPCR, ddPCR) approaches to analyze specific biomarker genes	
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.	(Lewin et al., 2018; Shakya et al., 2019)
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	(Cordier et al., 2019; Fulda et al., 2010; Mallick et al., 2021; Yates et al., 2021; Zhang et al., 2021)
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	(Klingenberg & Meinicke, 2017; Salazar et al., 2019; Yates et al., 2021; Zhang et al., 2021)