1   **Variation in the transcriptome response and detoxification gene diversity drives pesticide**

2   **tolerance in fishes**

3   *In prep for Nature Communications*

4   Lawrence, M.J.[1*†], and Grayson, P.[1,2†], Jeffrey, J.D.[1,3], Docker, M.F.[1], Garroway, C.J.[1], Wilson,

5   J.M.[4], Manzon, R.G.[5], Wilkie, M.P.[4], & Jeffries, K.M.[1]

6

7   [1]Department of Biological Sciences, University of Manitoba, Winnipeg, MB, R3T 2N2, Canada.

8   [2]Department of Biomedical Informatics, Harvard Medical School, Boston, MA, 02115, USA

9   [3]Department of Biology, University of Winnipeg, Winnipeg, MB, R3B 2E9, Canada

10   [4]Department of Biology, Wilfrid Laurier University, Waterloo, ON, N2L 3C5, Canada

11   [5]Department of Biology, University of Regina, Regina, SK, S4S 0A2, Canada

12

13

14

15

16

17

18

19

20

21

22   *Corresponding author: m_lawrence27@live.ca
23   †Co-first authors

24   **Key words:** invasive species control, lampricides, interspecific variation, biotransformation,
25   bluegill, sea lamprey

26

27   **Abstract**

28   Pesticides are critical for invasive species management, but often have negative effects on non-

29   target native biota. Tolerance to pesticides should have an evolutionary basis, but this is poorly

30   understood. Invasive sea lamprey (*Petromyzon marinus*) populations in North America have

31   been controlled with a pesticide lethal to them at lower concentrations than native fishes. We

32   addressed how interspecific variation in gene expression and detoxification gene diversity confer

33   differential pesticide sensitivity in two fish species. We exposed sea lamprey and bluegill

34   (*Lepomis macrochirus*), a tolerant native species, to TFM, a pesticide commonly used in sea

35   lamprey control, and used whole-transcriptome sequencing of gill and liver to characterize the

36   cellular response. Comparatively, bluegill exhibited a larger number of detoxification genes

37   expressed and a larger number of responsive transcripts overall, which likely contributes to

38   greater tolerance to TFM. Understanding the genetic and physiological basis for pesticide

39   tolerance is crucial for managing invasive species.

40

41   Invasive species represent a considerable threat to local biodiversity and ecosystem

42   functioning, having significant economic costs associated with lost ecosystem services and

43   control efforts[1]. Invasive species control is multifaceted and serves primarily to reduce, but not

44   necessarily eradicate, populations of the target organism[2]. Often, chemical pesticides are used to

45   kill the target species while attempting to minimize non-target impacts[2]. In aquatic environments,

46   this can be difficult as pesticides are often applied to large portions of the waterbody, thus

47   exposing all community members[3]. However, species-specific pesticides are relatively rare and,

48   consequently, there will likely be impacts on native species in most pesticide application

49   programs[4].

50   The degree to which different species are adversely affected by pesticide applications can

51   be highly species-specific. Interspecific variation in pesticide sensitivities occurs in amphibians[5],

52   invertebrates[6], and fishes[7], which suggests that species-specific differences in pesticide uptake,

53   biotransformation, and elimination (i.e., the disposition) are likely driving differential

54   sensitivity[8]. Further, phylogenetic signatures associated with toxicity tolerances have been

55   identified in several clades of aquatic ectotherms[9]. This suggests that toxicity tolerances likely

56   have an evolutionary basis and are in part genetically determined[10]. Recent transcriptional

57   profiling of several teleost fishes has shown considerable interspecific variation amongst genes

58   involved in detoxification of pesticides and other xenobiotics[11], underscoring an important

59   genomic role in toxicity tolerances. Consequently, using molecular toolsets for understanding

60   pesticide responses are valuable in providing insight into phylogenetic variation in physiological

61   response and in predicting community impacts of pesticide applications[9,10]. This is relevant in

62   contemporary invasion control efforts, where minimizing harm to native species is often a key

63   objective[12]. To date, our understanding of the genetic and evolutionary underpinnings of

64   interspecific variation in toxicity tolerance is limited, hindering our insight into community-level

65   effects of pesticide applications.

66   The sea lamprey (*Petromyzon marinus*) control program in the Laurentian Great Lakes of

67   North America (hereafter referred to as the Great Lakes) is an excellent model system to explore

68   the evolutionary and genetic foundations of differential toxicity tolerance to a pesticide in target

69   (i.e., invasive sea lamprey) and non-target (native) species. Sea lamprey, although native to the

70   Atlantic Ocean, likely gained access to Lake Ontario through canals in the mid-1800s; the

3

71 Welland Canal subsequently allowed access to the remaining Great Lakes by the early 1900s[13].

72 By the 1950s, sea lamprey populations in the Great Lakes had exploded, decimating native

73 fisheries, notably lake trout (*Salvelinus namaycush*), having severe impacts on the environment

74 and economy of the region[13]. An international Great Lakes Fishery Commission (GLFC)

75 involving Canadian and American institutions was formed in 1955 to co-manage fisheries and to

76 control sea lamprey populations. Since the 1950s, the GFLC's primary method of sea lamprey

77 control has involved treating streams and rivers with pesticides, commonly referred to as

78 lampricides, that target burrowing sea lamprey larvae in Great Lakes tributaries[13]. Because the

79 filter-feeding larval stage lasts for ~4–5 years before sea lamprey metamorphose into parasitic

80 juveniles that disperse widely in the lakes, lampricide treatments in streams containing sea

81 lamprey larvae kill multiple year classes at once.

82 Two pesticides are used to control sea lamprey, 3-trifluoromethyl-4-nitrophenol (TFM)

83 and niclosamide (2',5-dichloro-4'-nitrosalicylanilide), each of which contains a halogenated

84 phenol ring that contributes to their toxicity[3]. In most treatments, TFM is applied at

85 concentrations reaching the sea lamprey 9-h $LC_{99.9}$ of TFM[14], the concentration needed to kill

86 99.9% of the exposed population, and is co-applied with 1-2% niclosamide because these

87 compounds interact in a greater than additive fashion[15]. While niclosamide is broadly toxic to

88 fishes[3], TFM is highly toxic to lampreys relative to jawed fishes. The mode of action of TFM is

89 interference with oxidative phosphorylation in the mitochondria[16], which lowers aerobic ATP

90 production forcing the animal to increasingly rely on unsustainable anaerobic glycolysis[17–19].

91 However, TFM sensitivity varies greatly among jawed fishes, which makes understanding the

92 broader non-target impacts of TFM difficult to resolve[3]. For example, lake sturgeon (*Acipenser*

93 *fulvescens*), a non-teleost jawed fish, are moderately sensitive to TFM[3], while centrarchids,

94 which includes basses and bluegill (*Lepomis macrochirus*), have some of the highest reported

95 tolerances to TFM amongst teleost fishes in the Great Lakes, with reported LC values ~10-fold

96 higher than sea lamprey[19].

97 Little is understood about the physiological mechanisms driving varying tolerances to

98 TFM among fishes[3]. Detoxification of TFM is believed to occur through a combination of Phase

99 I and Phase II biotransformation processes, which collectively act to enhance the water solubility

100 of the compound to improve elimination[3]. It is generally believed that UDP-

4

101   glucuronosyltransferase (Ugt) is the main enzyme responsible for TFM detoxification, via

102   conjugation of TFM through the addition of a UDP-glucuronic acid[20–22]. The greater sensitivity

103   of lampreys to TFM is believed to be related to lower Ugt activities compared to other fishes[20–

104   22], but the other cellular aspects of TFM detoxification are poorly characterized in fishes.

105   Furthermore, interspecific variation in detoxification gene diversity is believed to be important in

106   driving differences in toxicity tolerances among fishes[11], which could suggest that available *ugt*

107   genes impart differential TFM tolerances between sea lamprey and non-target fishes. Indeed,

108   compared to zebrafish (*Danio rerio*; N =40), sea lamprey have far fewer *ugt* genes (N= 2) in

109   their genome[23]. A few comparative studies have been conducted looking at sea lamprey and

110   rainbow trout (*Oncorhynchus mykiss*), however little attention has been paid to the physiological

111   effects of TFM exposure in other fishes[16–19]. As TFM detoxification primarily occurs in the

112   liver, and the gills are one of the primary uptake sites, TFM physiological studies often focus on

113   these two tissue types in fishes[3]. However, a broader characterization of TFM toxicity,

114   particularly the expression patterns of Phase I and Phase II enzyme genes, in fishes is necessary.

115   We used a comparative transcriptomics approach to evaluate the effects of TFM on sea

116   lamprey and bluegill, species with the lowest and highest reported tolerances to TFM,

117   respectively[19]. The purpose of this study was to: 1) characterize interspecific variation in

118   detoxification gene families expressed in the transcriptome following TFM exposure, 2) assess

119   the potential mechanisms of TFM detoxification common to both species, and 3) identify key

120   taxon-specific differences in the physiological response to TFM between sea lamprey and a

121   TFM-tolerant teleost fish. We predicted that bluegill would have a greater diversity of

122   detoxification mechanisms alongside a large robust transcriptomic detoxification response that

123   contributes to their greater tolerance to TFM relative to sea lamprey. We also predicted that the

124   molecular signatures of TFM toxicity would become more evident with increasing exposure

125   duration as detoxification systems become exhausted; effects that would be more pronounced in

126   the sea lamprey. These results will help to identify the evolutionary and genetic underpinnings of

127   pesticide tolerance, which may be used to inform improvements in invasive species control

128   efforts across a diversity of contexts.

129   **Results**

130   *Detoxification gene expression patterns*

5

131   Sea lamprey and bluegill were exposed to TFM for 6, 12, and 24 h at their species-

132   specific 24 h-LC_{10} (sea lamprey = 2.21 mg L^{-1}, bluegill = 22.06 mg L^{-1}) with gill and liver tissue

133   being sampled at 6, 12, and 24 h of exposure[19]. The mRNA from the gills and liver were

134   extracted and sequenced using a whole transcriptome approach for the identification of gene

135   targets associated with TFM detoxification and gene ontology (GO) term enrichment. As

136   variation in TFM tolerance is believed to stem from differences in biotransformation capacities

137   in fishes[3,21], we filtered for Phase I-III biotransformation transcripts from each species'

138   transcriptome. Proteins listed as being involved with detoxification, xenobiotic removal, and/or

139   organic compound breakdown were retained as these functions likely made them important in

140   TFM detoxification. Briefly, Phase I biotransformation generally includes the hydrolysis of

141   organic compounds while Phase II biotransformation involves conjugation of organics with both

142   processes enhancing xenobiotic solubility for easier elimination. Phase III biotransformation

143   typically involves facilitating the excretion of the xenobiotic[24].

144   Several Phase I detoxification transcripts were affected by TFM treatment in both

145   species. In sea lamprey, transcripts associated with Phase I detoxification processes were only

146   differentially expressed in the gills and consisted mainly of cytochrome P450s (six unique *cyp*'s;

147   Table 1). Notably, this included both *cyp1a1* (6 h) and *cyp1b1* (24 h) with a paraoxonase (*pon1*)

148   being the only non-*cyp* transcript differentially expressed (Table 1). Most *cyp* transcripts were

149   upregulated in response to TFM exposure at 6 or 24 h, but *pon1*, *cyp2j2*, and *cyp2d3* were

150   downregulated at 24 h. Contrastingly, bluegill demonstrated a larger induction of Phase I

151   detoxification transcripts (nine unique transcripts, 28 differentially expressed transcripts) in both

152   tissues across all exposure durations (Table 1). In bluegill gill, *cyp1a1* and *cyp1b1* were

153   upregulated across all timepoints (Table 1). In the bluegill liver, several *cyp* transcripts

154   responded positively to TFM treatment across exposure durations (e.g., *cyp1a1*, *cyp1b1*,

155   *cyp2c31*), but three transcripts (*cyp2b2*, *cyp2k1*, *cyp2k6*) were downregulated at 24 h (Table 1).

156   In the transcriptomes of both species, we identified several *ugt*'s likely involved in the

157   detoxification of TFM (Table 2). In sea lamprey, six clusters from the superTranscriptome

158   mapped to *ugt*'s. In contrast, the bluegill transcriptome had 17 clusters which mapped to *ugt*'s

159   (Table 2). There were also noticeable differences between sea lamprey and bluegill in expression

160   of the *ugt* transcripts in response to TFM exposure. Treatment with TFM did not result in any *ugt*

161   being differentially expressed in sea lamprey. Conversely, in bluegill, *ugt3* was upregulated in

162   the gill across all three exposure durations (Table 3) and in the liver at 12 and 24 h of exposure.

163   Interestingly, *ugt3* was not present in the lamprey transcriptome (Table 2). Hepatic *ugt2b9*

164   transcript levels were also elevated under TFM exposure at 12 and 24 h in bluegill liver, and

165   TFM exposure led to several *ugt* transcripts (i.e., *ugt1a2*, *ugt1a5*, *ugt2a1*) being downregulated

166   in the liver of bluegill at 12 and 24 h of TFM exposure (Table 3).

167   There was also a noticeable difference between bluegill and sea lamprey in the

168   expression patterns for other transcripts involved in Phase II detoxification (Table 4). In sea

169   lamprey, there were only three differentially expressed transcripts for Phase II genes, including

170   glutathione S-transferase (*gstt3*), sulfotransferase (*sult6b1*), and nicotinamide N-

171   methyltransferase (*nnmt*), in both liver and gill, all at 24 h of exposure (Table 4). In bluegill,

172   several sulfotransferases and a single N-acetyltransferase (*nat8*; Table 4) transcripts were

173   upregulated in the gill. In the bluegill liver, Phase II detoxification transcripts were only

174   differentially regulated at 24 h of TFM exposure with N-acetyltransferases (*nat2* and *nat8*) being

175   the only upregulated transcripts and two glutathione S-transferases (*gstt3* and *gstt1*), a thiopurine

176   S-methyltransferase (*tpmt*), and a sulfotransferase (*sult1c1*) being downregulated (Table 4).

177   Differential expression of Phase III detoxification transcripts was limited to 24 h of

178   exposure in sea lamprey. In the gills, five ATP-binding cassettes (*abc*; *abcc3*, *abccc5*, *abcb1*,

179   *abcb11*, and *abcg2*) were upregulated in response to TFM, whereas the two solute carrier family

180   transcripts (*slc*; *slc35d2* and *slc22a15b*) were downregulated (Table 5). In the lamprey liver, a

181   single *abc* and *slc* were the only transcripts differentially expressed in response to TFM (Table

182   5). In contrast, the gills of bluegill exhibited differential expression of Phase III transcripts across

183   all three timepoints consisting of four unique *abc*'s (*abca3*, *abcb1*, *abcb6*, and *abcc4*) and nine

184   unique *slc*'s (Table 5). In most instances, these branchial *abc*'s and *slc*'s were upregulated in

185   bluegill. In the bluegill liver, a single *abc* (*abcb6*) was found to be upregulated at 12 and 24 h of

186   TFM exposure (Table 5). Aside from *slc47a1*, all of the five-remaining hepatic *slc*'s were found

187   be upregulated under TFM treatment and with differential expression occurring at 12 and 24 h of

188   exposure (Table 5).

189   *Overall transcriptome patterns*

7

190   The overall number of differentially expressed transcripts and their timing were different

191   in the two species. In the sea lamprey, there was a delayed transcriptional response in the gills

192   following TFM exposure (Fig. 1a). At 6 and 12 h, we observed only 139 and 9 differentially

193   expressed superTranscriptome clusters, respectively (hereafter referred to as 'clusters').

194   However, by 24 h, there were 7,055 differentially expressed clusters. In contrast, bluegill gill

195   expression patterns were marked by a stepwise increase in the number of differentially expressed

196   superTranscriptome clusters. A total of 595 and 2,662 differentially expressed clusters were

197   already observed at 6 and 12 h of TFM exposure, respectively, although the total at 24 h (4,370)

198   was less than the 24-h total in sea lamprey gills (Fig. 1b).

199   The livers of sea lamprey showed a muted pattern of transcriptional activity under TFM

200   exposure with no changes in transcription within the first 12 h of exposure, and only 1,940

201   differentially expressed clusters by 24 h (Fig. 1c). In bluegill, there were moderate levels of

202   differential expression in the liver within the first 12 h of TFM exposure (388 and 1,464 at 6 and

203   12 h, respectively), and the number had increased to 5,685 differentially expressed clusters by 24

204   h of exposure (Fig. 1d).

205   *Lamprey GO term enrichment*

206   In sea lamprey gills, TFM treatment resulted in an enrichment of GO terms associated

207   with cellular growth and proliferation, immune function, and metabolism by 24 h of exposure.

208   Specifically, in transcripts that were upregulated, GO terms associated with cellular proliferation

209   and growth included regulation of cell proliferation, Ras protein signaling, regulation of I-κB

210   kinase/signaling, and regulation of the cell cycle (Fig. 2a). GO terms related to immune function

211   including viral life cycle, antigen processing, and negative regulation of type I interferon were

212   also enriched for upregulated transcripts in sea lamprey (Fig. 2a). An upregulation in GO terms

213   associated with energy metabolism, including "cellular response to insulin stimulus", and

214   "regulation of gluconeogenesis", also featured prominently in the sea lamprey gill (Fig. 2a). Of

215   the gluconeogenic transcripts that were enriched, we specifically observed an upregulation in

216   mitochondrial pyruvate dehydrogenase kinase 2 (*pdk2*) and glycerol-3-phosphate phosphatase

217   (*pgp*).

218   In sea lamprey liver, there were no differentially expressed transcripts at 6 h and 12 h of

219   TFM exposure (Fig. 1c). At 24 h, there was an enrichment of upregulated transcripts associated

8

220   with the GO term "positive regulation of transcription" (Fig. 2b). Within this parent term, cell

221   growth and survival GO terms were enriched, which included regulation of cell proliferation, fat

222   cell differentiation, autophagy, and ubiquitination. Additionally, processes related to energy

223   metabolism appeared to be upregulated as there was enrichment of transcripts associated with

224   "cellular response to insulin stimulus" (e.g., *igf1r*, *insr*, *pck2*, *pdk2*) and "negative regulation of

225   lipid storage" (e.g., *abca1*, *nfkbia*, *osbpl8*; Fig. 2b).

226   *Bluegill GO term enrichment*

227   In bluegill gills, there was enrichment of upregulated transcripts associated with cellular

228   growth, proliferation, and death. For example, at 6 h of TFM exposure, there was enrichment of

229   apoptotic processes regulation (Fig. 3a). By 12 h, this included a "positive regulation of cell

230   proliferation", "transforming growth factor β (TGF-β) responsiveness", "regulation of

231   autophagy", "regulation of apoptotic process", and "ubiquitin dependent processes" (Fig. 3a). By

232   24 h, enrichment was largely restricted to cellular death including GO terms related to apoptosis

233   and macroautophagy (Fig. 3a). Several transcripts associated with apoptosis, including *akt2*,

234   several *bmp* and *map3k* transcripts, and *casp3*, were upregulated under TFM. For downregulated

235   transcripts in the gill, there was also further enrichment of processes related to cellular growth

236   and death (Fig. 3b). At 6 h, we observed a negative regulation of several growth-related

237   processes including a "positive regulation of cell proliferation" and "Wnt signaling" (Fig. 3b).

238   By 12 h, there was an enrichment of downregulated transcripts associated with "positive

239   regulation of intracellular signal transduction" in the gill which included several growth-related

240   processes (Fig. 3b). By 24 h, the parent term "cytokine-mediated signaling pathway" was the

241   largest enriched parent term and included the downregulation of transcripts associated with

242   several GO terms related to cell growth and survival, including cell proliferation, apoptosis, and

243   G2/M mitotic transition (Fig. 3b).

244   The bluegill gill's response to TFM treatment involved expression changes of transcripts

245   associated with immune function. For upregulated transcripts, "negative regulation of

246   macrophage derived foam cell differentiation" (12 h) and "cellular response to transforming

247   growth factor beta stimulus" (12 and 24 h) were the main responses (Fig. 3a). However, there

248   was a large enrichment of GO terms associated with the immune response in downregulated

249   transcripts in the gill (Fig. 3b). At 12 h, cytokine signaling, and T cell migration were enriched.

9

250   By 24 h, the parent term "cytokine-mediated signaling pathway" included GO terms associated

251   with immune function including "inflammatory response" and NF-κβ signaling (Fig. 3b). There

252   was also an enrichment of downregulated transcripts associated with the parent terms "antigen

253   presentation via MHC class I" and "viral life cycle" (Fig. 3b).

254   TFM exposure also affected metabolic processes in bluegill gills. For upregulated

255   transcripts, this included enrichment of lipid metabolism (e.g., *abca1*, *cyp1a1*, *angptl4*; 12 h),

256   gluconeogenesis (e.g., *pdk2*, *pgp*, *ptpn2*; 12 h), reactive oxygen species (ROS; e.g., *cdkn1a*,

257   *slc25a33*, *thbs1*; 12 h), and glucose import (*insr*, *irs2*, *c1qtnf12*; 24 h). In downregulated

258   transcripts, enriched metabolic processes included a regulation of primary metabolic processes

259   and cellular amino acid metabolic processes (Fig. 3b). Tangentially, there was also an

260   enrichment of GO terms in upregulated transcripts that were associated with a response to an

261   organic compound (6h; Fig. 3a), which included the upregulation of several transcript targets,

262   notably cytochrome P450s (*cyp1a1* and *cyp1b1*), caspases (*casp3* and *casp7*), and an aryl

263   hydrocarbon receptor (*ahr*).

264   In the bluegill liver, TFM effects were generally restricted to transcripts pertaining to

265   cellular growth, proliferation, and cellular death. Specifically, in upregulated transcripts, there

266   was an enrichment of the parent GO terms "regulation of autophagy" (6 h), "positive regulation

267   of apoptotic process" (12 h), "macroautophagy" (24 h), and "ubiquitin-dependent protein

268   catabolism" (24 h; Fig. 4a). At 24 h, there was also enrichment of cell growth/death GO terms

269   for downregulated transcripts including regulation of programmed cell death and cell cycle

270   G2/M transitions (Fig. 4b). TFM also led to an enrichment of metabolic functions in the bluegill

271   liver. Specifically, glycogen metabolic processes were enriched in downregulated transcripts

272   which corresponded with lower differential expression of a glycogen debrancher enzyme (*agl*),

273   α-glucosidase (*gaa*), glycogen synthase (*gys2*), and glycogen phosphorylase (*pygl*; Fig. 4b).

Discussion

275   Variation in the availability and functional status of different detoxification pathways

276   plays a critical role in determining the interspecific and intraspecific sensitivity of animals to

277   different pesticides[10]. Accordingly, we predicted that interspecific variation in the expression and

278   diversity of transcripts coding for biotransformation enzymes would partially explain the

279   differences in TFM tolerances between sea lamprey and bluegill. As predicted, we identified

280   considerable interspecific variation in the transcriptomic responses to TFM exposure where all

281   three suites of biotransformation transcripts (i.e., Phase I-III) were more diverse and responsive

282   in bluegill relative to sea lamprey, suggesting a greater inherent ability to detoxify TFM. The

283   differences in the transcriptome responses between the species supports an evolutionary basis of

284   TFM tolerances in fishes. These results likely explain why >15-fold higher concentrations of

285   TFM are required to produce a toxic effect in bluegill versus sea lamprey[3], as well as the 10-fold

286   higher 24 h LC_{10} observed in bluegill compared to sea lamprey[19]. A functional transcriptomics

287   approach has therefore provided us with greater insight into the molecular basis for the different

288   sensitivities of these two species to TFM, and how transcriptomic responses associated with

289   detoxification genes explains these observations (Fig. 5).

290   *Evolutionary considerations in lampricide tolerances*

291   We found that differences in the expression and diversity of biotransformation transcripts

292   could help to explain interspecific variation in TFM sensitivity among target sea lamprey and

293   non-target teleost fishes. Specifically, differences in *ugt* transcript diversity and expression were

294   likely a major factor dictating TFM tolerances in fishes. Variation in detoxification capacities

295   have been exploited to improve the specificity of pesticides used across a broad range of

296   agricultural, commercial, and domestic settings, while minimizing impacts on non-target species.

297   For example, current agricultural practices make use of a combination of bioengineered crop

298   plants that are tolerant to pesticides (e.g., Roundup Ready™ [25]) as well as using taxon-specific

299   pesticides (e.g., 2,4-dichlorophenoxyacetic acid in cereal crops[26]) to selectively control

300   dicotyledons (broadleaf plants) without harming valuable monocotyledon crops. In contrast,

301   many chemical insecticides (e.g., organophosphates and neonicotinoids) are broad-spectrum

302   rather than selective pesticides, with detrimental non-target effects generally being reduced by

303   applying them in a selective manner (e.g., via spot treatments or in baits)[27]. Thus, furthering our

304   understanding of the genetic, physiological, and evolutionary mechanisms driving variation in

305   pesticide sensitivities can allow for the development of more effective and targeted pesticide

306   applications that minimize impacts on non-target species[2]. As pesticides used as lampricides

307   cannot be applied in this manner, understanding genetic variation in detoxification mechanisms

308   is key in protecting local biodiversity.

11

309   The sea lamprey control program in the Laurentian Great Lakes is perhaps the most

310   notable and successful example of pesticides use in an invasive vertebrate[3]. High specificity was

311   key in ensuring minimal impacts on non-target species as the pesticide was being administered

312   directly into the natural environment. During the 1950's, researchers conducted a large-scale

313   series of toxicity tests, screening over 5,000 different compounds for an agent that could kill sea

314   lamprey with higher selectivity than native species. Eventually, compounds containing a

315   nitrophenol group were identified as the most selective agents, with TFM being the best

316   substance[28]. Stream treatment with TFM has resulted in an estimated 90% reduction in sea

317   lamprey populations in the Great Lakes[3].

318   The exact mechanism underpinning the variation in TFM sensitivity among fishes has

319   remained illusive. Phylogenetic variation in TFM sensitivity was apparent as closely related

320   species (i.e., within order) demonstrated similarities in TFM toxicity values[3,19]. Notably, native

321   lampreys had comparable toxicity values (i.e., minimum lethal concentration 99% [MLC99]) to

322   sea lamprey and were also adversely affected by lampricide treatments[29], suggesting a taxonomic

323   basis for the sea lamprey's high sensitivity to TFM. Indeed, it is believed that Ugt's are the

324   principal enzyme in TFM detoxification as part of Phase II biotransformation [3,20,21,30]. In sea

325   lamprey, Ugt activities and numbers of *ugt* gene isoforms appear to be low relative to teleost

326   fishes[21–23], suggesting an underlying genetic basis in dictating TFM tolerances. The sea

327   lamprey's phylogenetic position as a basal vertebrate could help explain the disparity in TFM

328   tolerances. Lampreys diverged from the lineage leading to jawed vertebrates ~500 million years

329   ago, most likely after one round of whole genome duplication (WGD), with the second WGD

330   occurring only in jawed vertebrates[31,32] and a third WGD in teleosts[31]; these WGD events are

331   important in allowing the generation of new traits that can facilitate adaptation[31]. Consequently,

332   this provided teleosts with a more diverse set of detoxification genes relative to lampreys, and

333   perhaps even to non-teleost fishes like lake sturgeon, which are moderately sensitive to TFM

334   (although less than lampreys[3]).

335   *Patterns of ugt expression*

336   We show for the first time that there is a genetic basis underlying differences in TFM

337   tolerances among target and non-target freshwater vertebrates. Specifically, bluegill had a larger

338   diversity of *ugt*'s, compared to sea lamprey, and only bluegill showed differential expression of

12

339   *ugt*'s in their gill and liver transcriptomes due to TFM exposure. In line with our predictions, we

340   showed that differences in *ugt* expression patterns between bluegill and sea lamprey are likely

341   one of the main factors driving interspecific variation in TFM tolerance in fishes.

342   Bluegill exhibited differential expression in the *ugt* gene family including *ugt1*, *ugt2*, and

343   *ugt3*. In bluegill, *ugt3* appeared to be the most important *ugt* transcript in both tissues as it was

344   expressed at the highest levels and underwent the greatest fold-increase (upwards of 28x),

345   compared to *ugt1* and *ugt2*. *ugt3* genes have not been identified in fishes previously and was

346   only recently discovered in humans with their functional role remaining uncertain[33]. However,

347   the Ugt3 enzyme appears to be involved in detoxifying polyaromatic hydrocarbons, and 4-

348   nitrophenol in mammals[33], suggesting a role in TFM detoxification in bluegill as well. We also

349   noted an upregulation of *ugt2b9* in the bluegill liver. Thus, hepatic biotransformation using *ugt3*

350   and, to a lesser extent, *ugt2b9* may be a key factor in the high tolerance of bluegill to TFM

351   compared to sea lamprey and perhaps other teleost fishes[3].

352   *Insight into detoxification pathways using transcriptomics*

353   Variation in the detoxification transcriptome between bluegill and sea lamprey was

354   evident in genes associated with Phase I biotransformation, predominantly genes coding for

355   *cyp*'s, which can detoxify organic compounds[24]. These enzymes may have a role in TFM

356   detoxification as evident by the presence of Cyp-conjugated TFM metabolites in fishes[21], but

357   their relative importance compared to Phase II processes is uncertain. In both bluegill and sea

358   lamprey, *cyp1a1* and *cyp1b1* appeared to be the dominant *cyp* transcripts that responded to TFM,

359   which is unsurprising given that *cyp1* genes are involved in the detoxification of a diversity of

360   organic compounds[24]. Interestingly, bluegill had a far greater magnitude of response, as evident

361   by the high fold-changes in expression (upwards of ~187x), and a greater diversity of

362   differentially expressed *cyp* transcripts, compared to sea lamprey. However, it is unclear how

363   important Cyp's are for detoxifying TFM, as no studies have detected the corresponding Phase I

364   metabolites *in vivo*, as compared to the Phase II sulfate and glucuronide conjugates[21,22,34].

365   Further studies, both *in vitro* and *in vivo*, are required to better understand the functional role of

366   Cyp's in TFM biotransformation in both species.

367   Phase II biotransformation transcripts also showed a clear taxonomic divergence that may

368   help explain interspecific variation in TFM tolerances. Bluegill exhibited an upregulation of

369   sulfotransferases (N=3) and N-acetyltransferases (N =2) across both tissue types, whereas only a

370   single sulfotransferase and glutathione S-transferase responded in sea lamprey gill. As

371   sulfotransferases are important in TFM detoxification[3,21], the greater capacity to conjugate via

372   sulfation as well as through glucuronidation in bluegill likely supports high TFM metabolism

373   when compared to sea lamprey. The presence of N-acetyltransferases in bluegill also suggests

374   that TFM is undergoing acetylation[24], which has been characterized in bluegill previously[21]. Our

375   results suggest a taxonomic disparity in Phase II biotransformation pathways involved in TFM

376   detoxification, with bluegill able to upregulate several Phase II biotransformation processes to

377   effectively eliminate TFM compared with sea lamprey.

378   Interspecific variation was also evident in Phase III biotransformation processes, which

379   are mainly involved in the transport and elimination of toxicants[24]. We detected a divergent

380   effect of TFM exposure on Phase III transcripts where numbers of responding *abc* and *slc*

381   transcripts were greater in sea lamprey and bluegill, respectively. This might suggest a disparity

382   in TFM elimination routes given that Abc's are generally active transporters while Slc's are

383   bidirectional passive transporters[24]. In the lamprey, the high *abc* response, as well as the general

384   lack of transporter expression in the liver, may suggest a predominantly branchial-mediated

385   elimination[3] as any TFM excretion would be against a concentration gradient (i.e. high

386   environmental TFM concentration). In bluegill, the high environmental concentrations of TFM

387   ($\sim$21 mg L^{-1} [101 μM] in water, 23 nmol g^{-1} ww in liver[19]) coupled with the lower branchial *abc*

388   expression, relative to lamprey, suggests that TFM elimination can be through biliary/renal

389   excretion as previously reported in teleosts[3,30]. Together, this appears to suggest that bluegill

390   have a more responsive and effective system for TFM elimination.

GO term enrichment and tissue responses to TFM

392   In fishes, studies of TFM toxicity have largely been restricted to characterizations of

393   energy metabolism and mortality[17,19]. However, we found several biomarkers associated with

394   inhibited cell cycle progression and growth, and higher levels of ubiquitination and apoptosis, in

395   both bluegill and sea lamprey, suggesting that TFM's effects extend beyond simply uncoupling

396   oxidative phosphorylation[16,35], indicating a new mode of TFM toxicity in fishes. Increases in

14

397   ROS or oxidative damage are likely the primary mechanism by which mitochondrial uncouplers

398   exert tissue-level repression of growth[36]. While TFM effects on ROS generation have not been

399   quantified, TFM does increase mitochondrial respiration rates[16], which may serve to elevate

400   cellular ROS levels[37]. Coupled with the GO term enrichment patterns, our results suggest that

401   TFM-induced ROS is likely mediating cellular arrest as seen in niclosamide[38,39] and 2,4-

402   dinitrophenol (DNP, a TFM analogue)[40,41] exposures. While fish-specific examples are limited,

403   zebrafish (*Danio rerio*) exposed to 4-nitrophenol experienced reduced cellular growth and

404   proliferation, and heightened cell death[36], aligning with our results. At a broader scale, increases

405   in oxidative damage by ROS can impair cellular growth and cell cycle progression as well as

406   inducing apoptosis in fishes[42,43]. Together, our results suggest that TFM has a potential role in

407   generating oxidative damage in the cell.

408   Despite TFM toxicity primarily affecting energy metabolism[3], enrichment of these

409   metabolic GO terms was limited. In sea lamprey, there was a negative regulation of GO terms

410   associated with lipid storage, gluconeogenesis, and insulin responsiveness. Likely, these changes

411   were needed to increase energy mobilization for supporting heightened metabolic demands under

412   TFM exposure[17–19]. In bluegill, metabolic changes included positive regulation on

413   gluconeogenesis, glucose import, and lipid metabolism and negative regulation of glycogen

414   metabolism suggesting that increased energetic expenditure during TFM exposure is likely

415   occurring, probably due to increased rates of mitochondrial respiration and detoxification costs[19].

416   Also, as predicted, the effects of TFM on both species' transcriptomes was more apparent with

417   longer exposure durations (i.e., more differentially expressed transcripts) suggesting an

418   exhaustion of detoxification systems.

419   *Conclusions*

420   Here, we showed how transcriptome responses demonstrate interspecific variation in

421   pesticide sensitivities, which are critical in developing more targeted invasive species control.

422   We identified differences in *ugt* diversity and responsiveness to TFM, which appeared to be

423   linked to differences in TFM sensitivities between bluegill and sea lamprey highlighting the

424   importance of evolutionary history in mediating toxicity resilience. In the context of sea lamprey

425   control, this knowledge will be useful in developing a predictive framework for assessing

426   community-scale impacts of lampricides. More broadly, similar principles are likely to apply to a

427   wide range of systems where invasive species control efforts use pesticides extensively. By

428   identifying key genes involved in detoxification in native fauna, we can develop methods for

429   addressing the relative impacts that pesticide treatments have on a system. For TFM specifically,

430   this would likely include developing screenings of Ugt's, Cyp's, and Sult's as potential

431   biomarkers of exposure. This approach could help shed light on more subtle interspecific

432   differences in sensitivity to TFM even among teleosts. As invasive species control efforts seek to

433   provide more effective and targeted pesticide applications[13], furthering our understanding of the

434   toxicological mode of action and evolutionary variation in tolerances among species is likely to

435   greatly improve the use of pesticides.

436   **Methods**

437   *Animal care and holding*

438   Fish collection and holding conditions are presented in Lawrence et al.[19]. Briefly,

439   juvenile bluegill ($n = 200$; total length [TL] = 97.3 ± 0.88 mm; mass = 25.51 ± 0.75 g) were

440   sourced from Kinmount Fish Farm (Kinmount, ON, Canada) in September 2018. Bluegill were

441   transported to the animal holding facility at Wilfrid Laurier University (Waterloo, ON, Canada;

442   ~1000 L holding tank; ~ 0.5 L min^{-1}; T = 12–14°C; pH 8.1–8.2, alkalinity ~255 mg L^{-1} as

443   CaCO$_3$), where they were provided with a combination of commercial fish feed (EWOS #1,

444   Cargill, ON, Canada) and bloodworms daily.

445   Larval sea lamprey ($n = 568$; TL = 104.94 ± 0.69 mm; mass = 1.47 ± 0.03 g) were

446   electrofished from a tributary of Lake Huron by the United States Fish and Wildlife Service in

447   April 2018 and were temporarily stored at the Hammond Bay Biological Station (Millersburg,

448   MI, USA). Sea lamprey were then transported to Wilfrid Laurier University and held under

449   similar water conditions to the bluegill, although the lamprey tank was smaller (~100 L) and

450   contained a 8–10 cm deep layer of sand to facilitate natural burrowing[44]. Lamprey were fed on a

451   diet of baker's yeast weekly (see Lawrence et al. [19] for specific details). All experimental series

452   were conducted under approval from the Wilfrid Laurier University Animal Care Committee

453   (Animal Use Protocol No. R18001) under the guidelines established by the Canadian Council of

454   Animal Care.

455   *TFM exposures*

16

456   Larval sea lamprey and bluegill were exposed to either control conditions (i.e., no

457   toxicant) or field grade TFM (35% active ingredient dissolved in isopropanol; Clariant,

458   Griesheim, Germany) for up to 24 h. The TFM exposure concentrations were equivalent to the

459   species-specific concentration of TFM that was lethal to 10% of each species over 24 h (24-h

460   LC_{10}), previously determined for the cohorts of sea lamprey (24-h LC_{10} = 2.21 mg L^{-1}) and

461   bluegill (24-h LC_{10} = 22.06 mg L^{-1})[19]. Exposures were performed in triplicate with animals being

462   held in glass aquaria (10 L for sea lamprey larvae, 14 L for bluegill) in groups of 2–3 or 6

463   individuals per tank for bluegill and sea lamprey larvae, respectively. The TFM exposures were

464   static with temperature being maintained at the fish's acclimation temperature (~14–15°C).

465   Tissue sampling at 6 h, 12 h, and 24 h of exposure. At these discrete intervals, fish were netted

466   and placed one at a time into buffered tricaine methanesulfonate (MS-222; Syndel, Nanaimo,

467   BC, Canada; 1.5 g L^{-1} MS-222 with 3.0 g L^{-1} NaHCO₃) to euthanize the animals. The livers and

468   gill of both species were excised, and placed into RNA*later* (Invitrogen, ThermoFisher

469   Scientific, Mississauga, ON, Canada), held at 4°C for at least 24 h, and then stored at -80°C for

470   later RNA extraction.

471   *RNA extraction and transcriptome sequencing*

472   Extractions of total RNA were performed using a commercially available kit (RNeasy

473   Plus Mini Kit; Qiagen, Toronto, ON, Canada) according to the manufacturer's specifications.

474   Following extraction, an initial quality control check was performed using the 260/280

475   absorbance ratio and the 260/230 absorbance ratio with a NanoDrop One Microvolume UV-Vis

476   Spectrophotometer (ThermoFisher, Mississauga, ON, Canada). Quality was further assessed

477   using a Qubit RNA IQ assay (ThermoFisher, Mississauga, ON, Canada), read on a Qubit 4

478   fluorometer (ThermoFisher, Mississauga, ON, Canada), as a final check of RNA integrity.

479   Samples were then diluted to a standardized volume (50 ng µL^{-1}) and stored at -80°C until

480   shipment to the sequencing facility.

481   RNA-seq library preparation and sequencing were performed by the Centre d'Expertise

482   et de Services Génome Québec (Montreal, QC, Canada). Before sequencing, samples were again

483   verified for integrity using both spectrophotometry (NanoDrop ND 1000; ThermoFisher,

484   Mississauga, ON, Canada) and electrophoresis (Bioanalyzer 2100; Agilent, Santa Clara, CA,

485   USA; RIN ranges: sea lamprey, 8.4–10.0; bluegill 8.5–9.8). Sequencing of cDNA libraries was

486   achieved using the NovaSeq 6000 sequencing system (Illumina, Vancouver, BC, Canada). Total

487   sequencing read counts can be found in Table S1.

488   *Transcriptome assembly, alignment, and annotation*

489   RNA-seq QC

490   Raw reads were downloaded from Genome Quebec for processing and downstream

491   analysis on a University of Manitoba personal Linux Server, as well as supercomputers

492   belonging to Westgrid (Grex) and Compute Canada (Beluga, Cedar, and Graham). Fastqc (v

493   0.11.9) was run on the raw sequencing data to assess quality, and MultiQC

494   (https://github.com/ewels/MultiQC) was used to generate reports from the Fastqc for easier

495   viewing across samples[45,46]. It is important to note here that RNA-seq data files contained

496   sequencing data from fish exposed to TFM as well as those exposed to niclosamide or a mixture

497   of TFM and niclosamide. The latter two exposures are part of a complementary but ongoing

498   study and are not included in the final results here. Raw sequencing quality and duplication rates

499   met our expectations across all 148 sea lamprey samples and all 163 bluegill samples (combined

500   numbers of gill and liver tissue for each species that were obtained from the exposure series).

501   Adapters were trimmed from the raw sequences using Trimmomatic (v0.36), Fastqc and

502   MultiQC were run again to ensure successful trimming[47].

503   *De novo* transcriptome assembly

504   Since bluegill does not have a reference genome to map reads back to, *de novo*

505   transcriptome assembly was carried out for both bluegill and sea lamprey to keep methods

506   consistent[48]. These *de novo* assemblies were each constructed from a list containing a single

507   library for each tissue and timepoint. These input libraries were selected based on the highest

508   read count, accounting for duplication rates for each condition. For bluegill ($n = 26$), we used

509   one paired-end library from liver and gill for each of the following treatments and timepoints: 6,

510   12, and 24 h for niclosamide ($n = 6$), TFM ($n = 6$) and TFM:niclosamide mixture ($n = 6$), as well

511   as 0, 6, 12, and 24 h for the control ($n = 8$). For lamprey ($n = 22$), paired-end libraries from liver

512   and gill for 6, 12, and 24 h for niclosamide ($n = 6$), TFM ($n = 6$), and control ($n = 6$) were used,

513   but the TFM:niclosamide mixture only had libraries at 6 and 12 h ($n = 4$), since the combination

514   killed 93% of all lamprey by 12 h of exposure (see Lawrence et al.[19]).

18

515   Each assembly was generated using Trinity (v2.8.5), run with the following command:

516 Trinity --max_memory 250G --seqType fq --samples_file samples.txt --KMER_SIZE 25 --

517 SS_lib_type RF --CPU 24 --bflyCalculateCPU --output full_trinity_assembly &>log_RF_2.txt

518 &[49]. Following Trinity, the completeness of the assemblies was assessed using BUSCO (v4.1.4)

519 to test for the presence of Benchmarking Universal Single Copy Orthologs that are conserved

520 within all eukaryotes using the eukaryota_odb10 database[50]. The bluegill assembly was 98.4%

521 complete with one fragmented and three missing BUSCOs out of the database of 255. The sea

522 lamprey was 96.1% complete with six fragmented and four missing BUSCOs (Table 1). These

523 results are comparable to a recently published *de novo* larval sea lamprey transcriptome

524 assembled from muscle, liver, and brain. The bluegill assembly had 559,416,397 assembled

525 bases in 498,997 transcripts with a median contig length of 413, contig N50 of 2,992, and %GC

526 of 44.38. The sea lamprey assembly had 377,365,173 assembled bases in 474,000 transcripts

527 with a median contig length of 358, contig N50 of 1,648, and %GC of 53.10.

528 *De novo* transcriptome annotation

529   Trinity assemblies were processed with the Trinotate (v3.2) annotation protocol

530 (http://trinotate.github.io), which incorporates many bioinformatics tools, namely TransDecoder

531 (v5.5.0), BLASTP (v 2.10.0), BLASTX (v2.10.0), HMMER (packaged in Trinotate v3.2),

532 TMHMM (v2.0c), SignalP (v4.1), and RNAmmer (packaged in Trinotate v3.2) to identify open

533 reading frames (ORFs) within the assembly and test these ORFs against databases of known

534 proteins, mRNA transcripts, transmembrane helices, signal peptides, and ribosomal RNA

535 sequences in order to provide candidate annotations and gene ontology (GO) terms for

536 transcripts within the assembly[49,51–55]. Predicted annotations were used in downstream

537 differential expression analyses.

538 Super transcriptome assembly

539 Although *de novo* assemblies are often essential when reference genomes do not exist,

540 carrying out differential expression analyses on raw assemblies is complicated by the fact that

541 multiple transcripts are often generated for each gene. To combat this issue, we used the Corset-

542 Lace pipeline (https://github.com/Oshlack/Lace) to generate SuperTranscripts from our *de novo*

543 assemblies. First, we mapped reads back to the Trinity assembly using Bowtie2, then carried out

544 hierarchical clustering of transcripts based on shared reads and expression across transcripts,

19

545   finally creating an assembly of SuperTranscripts based on clustering of gene groups[56–58]. For

546   bluegill, Corset-Lace reduced 498,997 transcripts to 109,702 SuperTranscripts, and 474,000

547   transcripts were reduced to 129,219 SuperTranscripts in sea lamprey. These

548   SuperTranscriptomes were run through BUSCO, as above, and bluegill SuperTranscriptome

549   displayed 84.7% completeness, whereas the lamprey reported 91% completeness. Although both

550   of these measures are slight decreases from the original 98.4% and 96.1% for bluegill and

551   lamprey, respectively, the biggest change in the BUSCO report was the number of complete and

552   single-copy BUSCOs compared to the number of complete and duplicated BUSCOs (Table S2).

553   Given that the genes in the BUSCO list represent Benchmarking Universal Single-Copy

554   Orthologs across eukaryotes, the high percentage of "complete duplicated BUSCOs (D)", 58.8%

555   and 60.4% in bluegill and lamprey, respectively, indicated that our *de novo* Trinity assemblies

556   contained numerous transcripts for the same gene, which would be problematic when assessing

557   differential expression. Our SuperTranscriptomes decreased the overall transcript number in our

558   assemblies by ~3.4–4.5 times, slightly decreased the overall percentage of complete BUSCOs

559   (the sum of single-copy and duplicated BUSCOs) by ~1.05–1.15 times, but most importantly,

560   decreased the number of duplicated BUSCOs by ~17–21 times. The Corset-Lace pipeline greatly

561   increased the number of complete single-copy BUSCOs (S) for both species. Although the

562   Corset-Lace pipeline resulted in slightly fewer genes overall, we believe that those remaining

563   were better assembled, with far fewer genes present across multiple transcripts in the assembly,

564   which increased our power for differential expression analyses.

565   *Differential expression analyses*

566   Reads from bluegill and lamprey TFM experimental and control conditions were mapped to

567   the corresponding SuperTranscriptome using STAR (v2.6.1a)[59]. The resulting BAM files from

568   STAR were used to generate transcript counts using featureCounts (part of subread v2.0.0)[60] for

569   both exon and transcript-based analyses.

570   Transcriptomic analysis was performed in R studio (v1.3.1093) using the R programming

571   language (v3.5.1)[61]. Principal analyses of differential expression patterns were performed using

572   the package 'edgeR' (v3.24.3)[62,63] using a quasi-likelihood pipeline[64]. Count data were first

573   filtered to remove lowly expressed genes with the "filterByExpr" function which used the

574   experimental design matrix to determine minimum gene count thresholds. Filtered count data

575   were then normalized using the trimmed mean of M-values[65]. At this point, normalized data were

576   visually inspected using principal component analysis (PCA; packages 'FactoMineR'[66] and

577   'factoextra'[67]; see Fig. S1) and a multi-dimensional scaling (MDS) plot (package 'Glimma'[68]).

578   Individual samples were also visually inspected using a mean difference (MD) plot.

579   To determine patterns of differential expression, count data were first modeled against a

580   negative binomial distribution[63]. From this, we obtained dispersal estimates using a Cox-Reid

581   profile adjusted likelihood method[63], which were visually inspected by use of a biological

582   coefficient of variation (BCV) plot. Quasi-likelihood (QL) dispersion estimates were determined

583   and were subject to a QL F-test to determine transcripts that had statistically significant

584   differential gene expression[69]. From the resulting QL F-tests, we pulled out pairwise contrasts of

585   interest which include control versus TFM exposed fish for each timepoint of exposure (i.e., 6,

586   12, 24 h). Statistical significance of pairwise comparison was accepted at $\alpha = 0.05$ where all p-

587   values were adjusted for false discovery rates (FDR) via a Benjamini-Hochberg correction[70].

588   *Enrichment analyses*

589   One of our primary goals was to ascertain specific pathways and processes that were

590   underlying the transcriptome response to TFM exposure. To do so, we conducted analyses in

591   pathway enrichment using the R package 'enrichR' (v2.1)[71], which allows R access to the

592   Enrichr databases (https://maayanlab.cloud/Enrichr/)[72]. This platform compares differentially

593   expressed genes in a dataset and determines a list of GO terms that are enriched under the

594   experimental treatment[72]. Here, we used three main databases for comparison:

595   GO_Biological_Process_2018, GO_Molecular_Function_2018, and

596   GO_Cellular_Component_2018. Before making database comparisons, unannotated

597   superTranscript clusters were removed from the analysis. For annotated transcripts, UniProt gene

598   names were retrieved using the UniProt's Retrieve/ID mapping tool

599   (https://www.uniprot.org/uploadlists/). In each transcripts list enrichR was run and GO term lists

600   were subsequently filtered for those that were statistically significant (adjusted p-value < 0.05)

601   and had four or more genes associated with a particular GO term[73].

602   To identify general patterns and biological features of interest, we simplified the resulting

603   enrichR GO term lists using a web-based platform, REVIGO (http://revigo.irb.hr/)[74], which

604   summarizes GO term patterns by grouping terms based on their similarity and then creates a

605   hierarchical structure of the terms[74]. Our REVIGO analyses were at a similarity level of 0.5 (i.e.,

606   small) and used the adjusted *p*-values generated from enrichR. Treemap files from the REVIGO

607   output were used to generate GO term enrichment plots in ggplot2[75]. To reduce term

608   redundancy, we opted to only use biological processes and molecular functions from the

609   REVIGO analysis. In cases where the biological processes and molecular functions provided

610   functionally similar results, only the biological processes were displayed in the Results section

611   with the molecular functions being available in the Supplementary Materials. Similarly,

612   summary plots that offered no insight on specific processes (e.g., DNA replication, protein

613   kinase activity, GTPase activities, etc.) were placed into Supplementary Materials.

614   *Identification of differentially expressed detoxification* transcripts

615   We next sought to identify specific mechanisms underlying TFM detoxification and

616   understand species-specific differences in detoxification ability (see Lawrence et al.[19] for a

617   review). To achieve these goals, we also incorporated a targeted approach, which required us to

618   identify known and suspected candidates for TFM detoxification, including *ugt*'s, one of the

619   principal enzymes believed to be involved in Phase II TFM detoxification (see Wilkie et al.[3]), as

620   well as other enzymes/proteins involved with Phases I, II, and III of detoxification. In the case of

621   *ugt* transcripts, *ugt*-associated transcripts were first identified in each species' annotated

622   transcriptome. These transcripts were then used to filter each list of species- and tissue-specific

623   differentially expressed transcripts to examine how these enzymes responded to TFM exposure.

624   In the case of the Phase I–III proteins, we opted to filter a broad list of common protein

625   families involved in detoxification from the differentially expressed transcripts lists and then

626   manually check their functional role based on their UniProt IDs. As discussed in the Results, we

627   filtered transcripts in the transcriptome associated with detoxification, xenobiotic removal,

628   and/or organic compound breakdown. In the case of Phase III transcripts, we also included

629   transporters that were involved in organic ion transport and those transporting UDP/glucuronide

630   substrates. Phase I search terms included cytochromes P450 (CYPs), alcohol dehydrogenase

631   (ADH), aldehyde dehydrogenase (ALDH), monoamine oxidases (MAO), and paraoxonases

632   (PON). Phase II protein search terms included sulfotransferases (SULT), glutathione transferases

633   (GST), glycine N-methyltransferase (GLYAT), N-terminal acetyltransferases (NAT), and

634   methyltransferases (MT). Phase III search terms included solute carrier's (SLC) and ATP-

635   binding cassette (ABC) transporters.

636   Code and Data Availability

637   All commands and scripts used for quality control (QC), assembly, annotation, read

638   mapping, transcript counting and RNAseq analyses are available at https://github.com/phil-

639   grayson/transcriptome_lawrence_jeffries/ . The raw sequence reads are available at the National

640   Center for Biotechnology Information Sequence Read Archive (accession number

641   SUB8714632). Gene lists and GO term list can be found in the Supplementary Materials.

650   **Contributions**

651   KMJ, MPW, RGM, JMW, CJG, and MFD were responsible for the experimental design. The

652   labs of KMJ and MPW carried out the exposures and sample collection. Data and sample

653   analyses were conducted by MJL, PG, and JDJ. The first draft of the manuscript was written by

654   MJL and PG, with all authors contributing to the refinement and production of the final

655   manuscript.

656

657

658

659

660

23

References:

1. Lovell, S. J., Stone, S. F. & Fernandez, L. The economic impacts of aquatic invasive species: a review of the literature. *Agricultural and Resource Economics Review* **35**, 195–208 (2006).

2. Gherardi, F. & Angiolini, C. Eradication and control of invasive species. in *Biodiversity Conservation and Habitat Management.* (eds. Gherardi, F., Corti, C. & Gualtieri, M.) vol. 2 271–299 (2007).

3. Wilkie, M. P., Hubert, T. D., Boogaard, M. A. & Birceanu, O. Control of invasive sea lampreys using the piscicides TFM and niclosamide: Toxicology, successes & future prospects. *Aquatic Toxicology* **211**, 235–252 (2019).

4. Mahmood, I., Imadi, S. R., Shazadi, K., Gul, A. & Hakeem, K. R. Effects of pesticides on environment. in *Plant, Soil and Microbes* (eds. Hakeem, K. R., Akhtar, M. S. & Abdullah, S. N. A.) 253–269 (Springer International Publishing, 2016). doi:10.1007/978-3-319-27455-3_13.

5. Bridges, C. M. & Semlitsch, R. D. Variation in pesticide tolerance of tadpoles among and within species of Ranidae and patterns of amphibian decline. *Conservation Biology* **14**, 1490–1499 (2000).

6. Daam, M. A. & Rico, A. Freshwater shrimps as sensitive test species for the risk assessment of pesticides in the tropics. *Environmental Science and Pollution Research* **25**, 13235–13243 (2018).

7. Pickering, Q. H., Henderson, C. & Lemke, A. E. The Toxicity of Organic Phosphorus Insecticides to Different Species of Warmwater Fishes. *Transactions of the American Fisheries Society* **91**, 175–184 (1962).

24

683   8. Nyman, A.-M., Schirmer, K. & Ashauer, R. Importance of toxicokinetics for interspecies

684   variation in sensitivity to chemicals. *Environmental Science & Technology* **48**, 5946–5954

685   (2014).

686   9. Guénard, G., Ohe, P. C. von der, de Zwart, D., Legendre, P. & Lek, S. Using phylogenetic

687   information to predict species tolerances to toxic chemicals. *Ecological Applications* **21**,

688   3178–3190 (2011).

689   10. Reid, N. M. & Whitehead, A. Functional genomics to assess biological responses to marine

690   pollution at physiological and evolutionary timescales: toward a vision of predictive

691   ecotoxicology. *Briefings in Functional Genomics* **15**, 358–364 (2016).

692   11. Eide, M. *et al.* The chemical defensome of five model teleost fish. *Scientific Reports* **11**,

693   10546 (2021).

694   12. Byers, J. E. *et al.* Directing research to reduce the impacts of nonindigenous species.

695   *Conservation Biology* **16**, 630–640 (2002).

696   13. Marsden, J. E. & Siefkes, M. J. Control of invasive sea lamprey in the Great Lakes, Lake

697   Champlain, and Finger Lakes of New York. in *Lampreys: biology, conservation and control*

698   (ed Docker, M.F.) 411–479 (Springer, 2019).

699   14. McDonald, D. G. & Kolar, C. S. Research to guide the use of lampricides for controlling sea

700   lamprey. *Journal of Great Lakes Research* **33**, 20–34 (2007).

701   15. Hepditch, S. L. J., Birceanu, O. & Wilkie, M. P. A toxic unit and additive index approach to

702   understanding the interactions of 2 piscicides, 3-trifluoromethyl-4-nitrophenol and

703   niclosamide, in rainbow trout. *Environmental Toxicology and Chemistry* **40**, 1419–1430

704   (2021).

25

705   16. Birceanu, O., McClelland, G. B., Wang, Y. S., Brown, J. C. L. & Wilkie, M. P. The

706   lampricide 3-trifluoromethyl-4-nitrophenol (TFM) uncouples mitochondrial oxidative

707   phosphorylation in both sea lamprey (*Petromyzon marinus*) and TFM-tolerant rainbow trout

708   (*Oncorhynchus mykiss*). *Comparative Biochemistry and Physiology Part C: Toxicology &*

709   *Pharmacology* **153**, 342–349 (2011).

710   17. Wilkie, M. P., Holmes, J. A. & Youson, J. H. The lampricide 3-trifluoromethyl-4-

711   nitrophenol (TFM) interferes with intermediary metabolism and glucose homeostasis, but not

712   with ion balance, in larval sea lamprey (*Petromyzon marinus*). *Canadian Journal of*

713   *Fisheries and Aquatic Sciences* **64**, 1174–1182 (2007).

714   18. Birceanu, O., McClelland, G. B., Wang, Y. S. & Wilkie, M. P. Failure of ATP supply to

715   match ATP demand: The mechanism of toxicity of the lampricide, 3-trifluoromethyl-4-

716   nitrophenol (TFM), used to control sea lamprey (*Petromyzon marinus*) populations in the

717   Great Lakes. *Aquatic Toxicology* **94**, 265–274 (2009).

718   19. Lawrence, M. J. *et al.* Contrasting physiological responses between invasive sea lamprey and

719   non-target bluegill in response to acute lampricide exposure. *Aquatic Toxicology* **237**,

720   105848 (2021).

721   20. Lech, J. J. & Statham, C. N. Role of glucuronide formation in the selective toxicity of 3-

722   trifluoromethyl-4-nitrophenol (TFM) for the sea lamprey: comparative aspects of TFM

723   uptake and conjugation in sea lamprey and rainbow trout. *Toxicology and Applied*

724   *Pharmacology* **31**, 150–158 (1975).

725   21. Bussy, U. *et al.* Metabolism of a sea lamprey pesticide by fish liver enzymes part A:

726   identification and synthesis of TFM metabolites. *Analytical and Bioanalytical Chemistry*

727   **410**, 1749–1761 (2018).

728   22. Kane, A. S., Kahng, M. W., Reimschuessel, R., Nhamburo, P. T. & Lipsky, M. M. UDP-

729   glucuronyltransferase kinetics for 3-trifluoromethyl-4-nitrophenol (TFM) in fish.

730   *Transactions of the American Fisheries Society* **123**, 217–222 (1994).

731   23. Lantz, S. R. *et al.* Next Generation Lampricides: A three-stage process to develop improved

732   control tools for invasive sea lamprey. *Canadian Journal of Fisheries and Aquatic Sciences*

733   (2021).

734   24. Smart, R.C., & Hodgson, E. *Molecular and Biochemical Toxicology*. (Wiley, 2018).

735   25. Funke, T., Han, H., Healy-Fried, M. L., Fischer, M. & Schönbrunn, E. Molecular basis for

736   the herbicide resistance of Roundup Ready crops. *PNAS* **103**, 13010–13015 (2006).

737   26. Peterson, M. A., McMaster, S. A., Riechers, D. E., Skelton, J. & Stahlman, P. W. 2,4-D Past,

738   present, and future: a review. *Weed Technology* **30**, 303–345 (2016).

739   27. Costa, L. G. Current issues in organophosphate toxicology. *Clinica Chimica Acta* **366**, 1–13

740   (2006).

741   28. Brant, C. *Great Lakes sea lamprey: the 70 year war on a biological invader*. (University of

742   Michigan Press, 2019).

743   29. Neave, F. B. *et al.* Changes in native lamprey populations in the Great Lakes since the onset

744   of sea lamprey (*Petromyzon marinus*) control. *Journal of Great Lakes Research*

745   S0380133020302422 (2020) doi:10.1016/j.jglr.2020.10.005.

746   30. Lech, J. J. & Costrini, N. V. In vitro and in vivo metabolism of 3-trifluoromethyl-4-

747   nitrophenol (TFM) in rainbow trout. *Comparative and General Pharmacology* **3**, 160–166

748   (1972).

749   31. Glasauer, S. M. K. & Neuhauss, S. C. F. Whole-genome duplication in teleost fishes and its

750   evolutionary consequences. *Molecular Genetics and Genomics* **289**, 1045–1060 (2014).

27

751   32. Smith, J. J. *et al.* The sea lamprey germline genome provides insights into programmed

752   genome rearrangement and vertebrate evolution. *Nature Genetics* **50**, 270–277 (2018).

753   33. Meech, R. *et al.* The UDP-glycosyltransferase (UGT) superfamily: new members, new

754   functions, and novel paradigms. *Physiological Reviews* **99**, 1153–1222 (2019).

755   34. Bussy, U. *et al.* Metabolism of a sea lamprey pesticide by fish liver enzymes part B: method

756   development and application in quantification of TFM metabolites formed in vivo.

757   *Analytical and Bioanalytical Chemistry* **410**, 1763–1774 (2018).

758   35. Huerta, B. *et al.* Sea lamprey cardiac mitochondrial bioenergetics after exposure to TFM and

759   its metabolites. *Aquatic Toxicology* **219**, 105380 (2020).

760   36. Lam, S. H. *et al.* Molecular insights into 4-nitrophenol-induced hepatotoxicity in zebrafish:

761   Transcriptomic, histological and targeted gene expression analyses. *Biochimica et*

762   *Biophysica Acta (BBA) - General Subjects* **1830**, 4778–4789 (2013).

763   37. Nickel, A., Kohlhaas, M. & Maack, C. Mitochondrial reactive oxygen species production

764   and elimination. *Journal of Molecular and Cellular Cardiology* **73**, 26–33 (2014).

765   38. Jin, Y. *et al.* Antineoplastic mechanisms of niclosamide in acute myelogenous leukemia stem

766   cells: inactivation of the NF-κB pathway and generation of reactive oxygen species. *Cancer*

767   *Research* **70**, 2516–2527 (2010).

768   39. Hamdoun, S., Jung, P. & Efferth, T. Drug repurposing of the anthelmintic niclosamide to

769   treat multidrug-resistant leukemia. *Frontiers in Pharmacology* **8**, (2017).

770   40. Miyoshi, N., Oubrahim, H., Chock, P. B. & Stadtman, E. R. Age-dependent cell death and

771   the role of ATP in hydrogen peroxide-induced apoptosis and necrosis. *Proceedings of the*

772   *National Academy of Sciences* **103**, 1727–1731 (2006).

28

41. Han, Y. H., Kim, S. Z., Kim, S. H. & Park, W. H. 2,4-Dinitrophenol induces apoptosis in As4.1 juxtaglomerular cells through rapid depletion of GSH. *Cell Biology International* **32**, 1536–1545 (2008).

42. Espinosa Ruiz, C. *et al.* Sub-lethal doses of polybrominated diphenyl ethers affect some biomarkers involved in energy balance and cell cycle, via oxidative stress in the marine fish cell line SAF-1. *Aquatic Toxicology* **210**, 1–10 (2019).

43. Morcillo, P., Esteban, M. Á. & Cuesta, A. Heavy metals produce toxicity, oxidative stress and apoptosis in the marine teleost fish SAF-1 cell line. *Chemosphere* **144**, 225–233 (2016).

44. Scott, W. B. & Crossman, E. J. *Freshwater Fishes of Canada*. (Fisheries Research Board of Canada, 1973).

45. Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. (2010).

46. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).

47. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

48. Smith, J. J. *et al.* Sequencing of the sea lamprey (*Petromyzon marinus*) genome provides insights into vertebrate evolution. *Nature Genetics* **45**, 415–421 (2013).

49. Haas, B. J. *et al. De novo transcript sequence recostruction from RNA-Seq: reference generation and analysis with Trinity. Nature Protocols* vol. 8 (2013).

50. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).

29

796   51. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment

797   search tool. *Journal of Molecular Biology* **215**, 403–410 (1990).

798   52. Eddy, S. R. Profile hidden Markov models. *Bioinformatics* **14**, 755–763 (1998).

799   53. Krogh, A., Larsson, B., Von Heijne, G. & Sonnhammer, E. L. L. Predicting transmembrane

800   protein topology with a hidden Markov model: Application to complete genomes. *Journal of*

801   *Molecular Biology* **305**, 567–580 (2001).

802   54. Nielsen, H. Predicting secretory proteins with signaIP. *Methods in Molecular Biology* **1611**,

803   59–73 (2017).

804   55. Lagesen, K. *et al.* RNAmmer: Consistent and rapid annotation of ribosomal RNA genes.

805   *Nucleic Acids Research* **35**, 3100–3108 (2007).

806   56. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods*

807   **9**, 357–359 (2012).

808   57. Davidson, N. M. & Oshlack, A. Corset: Enabling differential gene expression analysis for de

809   novo assembled transcriptomes. *Genome Biology* **15**, 1–14 (2014).

810   58. Davidson, N. M., Hawkins, A. D. K. & Oshlack, A. SuperTranscripts: A data driven

811   reference for analysis and visualisation of transcriptomes. *Genome Biology* **18**, 1–10 (2017).

812   59. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21

813   (2013).

814   60. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: An efficient general purpose program for

815   assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).

816   61. R Core Team (2020). — European Environment Agency. https://www.eea.europa.eu/data-

817   and-maps/indicators/oxygen-consuming-substances-in-rivers/r-development-core-team-

818   2006.

819   62. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for

820   differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140

821   (2010).

822   63. McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor

823   RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research* **40**,

824   4288–4297 (2012).

825   64. Chen, Y., Lun, A. T. L. & Smyth, G. K. From reads to genes to pathways: differential

826   expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood

827   pipeline. *F1000Res* **5**, 1438 (2016).

828   65. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression

829   analysis of RNA-seq data. *Genome Biology* **11**, R25 (2010).

830   66. Lê, S., Josse, J. & Husson, F. FactoMineR : An *R* package for multivariate analysis. *Journal*

831   *of Statistical Software* **25**, (2008).

832   67. Kassambara, A. & Mundt, F. *factoextra: Extract and Visualize the Results of Multivariate*

833   *Data Analyses*. (2020).

834   68. Su, S. *et al.* Glimma: interactive graphics for gene expression analysis. *Bioinformatics* **33**,

835   2050–2052 (2017).

836   69. Lun, A. T. L., Chen, Y. & Smyth, G. K. It's DE-licious: A recipe for differential expression

837   analyses of RNA-seq experiments using quasi-likelihood methods in edgeR. in *Statistical*

838   *Genomics* (eds. Mathé, E. & Davis, S.) vol. 1418 391–416 (Springer New York, 2016).

839   70. Kvam, V. M., Liu, P. & Si, Y. A comparison of statistical methods for detecting

840   differentially expressed genes from RNA-seq data. *American Journal of Botany* **99**, 248–256

841   (2012).

842   71. Jawaid, W. *enrichR: Provides an R Interface to 'Enrichr'*. (2020).

843   72. Chen, E. Y. *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment

844   analysis tool. *BMC Bioinformatics* **14**, 128 (2013).

845   73. Jeffries, K. M. *et al.* Divergent transcriptomic signatures in response to salinity exposure in

846   two populations of an estuarine fish. *Evolutionary Application* **12**, 1212–1226 (2019).

847   74. Supek, F., Bošnjak, M., Škunca, N. & Šmuc, T. REVIGO summarizes and visualizes long

848   lists of gene ontology terms. *PLOS ONE* **6**, e21800 (2011).

849   75. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer International

850   Publishing : Imprint: Springer, 2016). doi:10.1007/978-3-319-24277-4.

851

852

Table 1: Differentially expressed transcripts associated with phase I detoxification for sea lamprey larvae (*Petromyzon marinus*) and bluegill (*Lepomis macrochirus*) following 6, 12, and 24 h TFM (3-trifluoromethyl-4'-nitrophenol) exposure.

Species	Tissue	Exposure Duration (h)	Cluster ID	UniProt Gene Name	FDR	Fold Change
Sea Lamprey						
	Gill	6	Cluster_41283.36015	*cyp1a1*	0.041	5.06
	Gill	6	Cluster_41283.41047	*cyp1a1*	0.049	1.65
	Gill	24	Cluster_41283.1293	*cyp1b1*	0.002	2.57
	Gill	24	Cluster_41283.2130	*cyp1a5*	<0.001	6.95
	Gill	24	Cluster_41283.54788	*cyp3a21*	0.003	1.46
	Gill	24	Cluster_41283.54890	*pon1*	0.033	-1.19
	Gill	24	Cluster_41283.8821	*cyp2j2*	<0.001	-2.75
	Gill	24	Cluster_41283.60437	*cyp2d3*	0.001	-2.01
Bluegill	Gill	6	Cluster_20758.1	*cyp1b1*	<0.001	11.57
	Gill	6	Cluster_49386.0	*cyp1b1*	<0.001	16.23
	Gill	6	Cluster_51425.1737	*cyp1a1*	<0.001	108.31
	Gill	12	Cluster_20758.1	*cyp1b1*	<0.001	12.17
	Gill	12	Cluster_49386.0	*cyp1b1*	<0.001	18.85
	Gill	12	Cluster_51425.1737	*cyp1a1*	<0.001	117.16
	Gill	24	Cluster_11231.2	*mao*	0.011	1.347
	Gill	24	Cluster_20758.1	*cyp1b1*	<0.001	18.37
	Gill	24	Cluster_49386.0	*cyp1b1*	<0.001	40.30
	Gill	24	Cluster_51425.1737	*cyp1a1*	<0.001	186.76
	Liver	6	Cluster_51425.1737	*cyp1a1*	<0.001	6.42
	Liver	12	Cluster_49386.0	*cyp1b1*	<0.001	12.93
	Liver	12	Cluster_51425.10202	*cyp2c31*	0.012	1.75
	Liver	12	Cluster_51425.1737	*cyp1a1*	<0.001	12.80

Liver	24	Cluster_20758.1	*cyp1b1*	0.044	2.14
Liver	24	Cluster_49386.0	*cyp1b1*	<0.001	12.36
Liver	24	Cluster_51425.10202	*cyp2c31*	0.003	1.78
Liver	24	Cluster_51425.1737	*cyp1a1*	<0.001	11.63
Liver	24	Cluster_51425.22684	*cyp2k1*	<0.001	2.46
Liver	24	Cluster_51425.24180	*cyp2k1*	0.027	2.56
Liver	24	Cluster_51425.30529	*cyp3a40*	<0.001	2.19
Liver	24	Cluster_51425.31208	*cyp2k1*	0.005	1.71
Liver	24	Cluster_51425.36982	*cyp2k1*	<0.001	2.18
Liver	24	Cluster_51425.6113	*cyp2k6*	0.020	1.68
Liver	24	Cluster_51425.29803	*cyp2b2*	0.013	-3.08
Liver	24	Cluster_51425.38557	*cyp2k1*	0.008	-1.66
Liver	24	Cluster_51425.7362	*cyp2k6*	0.010	-1.70
Liver	24	Cluster_51425.38638	*cyp2b4*	0.002	2.20

Transcripts were selected using UniProt descriptions where the enzyme was noted to be involved in xenobiotic and/or drug detoxification processes. FDR represents the false discovery rate of the pairwise contrasts and the fold change is against respective controls for a specific timepoint. A positive and negative fold change represents an upregulation or downregulation of the transcript relative to the control. Abbreviations for gene names are as follows: cytochrome P450 (*cyp*), monoamine oxidase (*mao*), and paraoxonase (*pon*).

Table 2: Identification of all UDP-glucuronosyltransferase (*ugt*) transcripts in the annotated transcriptomes of sea lamprey larvae (*Petromyzon marinus*) and bluegill (*Lepomis macrochirus*).

Species	Cluster ID	UniProt Gene Name	Uniprot Entry ID	Expect Value (E)	Percent ID (%)	Protein Name
Sea lamprey						
	Cluster_41283.10908	*ugt2c1*	P36514	2.93×10^{-24}	62.50	UDP-glucuronosyltransferase 2c1
	Cluster_41283.12396	*ugt2a1*	P0DTE4	1.45×10^{-86}	36.96	UDP-glucuronosyltransferase 2a1
	Cluster_41283.4253	*ugt2a2*	Q6PDD0	2.35×10^{-57}	40.35	UDP-glucuronosyltransferase 2a2
	Cluster_41283.48705	*ugt2a2*	Q6PDD0	1.13×10^{-24}	56.99	UDP-glucuronosyltransferase 2a2
	Cluster_41283.5366	*ugt2b1*	P09875	3.84×10^{-14}	63.27	UDP-glucuronosyltransferase 2b1
	Cluster_41283.69509	*ugt2a2*	Q6PDD0	0.00	51.15	UDP-glucuronosyltransferase 2a2
Bluegill						
	Cluster_21009.0	*ugt3*	Q91280	1.22×10^{-28}	41.61	UDP-glucuronosyltransferase
	Cluster_3692.0	*ugt3a1*	Q3UP75	3.36×10^{-85}	46.47	UDP-glucuronosyltransferase 3a1
	Cluster_51425.15463	*ugt2a1*	P0DTE4	1.56×10^{-68}	36.47	UDP-glucuronosyltransferase 2a1
	Cluster_51425.15468	*ugt2a1*	P0DTE4	1.67×10^{-63}	59.87	UDP-glucuronosyltransferase 2a1
	Cluster_51425.15474	*ugt2a1*	Q80X89	1.26×10^{-110}	39.82	UDP-glucuronosyltransferase 2a1
	Cluster_51425.15477	*ugt2b20*	O77649	3.76×10^{-53}	53.57	UDP-glucuronosyltransferase 2b20
	Cluster_51425.15481	*ugt2a1*	P0DTE4	1.40×10^{-62}	59.63	UDP-glucuronosyltransferase 2a1
	Cluster_51425.17384	*ugt3*	Q91280	0.00	67.26	UDP-glucuronosyltransferase
	Cluster_51425.25248	*ugt2b9*	O02663	2.49×10^{-6}	96.00	UDP-glucuronosyltransferase 2b9
	Cluster_51425.26320	*ugt2a1*	P0DTE4	1.15×10^{-126}	38.21	UDP-glucuronosyltransferase 2a1
	Cluster_51425.32583	*ugt1a5*	Q64638	1.09×10^{-7}	50.00	UDP-glucuronosyltransferase 1a5
	Cluster_51425.33579	*ugt2a1*	P36510	0.00	58.26	UDP-glucuronosyltransferase 2a1

Cluster_51425.33985	*ugt2b9*	O02663	1.35×10^{-14}	39.13	UDP-glucuronosyltransferase 2b9
Cluster_51425.36243	*ugt1a2*	P20720	2.74×10^{-124}	42.06	UDP-glucuronosyltransferase 1-2
Cluster_51425.37011	*ugt2a1*	Q80X89	4.91×10^{-132}	42.56	UDP-glucuronosyltransferase 2a1
Cluster_51425.38640	*ugt2b20*	O77649	1.39×10^{-9}	58.49	UDP-glucuronosyltransferase 2b20
Cluster_51425.5745	*ugt2a1*	P0DTE4	1.09×10^{-56}	64.44	UDP-glucuronosyltransferase 2a1

Table 3: Summary of the differentially expressed UDP-glucuronosyltransferase (UGT) transcripts for bluegill (*Lepomis macrochirus*) following 6, 12, and 24 h TFM (3-trifluoromethyl-4'-nitrophenol) exposure.

Tissue	Exposure Duration (h)	Cluster ID	UniProt Gene Name	FDR	Fold Change
Gill	6	Cluster_51425.17384	*ugt3*	<0.001	5.69
Gill	12	Cluster_51425.17384	*ugt3*	<0.001	11.22
Gill	24	Cluster_51425.17384	*ugt3*	<0.001	15.88
Liver	12	Cluster_51425.17384	*ugt3*	<0.001	13.06
Liver	12	Cluster_51425.33985	*ugt2b9*	0.003	1.79
Liver	12	Cluster_51425.36243	*ugt1a2*	0.004	-1.90
Liver	24	Cluster_21009.0	*ugt3*	0.024	1.56
Liver	24	Cluster_51425.17384	*ugt3*	<0.001	28.14
Liver	24	Cluster_51425.33985	*ugt2b9*	<0.001	2.67
Liver	24	Cluster_51425.15468	*ugt2a1*	<0.001	-7.19
Liver	24	Cluster_51425.15481	*ugt2a1*	0.001	-3.14
Liver	24	Cluster_51425.32583	*ugt1a5*	0.038	-1.53
Liver	24	Cluster_51425.36243	*ugt1a2*	<0.001	-2.55

FDR represents the false discovery rate of the pairwise contrasts and the fold change is against respective controls for a specific timepoint. A positive and negative fold change represents an upregulation or downregulation of the transcript relative to the control.

Table 4: Differentially expressed transcripts associated with phase II detoxification for sea lamprey larvae (*Petromyzon marinus*) and bluegill (*Lepomis macrochirus*) following 6, 12, and 24 h TFM (3-trifluoromethyl-4'-nitrophenol) exposure.

Species	Tissue	Exposure Duration (h)	Cluster ID	UniProt Gene Name	FDR	Fold Change
Sea Lamprey						
	Gill	24	Cluster_41283.24751	*gstt3*	<0.001	2.26
	Gill	24	Cluster_41283.69009	*sult6b1*	<0.001	2.29
	Liver	24	Cluster_41283.18986	*nnmt*	0.026	-1.76
Bluegill						
	Gill	6	Cluster_51425.33237	*sult6b1*	0.006	1.54
	Gill	12	Cluster_3656.0	*nat8*	0.005	2.07
	Gill	12	Cluster_51425.33237	*sult6b1*	<0.001	2.19
	Gill	12	Cluster_52314.5	*tpmt*	0.032	-1.30
	Gill	24	Cluster_51425.10275	*sult1st2*	0.025	1.29
	Gill	24	Cluster_51425.33237	*sult6b1*	<0.001	3.81
	Gill	24	Cluster_36463.1	*nat2*	0.026	-1.57
	Liver	24	Cluster_36463.1	*nat2*	0.022	1.70
	Liver	24	Cluster_3656.0	*nat8*	<0.001	2.34
	Liver	24	Cluster_51425.35293	*gstt3*	0.048	-1.37
	Liver	24	Cluster_51425.6836	*sult1c1*	<0.001	-1.63
	Liver	24	Cluster_52314.5	*tpmt*	<0.001	-1.81
	Liver	24	Cluster_7193.0	*gstt1*	0.019	-1.70

Transcripts were selected using UniProt descriptions where the enzyme was noted to be involved in xenobiotic and/or drug detoxification processes. Transcripts excluded UDP-glucuronosyltransferases (see Table 3). FDR represents the false discovery rate of the pairwise contrasts and the fold change is against respective controls for a specific timepoint. A positive and negative fold change represents an upregulation or downregulation of the transcript relative to the control. Abbreviations for gene names are as follows: Glutathione S-transferase (*gstt*), sulfotransferase (*sult*), N-acetyltransferase (*nat*), nicotinamide N-methyltransferase (*nnmt*), and thiopurine S-methyltransferase (*tpmt*).

Table 5: Differentially expressed transcripts associated with phase III detoxification for sea lamprey larvae (*Petromyzon marinus*) and bluegill (*Lepomis macrochirus*) following 6, 12, and 24 h TFM (3-trifluoromethyl-4'-nitrophenol) exposure.

Species	Tissue	Exposure Duration (h)	Cluster ID	UniProt Gene Name	FDR	Fold Change
Sea Lamprey						
	Gill	24	Cluster_34003.4	*abcc3*	0.008	1.55
	Gill	24	Cluster_41283.10772	*abcc5*	<0.001	1.95
	Gill	24	Cluster_41283.22177	*abcb1*	<0.001	5.56
	Gill	24	Cluster_41283.58441	*abcb11*	<0.001	4.31
	Gill	24	Cluster_41283.61313	*abcb1*	<0.001	4.95
	Gill	24	Cluster_41283.62778	*slc7a5*	<0.001	1.73
	Gill	24	Cluster_41283.9558	*abcg2*	<0.001	2.90
	Gill	24	Cluster_41283.1906	*slc35d2*	<0.001	-1.50
	Gill	24	Cluster_41283.24514	*slc22a15b*	0.014	-1.38
	Liver	24	Cluster_41283.10772	*abcc5*	0.017	2.34
	Liver	24	Cluster_41283.69161	*slc35a3*	0.016	-2.10
Bluegill						
	Gill	6	Cluster_49443.0	*slco3a1*	0.012	-1.30
	Gill	12	Cluster_50801.0	*slc29a3*	0.011	1.23
	Gill	12	Cluster_51425.13125	*slco2b1*	0.033	1.31
	Gill	12	Cluster_51425.24527	*slc35d1*	0.046	1.19
	Gill	12	Cluster_51425.35543	*slc7a8*	0.046	1.32
	Gill	12	Cluster_51425.4852	*slc35a3*	0.002	1.32
	Gill	12	Cluster_47743.0	*slc35d2*	0.034	-1.25
	Gill	12	Cluster_49443.0	*slco3a1*	0.004	-1.24
	Gill	24	Cluster_1028.4	*abca3*	<0.001	1.65
	Gill	24	Cluster_10382.9	*abcc4*	0.039	1.23
	Gill	24	Cluster_38275.4	*slc22a15b*	0.032	1.16

Gill	24	Cluster_50801.0	*slc29a3*	<0.001	1.49
Gill	24	Cluster_51425.13125	*slco2b1*	0.007	1.46
Gill	24	Cluster_51425.3133	*slc17a5*	0.006	1.21
Gill	24	Cluster_51425.4852	*slc35a3*	<0.001	1.68
Gill	24	Cluster_52136.2	*abcb6*	<0.001	1.44
Gill	24	Cluster_47743.0	*slc35d2*	0.001	-1.47
Gill	24	Cluster_49443.0	*slco3a1*	0.026	-1.22
Gill	24	Cluster_51425.1412	*abcb1*	0.022	-1.70
Liver	12	Cluster_43105.2	*slc17a5*	0.005	1.44
Liver	12	Cluster_51425.33972	*slc7a5*	0.027	1.58
Liver	12	Cluster_51425.35543	*slc7a8*	0.014	1.83
Liver	12	Cluster_52136.2	*abcb6*	0.001	2.66
Liver	24	Cluster_38275.4	*slc22a15b*	<0.001	1.56
Liver	24	Cluster_43105.2	*slc17a5*	<0.001	1.98
Liver	24	Cluster_51425.3133	*slc17a5*	0.022	1.21
Liver	24	Cluster_51425.33972	*slc7a5*	<0.001	1.88
Liver	24	Cluster_51425.4852	*slc35a3*	0.037	1.52
Liver	24	Cluster_52136.2	*abcb6*	<0.001	2.70
Liver	24	Cluster_17571.3	*slc47a1*	0.002	-1.70
Liver	24	Cluster_51425.13125	*slco2b1*	0.002	-1.87

Transcripts were selected using UniProt descriptions where the enzyme was noted to be involved in xenobiotic and/or drug detoxification processes as well as those capable of transporting conjugate molecules for glucuronidation. FDR represents the false discovery rate of the pairwise contrasts and the fold change is against respective controls for a specific timepoint. A positive and negative fold change represents an upregulation or downregulation of the transcript relative to the control. Abbreviations for gene names are as follows: ATP-binding cassette transporter (*abc*) and solute carrier (*slc*).

24
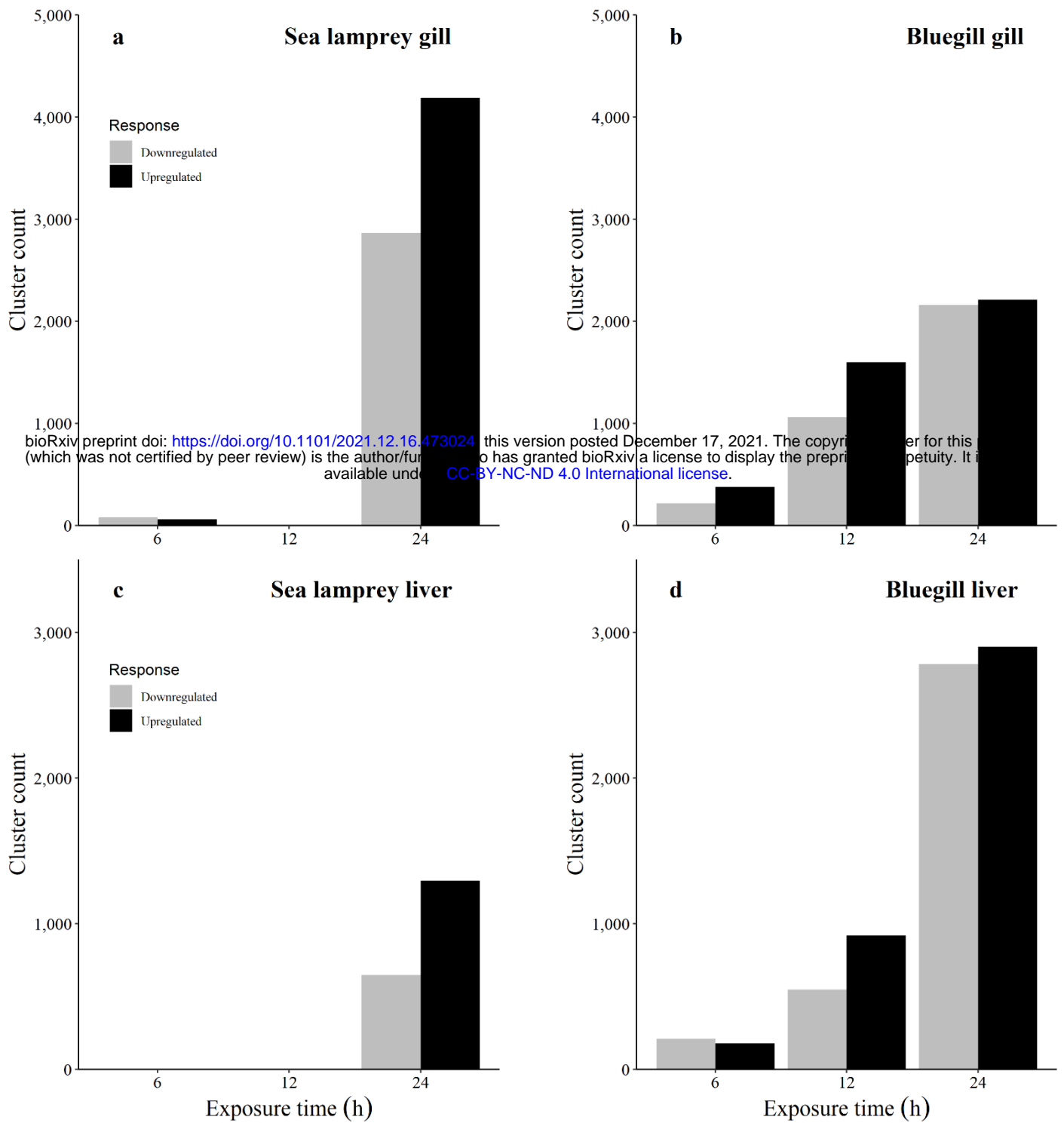
Figure 1: Total counts of differentially expressed superTranscriptome clusters in the gills (a,b) and liver (c,d) of sea lamprey larvae and bluegill following a 24 h TFM exposure. Grey bars represent differentially expressed clusters that were downregulated whereas black bars represent upregulated clusters. All clusters were deemed to be significant at a false discovery rate (FDR) of $\alpha = 0.05$.

30

31

32

33

34

35

36

37

38

39

40

41
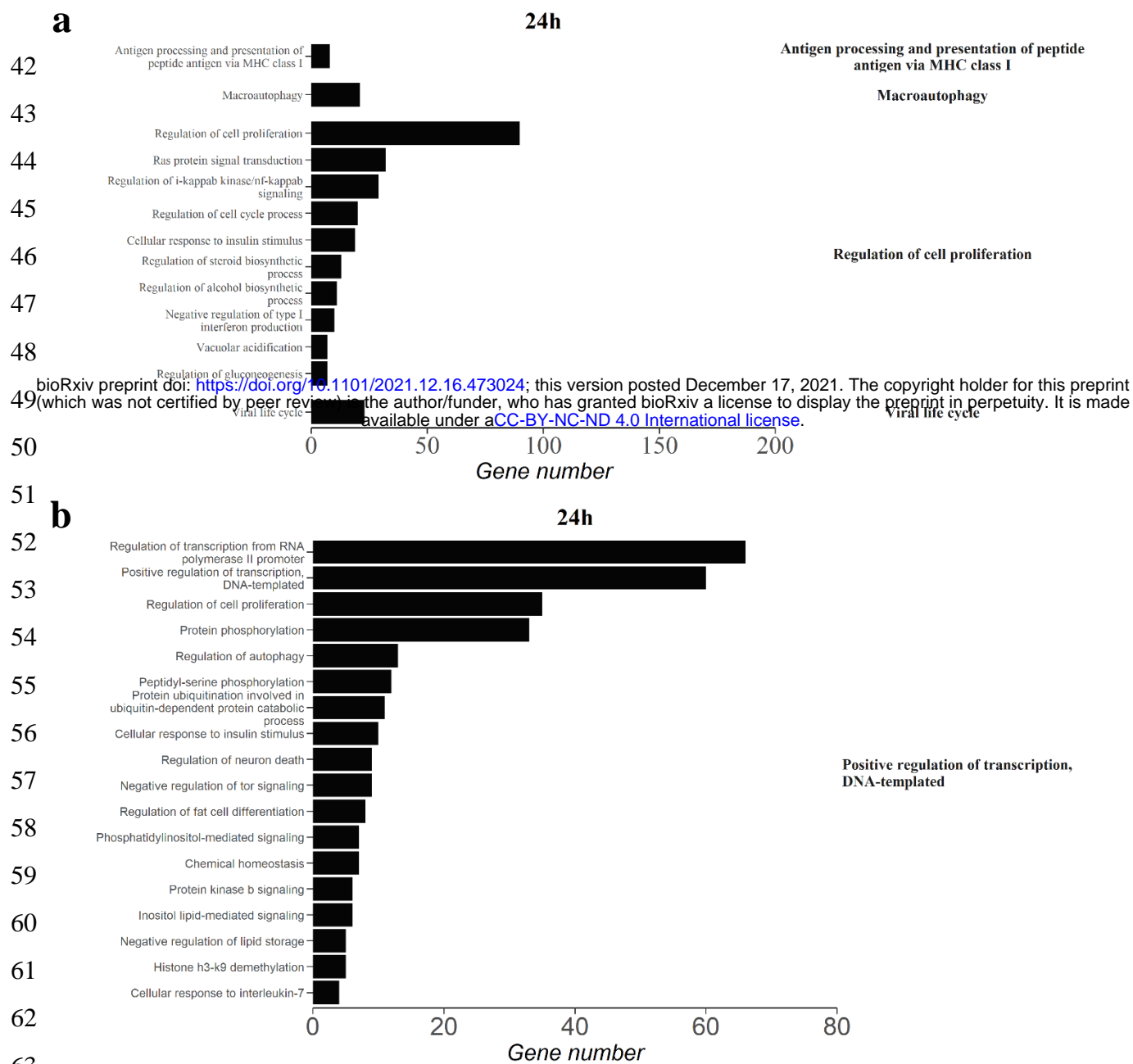
Figure 2: Summary of enriched gene ontology (GO) terms with transcripts that were upregulated following 24 h of TFM (3-trifluoromethyl-4'-nitrophenol; 2.21 mg L^{-1} nominal) exposure in the gills (**a**) and liver (**b**) of sea lamprey (*Petromyzon marinus*) larvae. Transcripts were considered differentially regulated at a false discovery rate < 0.05. Only GO terms from the functional analysis with an adjusted p < 0.05 and at least four transcripts were considered significantly enriched. REVIGO was used to summarize GO terms to reduce redundancy and group according to similarity (right labels).

a

b

6h

Positive regulation of cell proliferation

Regulation of protein phosphorylation

Negative regulation of multicellular organismal process

Wnt signaling pathway

Regulation of phosphatidylinositol 3-kinase signaling

0   5   10   15   20

Negative regulation of multicellular organismal process

12h

Regulation of cell proliferation

Negative regulation of transcription

Cytokine-mediated signaling pathway

Positive regulation of intracellular signal transduction

Regulation of protein phosphorylation

Fat cell differentiation

Regulation of T cell migration

Regulation of cell junction assembly

Embryonic digestive tract development

Regulation of morphogenesis of an epithelium

0   10   20   30   40   50   60

Positive regulation of intracellular signal transduction

24h

Neutrophil mediated immunity
Antigen processing and presentation of exogenous peptide antigen via MHC class 1
Cytokine-mediated signaling pathway
Regulation of cell proliferation
Regulation of apoptotic process
MAPK cascade
Inflammatory response
Transmembrane transport
Regulation of mRNA stability
Regulation of i-kappab kinase/nf-kappab signaling
Regulation of GTPase activity
Nik/nf-kappab signaling
Regulation of primary metabolic process
Regulation of transcription from RNA polymerase II promoter in response to stress
Regulation of stem cell differentiation
Negative regulation of g2/m transition of mitotic cell cycle
Regulation of cellular amino acid metabolic process
Response to endoplasmic reticulum stress
Termination of RNA polymerase II transcription
Second-messenger-mediated signaling
Regulation of calcium ion transport
I-kappab kinase/nf-kappab signaling
Nucleocytoplasmic transport
Cellular response to mechanical stimulus
Anatomical structure homeostasis
Regulation of pri-miRNA transcription from rna polymerase II promoter
Establishment of endothelial barrier
Regulation of protein heterodimerization activity
Pri-miRNA transcription from rna polymerase II promoter

Post-translational protein modification
Scf-dependent proteasomal ubiquitin-dependent protein catabolic process
Protein n-linked glycosylation
Membrane protein proteolysis
Nucleotide-sugar biosynthetic process

Viral life cycle

0   50   100   150

Gene number

Antigen processing and presentation of exogenous peptide antigen via MHC class 1

Cytokine-mediated signaling pathway

Scf-dependent proteasomal ubiquitin-dependent protein catabolism

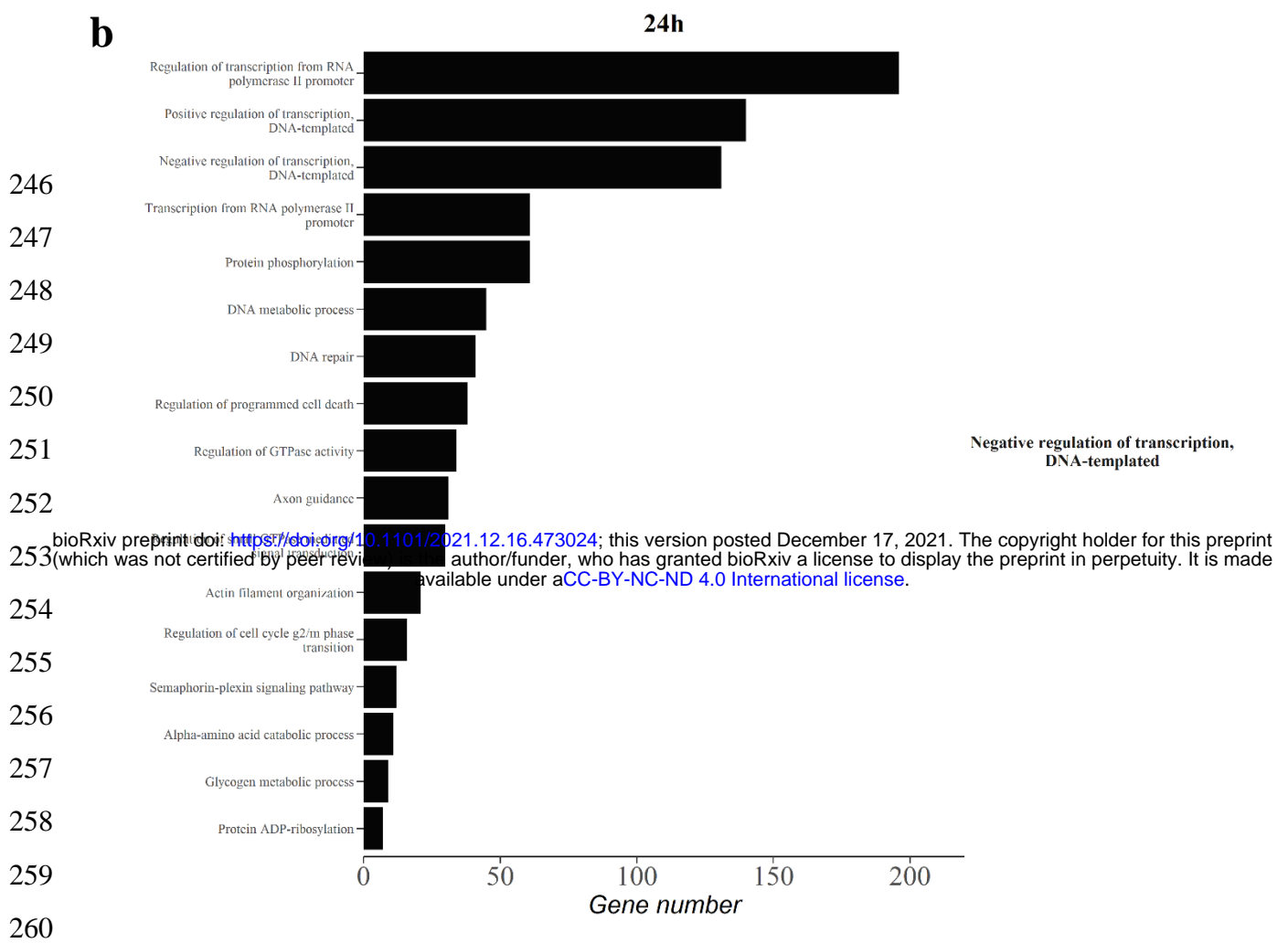**Viral life cycle**

4

164   **Figure 3:** Summary of enriched gene ontology (GO) terms with transcripts that were upregulated
165   (**a**) and downregulated (**b**) following 6 h (light grey), 12 h (dark grey), and 24 h (black) of TFM
166   (3-trifluoromethyl-4'-nitrophenol; 22.06 mg L^{-1} nominal) exposure in the gills of bluegill
167   (*Lepomis macrochirus*). Transcripts were considered differentially regulated at a false discovery
168   rate < 0.05. Only GO terms from the functional analysis with an adjusted $p < 0.05$ and at least
169   four transcripts were considered significantly enriched. REVIGO was used to summarize GO
170   terms to reduce redundancy and group according to similarity (right labels).

171

172

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

a

b

24h

Negative regulation of transcription, DNA-templated

Gene number
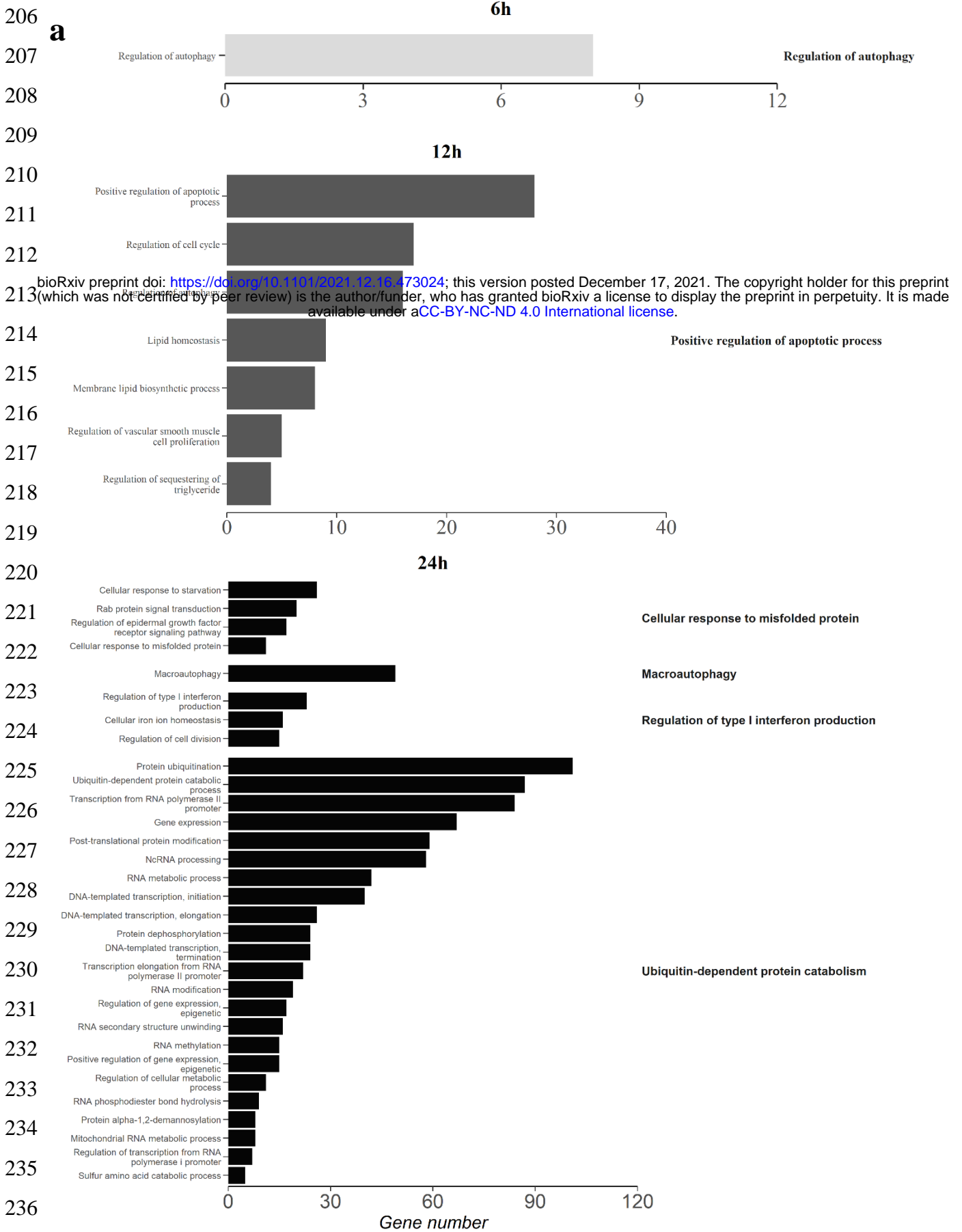
Figure 4: Summary of enriched gene ontology (GO) terms with transcripts that were upregulated (**a**) and downregulated (**b**) following 6 h (light grey), 12 h (dark grey) and 24 h (black) of TFM (3-trifluoromethyl-4'-nitrophenol; 22.06 mg L^{-1} nominal) exposure in the liver of bluegill (*Lepomis macrochirus*). Transcripts were considered differentially regulated at a false discovery rate < 0.05. Only GO terms from the functional analysis with an adjusted p < 0.05 and at least four transcripts were considered significantly enriched. REVIGO was used to summarize GO terms to reduce redundancy and group according to similarity (right labels).

7

288
289
290
291
292
293
294

296
297
298
299
300
301
302
303
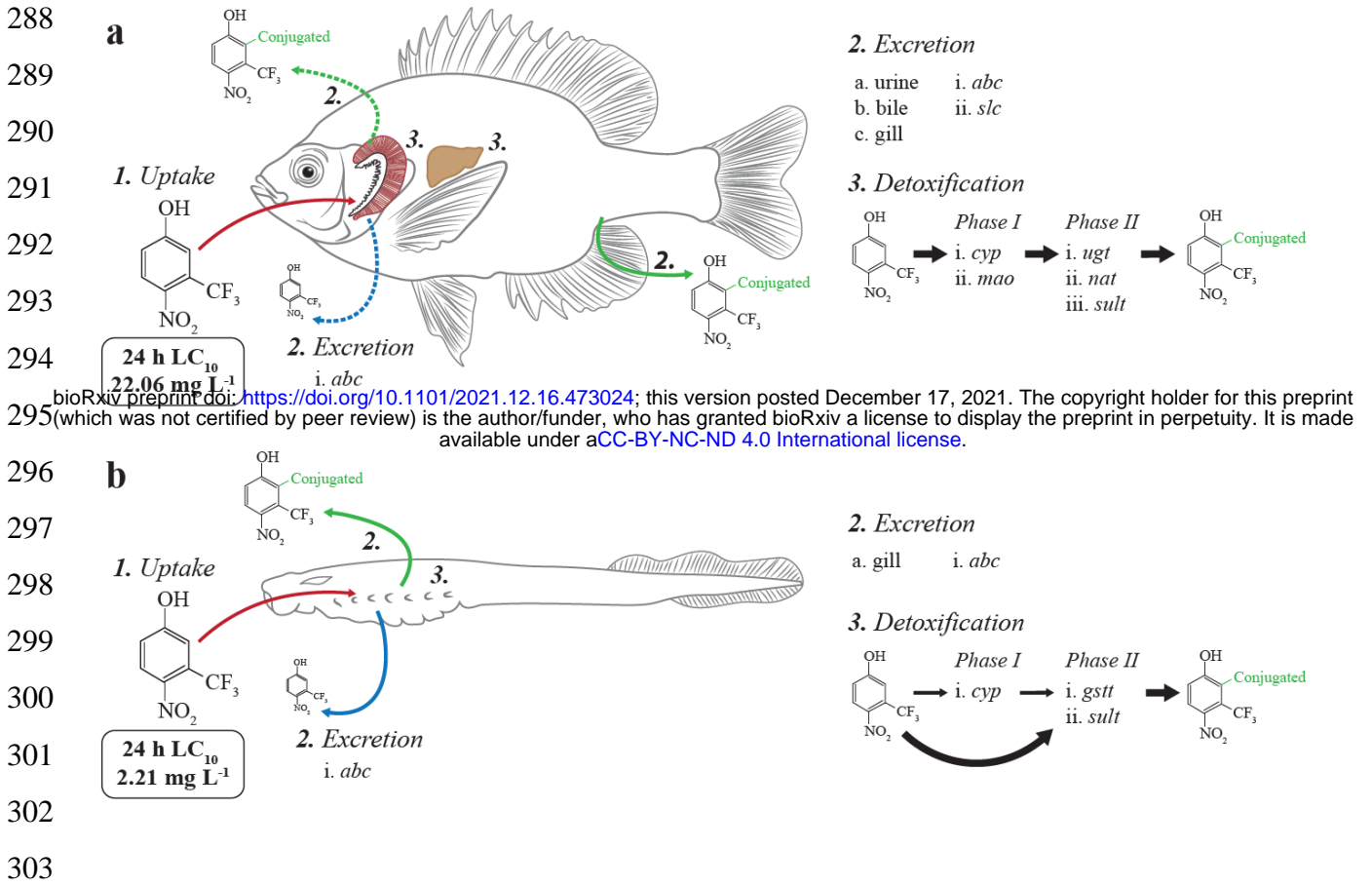
Figure 5: Overview of the pathways of TFM (3-trifluoromethyl-4'-nitrophenol) uptake, excretion, biotransformation, and elimination in bluegill (*Lepomis macrochirus*; **a**) and sea lamprey larvae (*Petromyzon marinus*; **b**). Detoxification transcripts that are presented here were found to have higher differential expression following exposure to TFM (bluegill = 22.06 mg L^{-1}, lamprey = 2.21 mg L^{-1}) over a 24 h period.

8

1

2   **Table S1:** Total sequencing reads counts by species and tissue.

Species	Tissue	Total Reads
Sea lamprey		
	Gill	1.73×10^9
	Liver	1.19×10^9
Bluegill		
	Gill	1.59×10^9
	Liver	1.81×10^9

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25   **Table S2.** BUSCO reports comparing the initial Trinity assembly to the Corset-Lace assembly

Species	Assembly	Complete BUSCOs (C) = S+D	Complete Single-Copy BUSCOs (S)	Complete duplicated BUSCOs (D)	Fragmented BUSCOs (F)	Missing BUSCOs (M)
Bluegill	Trinity	98.4%	39.6%	58.8%	0.4%	1.2%
	Corset/Lace	84.7%	82.0%	2.7%	9.4%	5.9%
Sea lamprey	Trinity	96.1%	35.7%	60.4%	2.4%	1.5%
	Corset/Lace	91%	87.5%	3.5%	6.3%	2.7%

26

27

28

29

30

31

32

33

34

35
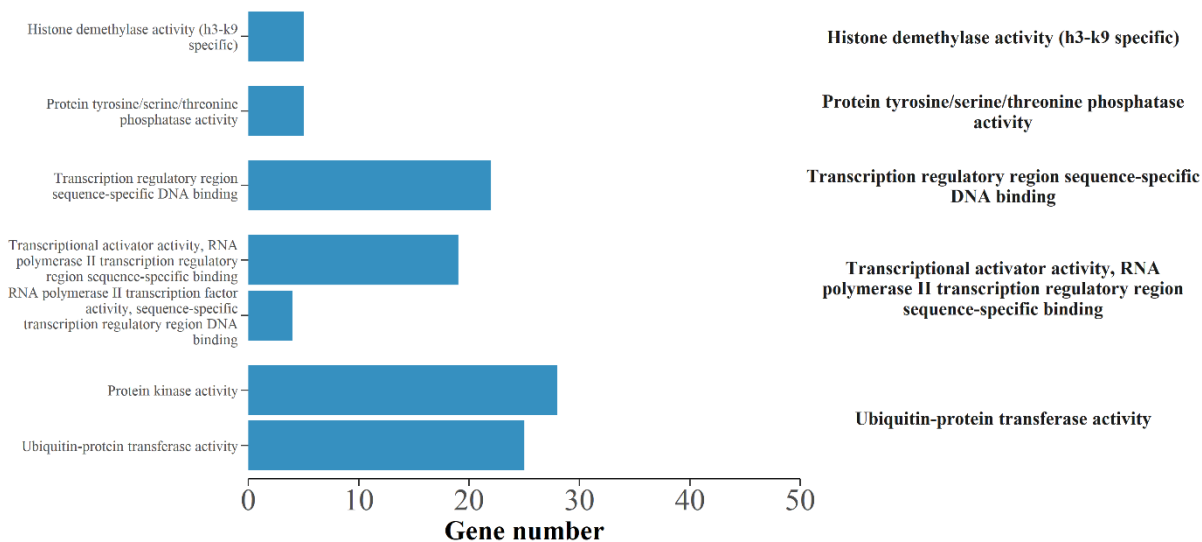
36

37

38

39

40

41

42

43

44

45 **Figure S1:** Summary of enriched gene ontology (GO) terms associated with biological processes
46 in transcripts that were upregulated following 6 h of TFM (3-trifluoromethyl-4'-nitrophenol; 2.21
47 mg L^{-1} nominal) exposure in the gills of sea lamprey (*Petromyzon marinus*) larvae. Transcripts
48 were considered differentially regulated at a false discovery rate < 0.05. Only GO terms from the
49 functional analysis with an adjusted $p < 0.05$ and at least four transcripts were considered
50 significantly enriched. REVIGO was used to summarize GO terms to reduce redundancy and
51 group according to similarity (right labels).

52

53

54

55

56

57

58

59

60

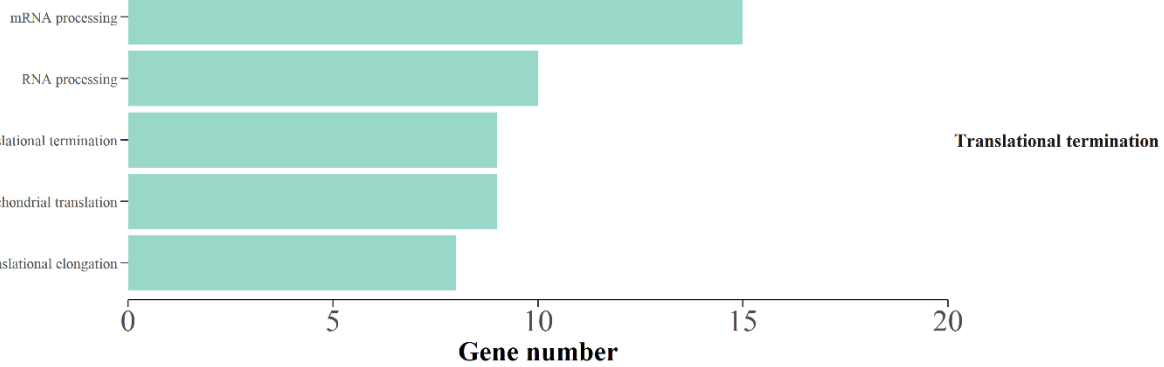61

62

63

64

65

66

67

68

3
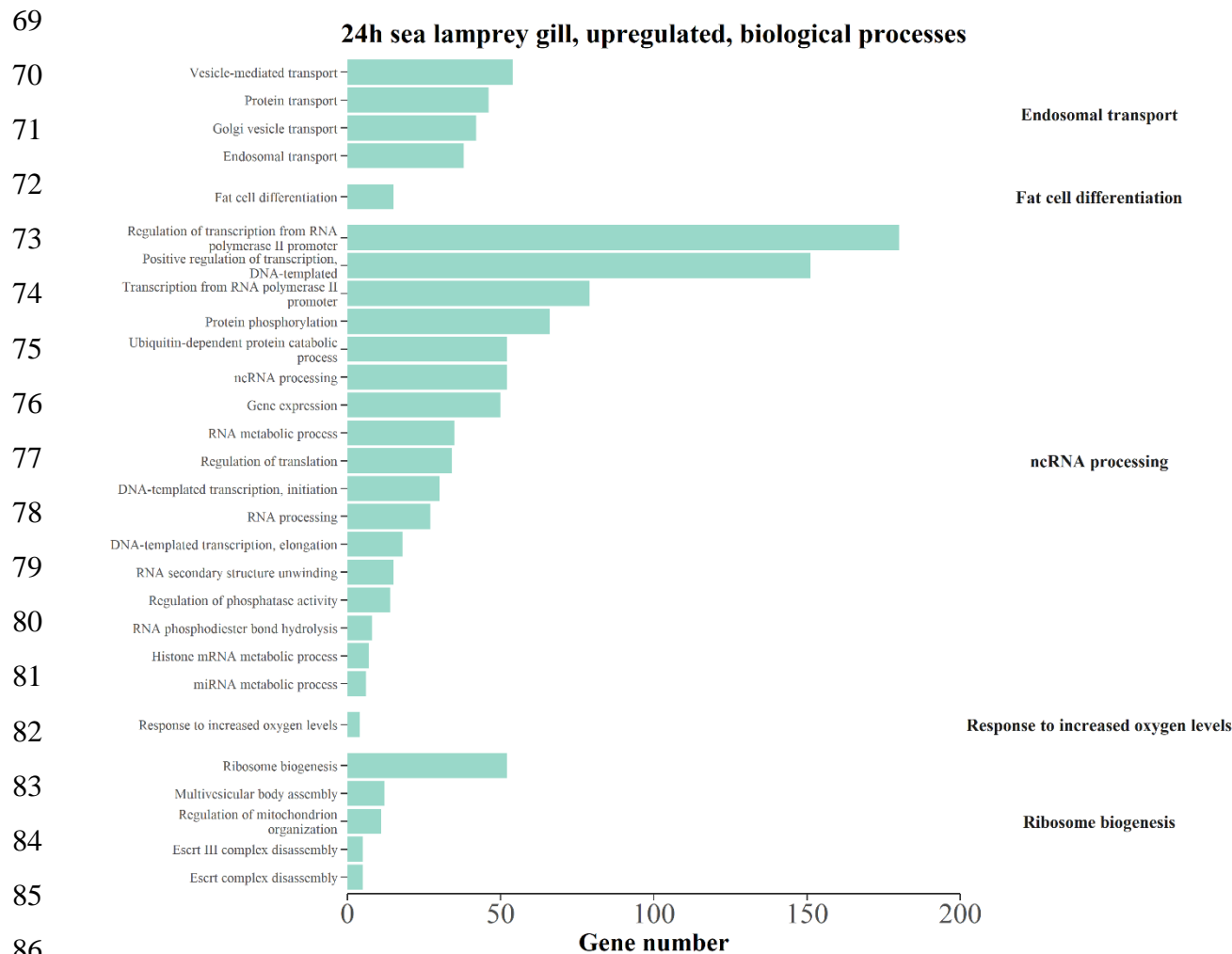
Figure S2: Summary of enriched gene ontology (GO) terms associated with biological processes in transcripts that were upregulated following 24 h of TFM (3-trifluoromethyl-4'-nitrophenol; 2.21 mg L^{-1} nominal) exposure in the gills of sea lamprey (*Petromyzon marinus*) larvae. Transcripts were considered differentially regulated at a false discovery rate < 0.05. Only GO terms from the functional analysis with an adjusted p < 0.05 and at least four transcripts were considered significantly enriched. REVIGO was used to summarize GO terms to reduce redundancy and group according to similarity (right labels).
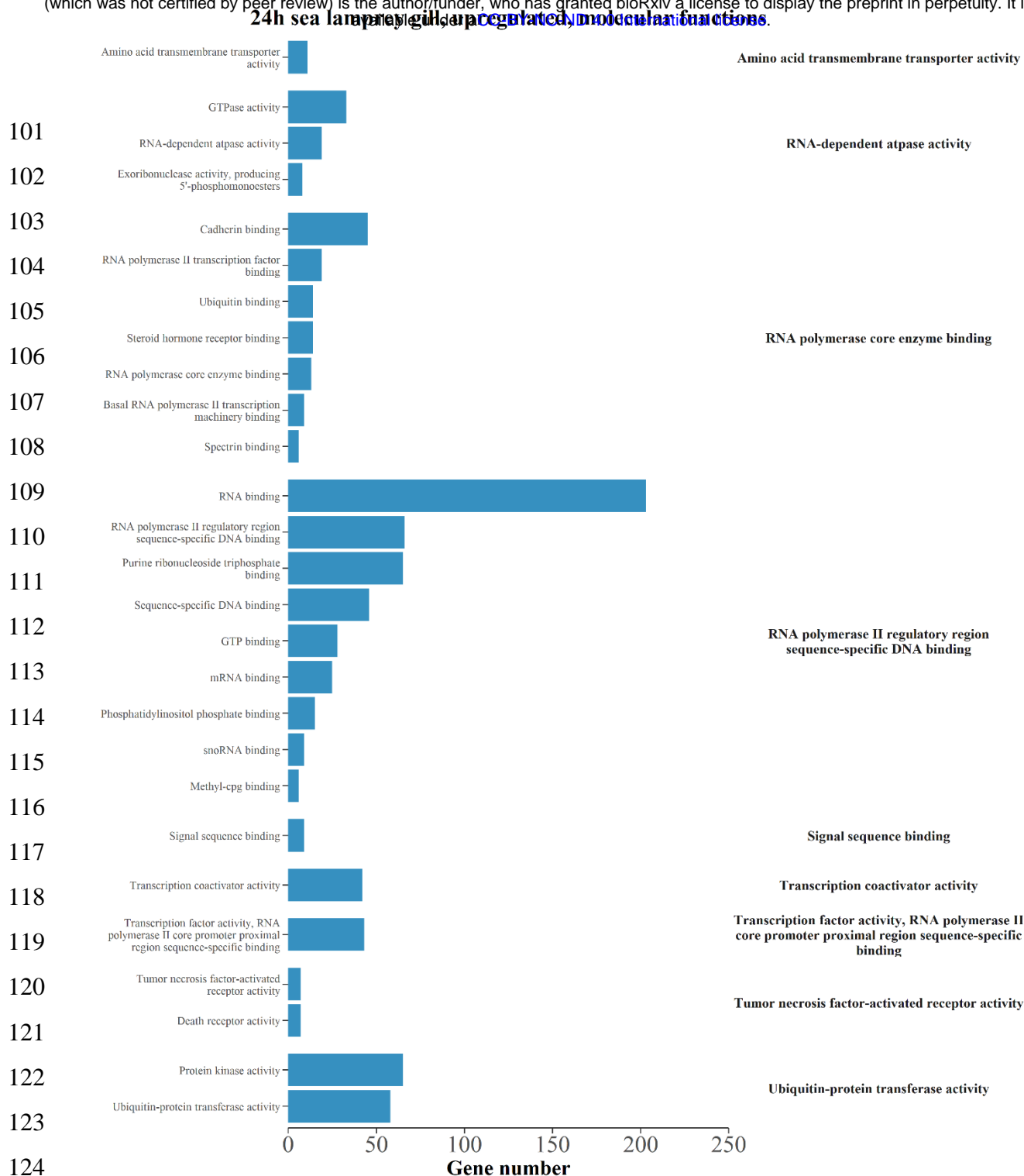
Figure S3: Summary of enriched gene ontology (GO) terms associated with molecular functions in transcripts that were upregulated following 24 h of TFM (3-trifluoromethyl-4'-nitrophenol; 2.21 mg L^{-1} nominal) exposure in the gills of sea lamprey (*Petromyzon marinus*) larvae. Transcripts were considered differentially regulated at a false discovery rate < 0.05. Only GO terms from the functional analysis with an adjusted p < 0.05 and at least four transcripts were considered significantly enriched. REVIGO was used to summarize GO terms to reduce redundancy and group according to similarity (right labels).
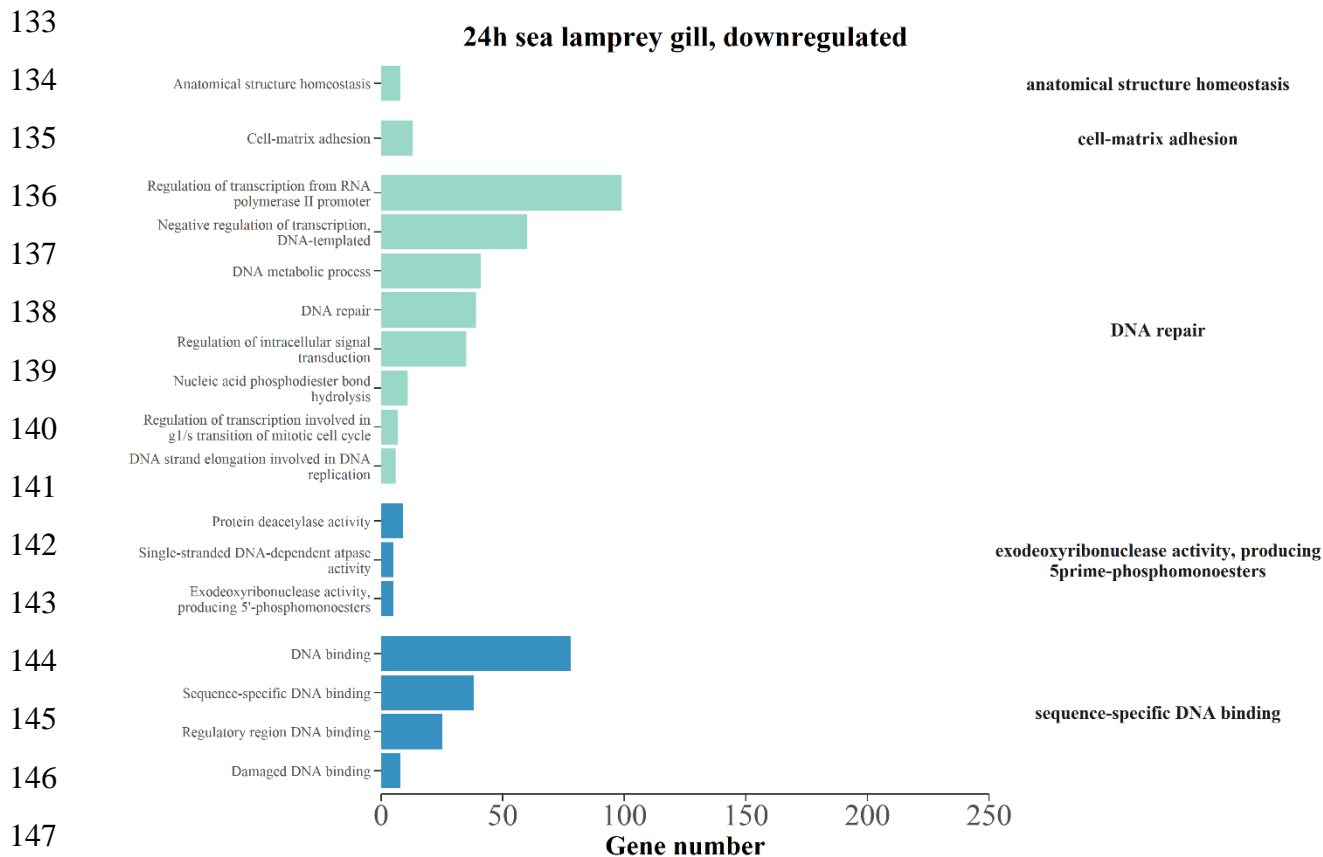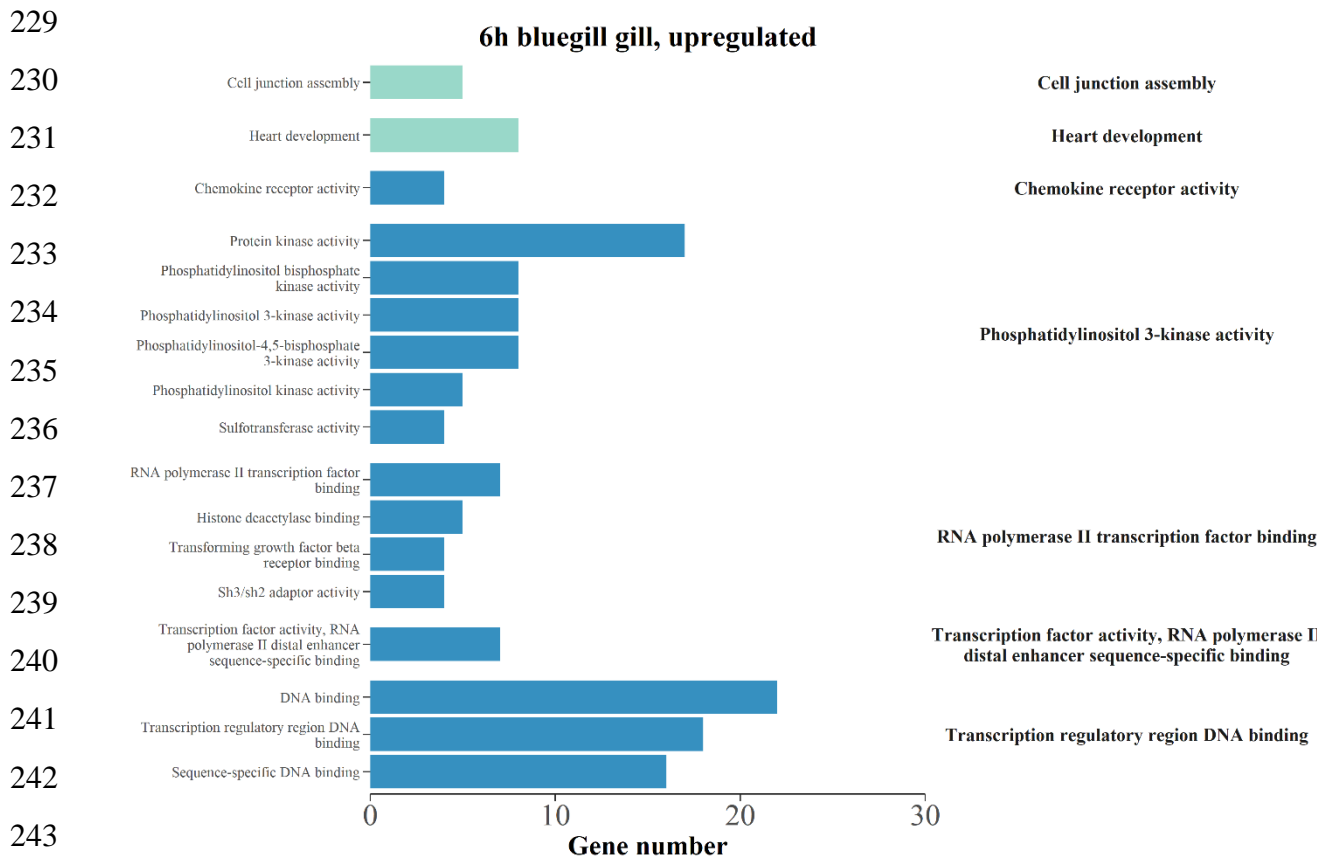
Figure S4: Summary of enriched gene ontology (GO) terms associated with biological processes (green) and molecular functions (blue) in transcripts that were downregulated following 24 h of TFM (3-trifluoromethyl-4'-nitrophenol; 2.21 mg L^{-1} nominal) exposure in the gills of sea lamprey (*Petromyzon marinus*) larvae. Transcripts were considered differentially regulated at a false discovery rate < 0.05. Only GO terms from the functional analysis with an adjusted p < 0.05 and at least four transcripts were considered significantly enriched. REVIGO was used to summarize GO terms to reduce redundancy and group according to similarity (right labels).

24h sea lamprey liver, upregulated, molecular functions

Figure S5: Summary of enriched gene ontology (GO) terms associated with molecular functions in transcripts that were upregulated following 24 h of TFM (3-trifluoromethyl-4'-nitrophenol; 2.21 mg L^{-1} nominal) exposure in the liver of sea lamprey (*Petromyzon marinus*) larvae. Transcripts were considered differentially regulated at a false discovery rate < 0.05. Only GO terms from the functional analysis with an adjusted p < 0.05 and at least four transcripts were considered significantly enriched. REVIGO was used to summarize GO terms to reduce redundancy and group according to similarity (right labels).

197

Figure S6: Summary of enriched gene ontology (GO) terms associated with biological processes in transcripts that were downregulated following 24 h of TFM (3-trifluoromethyl-4'-nitrophenol; 2.21 mg L^{-1} nominal) exposure in the liver of sea lamprey (*Petromyzon marinus*) larvae. Transcripts were considered differentially regulated at a false discovery rate < 0.05. Only GO terms from the functional analysis with an adjusted p < 0.05 and at least four transcripts were considered significantly enriched. REVIGO was used to summarize GO terms to reduce redundancy and group according to similarity (right labels).

Figure S7: Summary of enriched gene ontology (GO) terms associated with biological processes (green) and molecular functions (blue) in transcripts that were upregulated following 6 h of TFM (3-trifluoromethyl-4'-nitrophenol; 22.06 mg L^{-1} nominal) exposure in the gills of bluegill (*Lepomis macrochirus*). Transcripts were considered differentially regulated at a false discovery rate < 0.05. Only GO terms from the functional analysis with an adjusted p < 0.05 and at least four transcripts were considered significantly enriched. REVIGO was used to summarize GO terms to reduce redundancy and group according to similarity (right labels).

9

Figure S8: Summary of enriched gene ontology (GO) terms associated with biological processes (green) and molecular functions (blue) in transcripts that were downregulated following 6 h of TFM (3-trifluoromethyl-4'-nitrophenol; 22.06 mg L^{-1} nominal) exposure in the gills of bluegill (*Lepomis macrochirus*). Transcripts were considered differentially regulated at a false discovery rate < 0.05. Only GO terms from the functional analysis with an adjusted $p < 0.05$ and at least four transcripts were considered significantly enriched. REVIGO was used to summarize GO terms to reduce redundancy and group according to similarity (right labels).
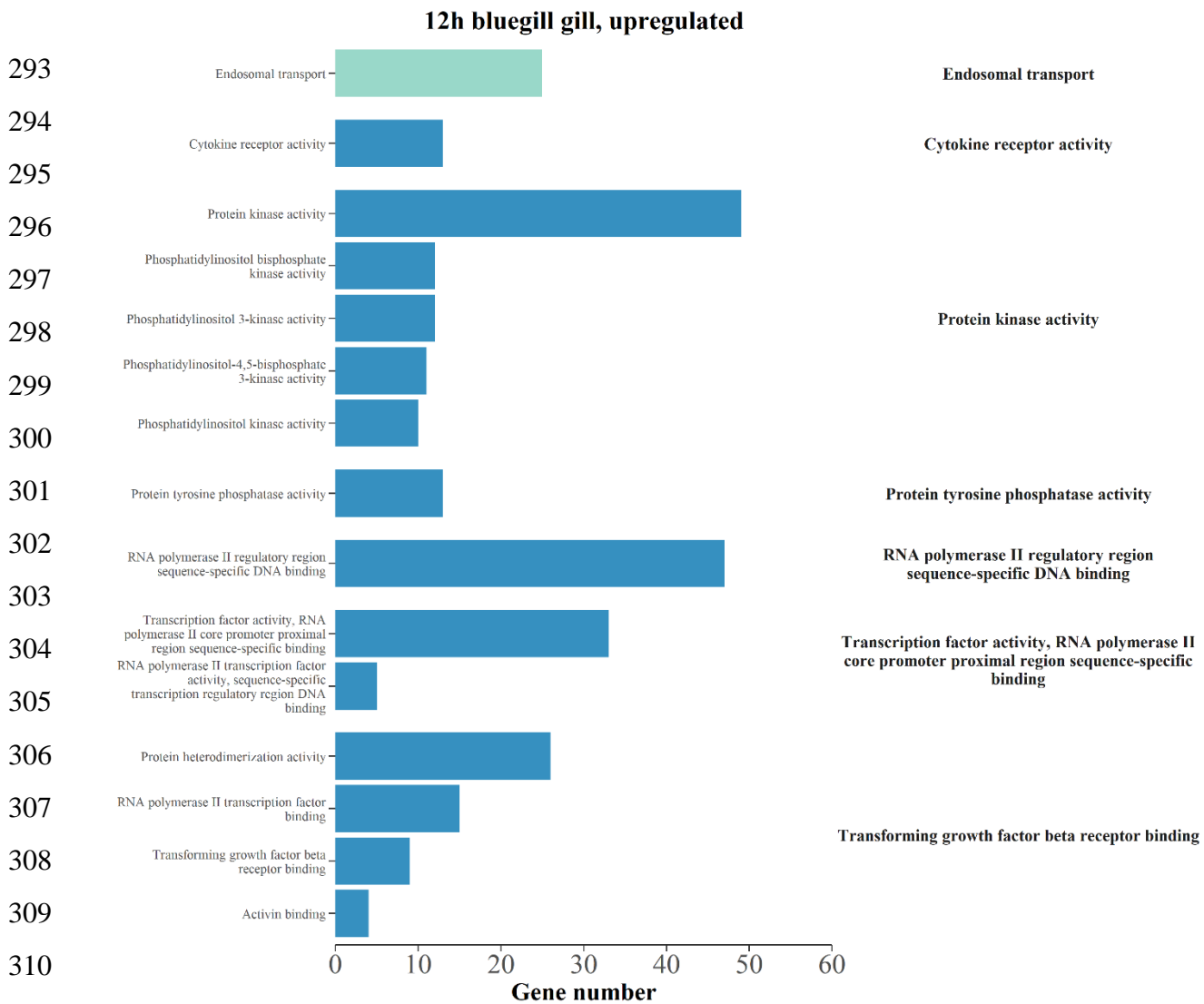
Figure S9: Summary of enriched gene ontology (GO) terms associated with biological processes (green) and molecular functions (blue) in transcripts that were upregulated following 12 h of TFM (3-trifluoromethyl-4'-nitrophenol; 22.06 mg L^{-1} nominal) exposure in the gills of bluegill (*Lepomis macrochirus*). Transcripts were considered differentially regulated at a false discovery rate < 0.05. Only GO terms from the functional analysis with an adjusted p < 0.05 and at least four transcripts were considered significantly enriched. REVIGO was used to summarize GO terms to reduce redundancy and group according to similarity (right labels).

325

326

12h bluegill gill, downregulated, biological processes

327

328

L-serine metabolic process   **L-serine metabolism**

329

| 0 | 1 | 2 | 3 | 4 | 5 |

Gene number

330

331   **Figure S10:** Summary of enriched gene ontology (GO) terms associated with biological
332   processes in transcripts that were downregulated following 12 h of TFM (3-trifluoromethyl-4'-
333   nitrophenol; 22.06 mg L^{-1} nominal) exposure in the gills of bluegill (*Lepomis macrochirus*).
334   Transcripts were considered differentially regulated at a false discovery rate < 0.05. Only GO
335   terms from the functional analysis with an adjusted $p < 0.05$ and at least four transcripts were
336   considered significantly enriched. REVIGO was used to summarize GO terms to reduce
337   redundancy and group according to similarity (right labels).

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

12
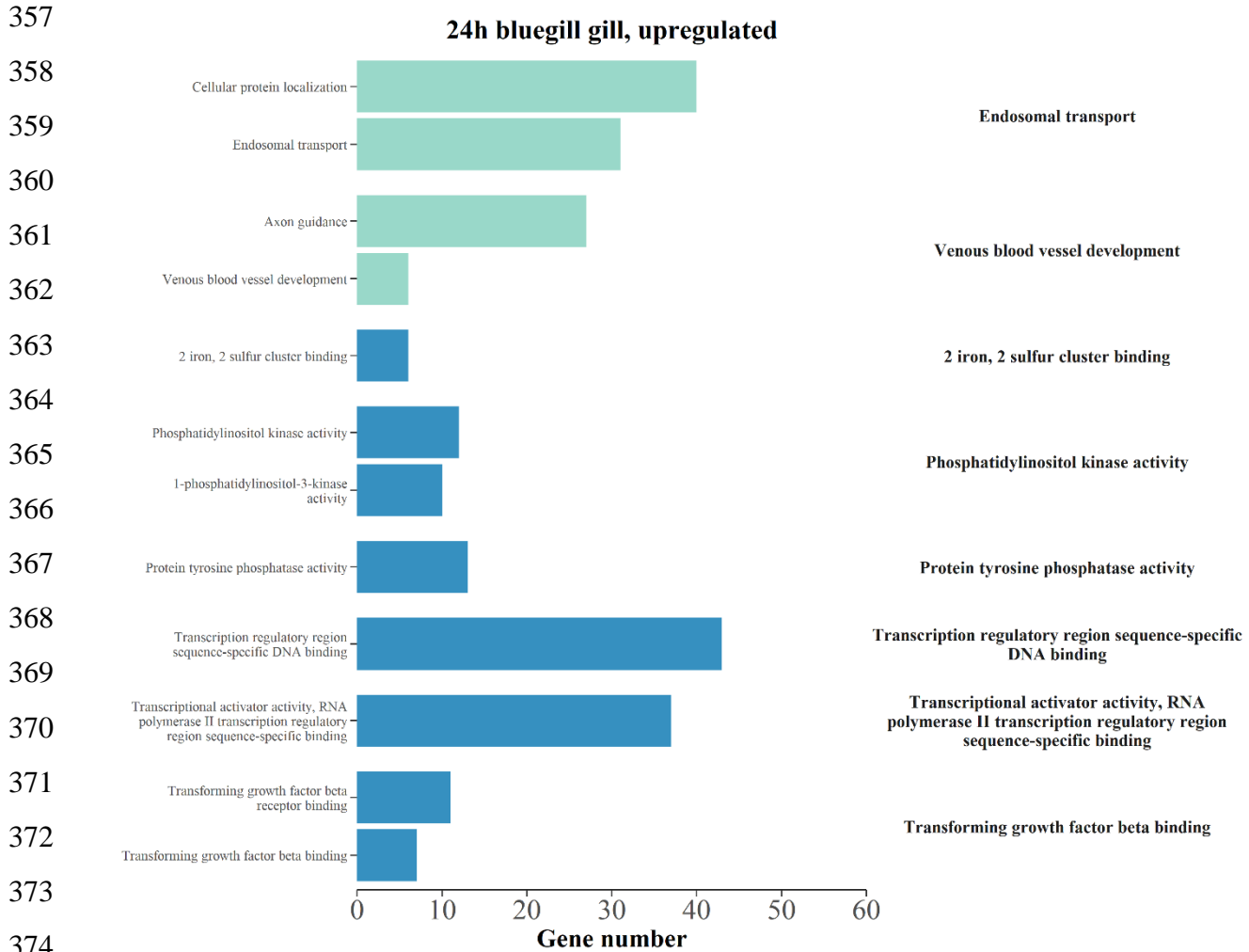
Figure S11: Summary of enriched gene ontology (GO) terms associated with biological processes (green) and molecular functions (blue) in transcripts that were upregulated following 24 h of TFM (3-trifluoromethyl-4'-nitrophenol; 22.06 mg L^{-1} nominal) exposure in the gills of bluegill (*Lepomis macrochirus*). Transcripts were considered differentially regulated at a false discovery rate < 0.05. Only GO terms from the functional analysis with an adjusted p < 0.05 and at least four transcripts were considered significantly enriched. REVIGO was used to summarize GO terms to reduce redundancy and group according to similarity (right labels).

13

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

24h bluegill gill, downregulated

419   **Figure S12:** Summary of enriched gene ontology (GO) terms associated with biological
420   processes (green) and molecular functions (blue) in transcripts that were downregulated
421   following 24 h of TFM (3-trifluoromethyl-4'-nitrophenol; 22.06 mg L^{-1} nominal) exposure in the
422   gills of bluegill (*Lepomis macrochirus*). Transcripts were considered differentially regulated at a
423   false discovery rate < 0.05. Only GO terms from the functional analysis with an adjusted p <
424   0.05 and at least four transcripts were considered significantly enriched. REVIGO was used to
425   summarize GO terms to reduce redundancy and group according to similarity (right labels).

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

6h bluegill liver, downregulated, biological processes
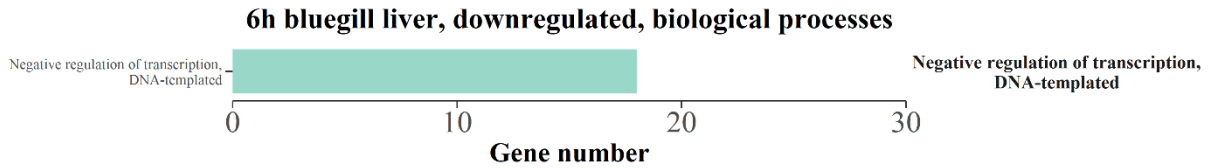
Figure S13: Summary of enriched gene ontology (GO) terms associated with biological processes in transcripts that were downregulated following 6 h of TFM (3-trifluoromethyl-4'-nitrophenol; 22.06 mg L^{-1} nominal) exposure in the liver of bluegill (*Lepomis macrochirus*). Transcripts were considered differentially regulated at a false discovery rate < 0.05. Only GO terms from the functional analysis with an adjusted $p < 0.05$ and at least four transcripts were considered significantly enriched. REVIGO was used to summarize GO terms to reduce redundancy and group according to similarity (right labels).
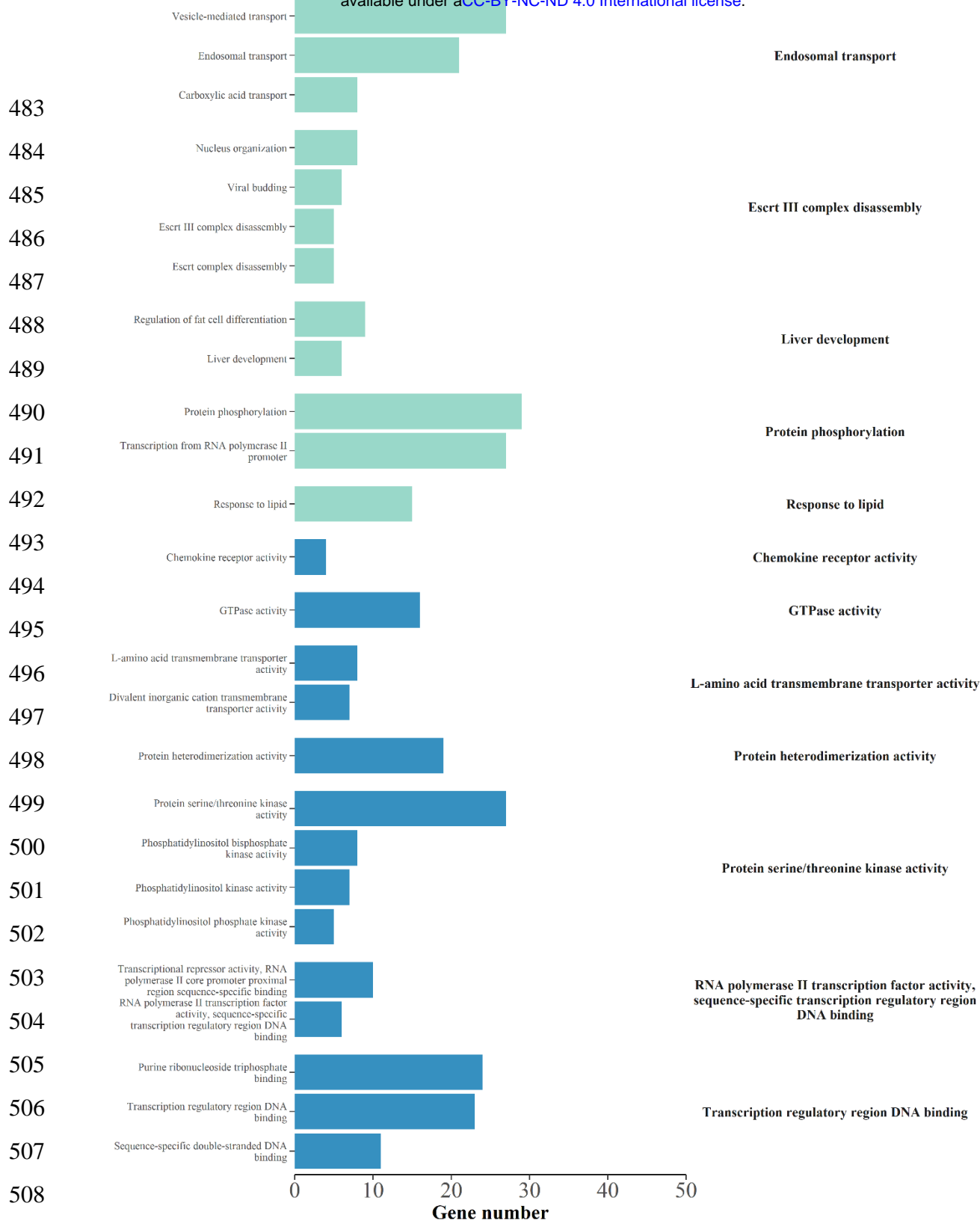
Figure S14: Summary of enriched gene ontology (GO) terms associated with biological processes (green) and molecular functions (blue) in transcripts that were upregulated following 12 h of TFM (3-trifluoromethyl-4'-nitrophenol; 22.06 mg L^{-1} nominal) exposure in the liver of bluegill (*Lepomis macrochirus*). Transcripts were considered differentially regulated at a false

514   discovery rate < 0.05. Only GO terms from the functional analysis with an adjusted p < 0.05 and
515   at least four transcripts were considered significantly enriched. REVIGO was used to summarize
516   GO terms to reduce redundancy and group according to similarity (right labels).

517

518

519

520

521

522

523

524

525

526

527

528

529

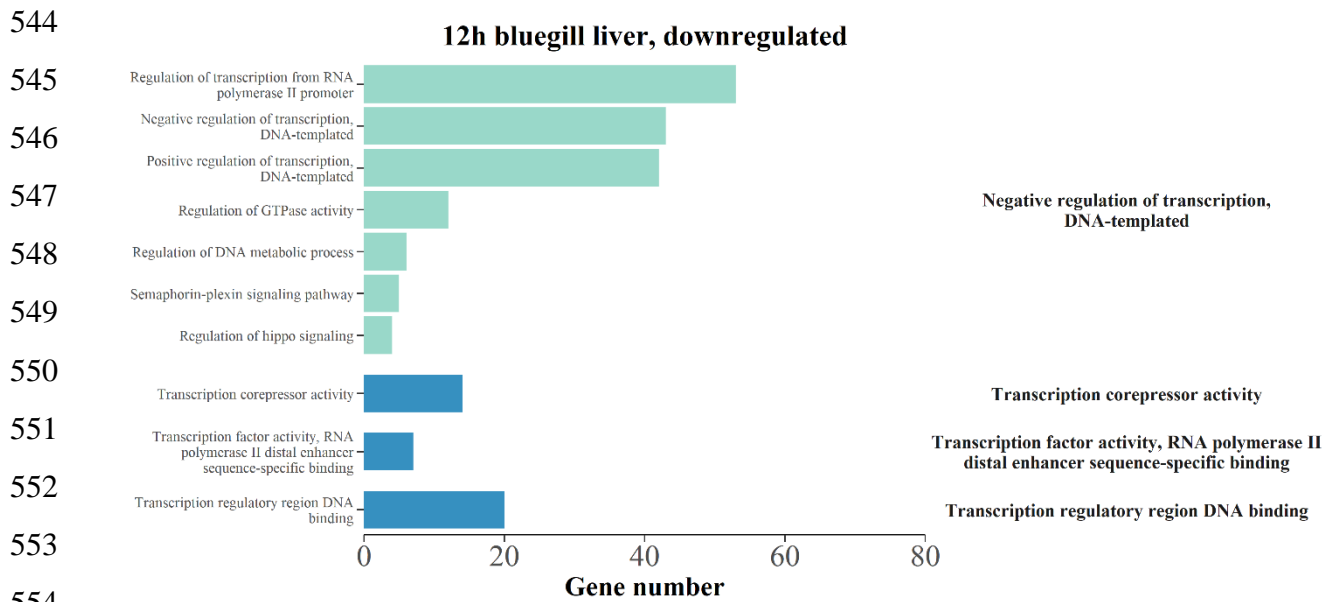530

531

532

533

534

535

536

537

538

539

540
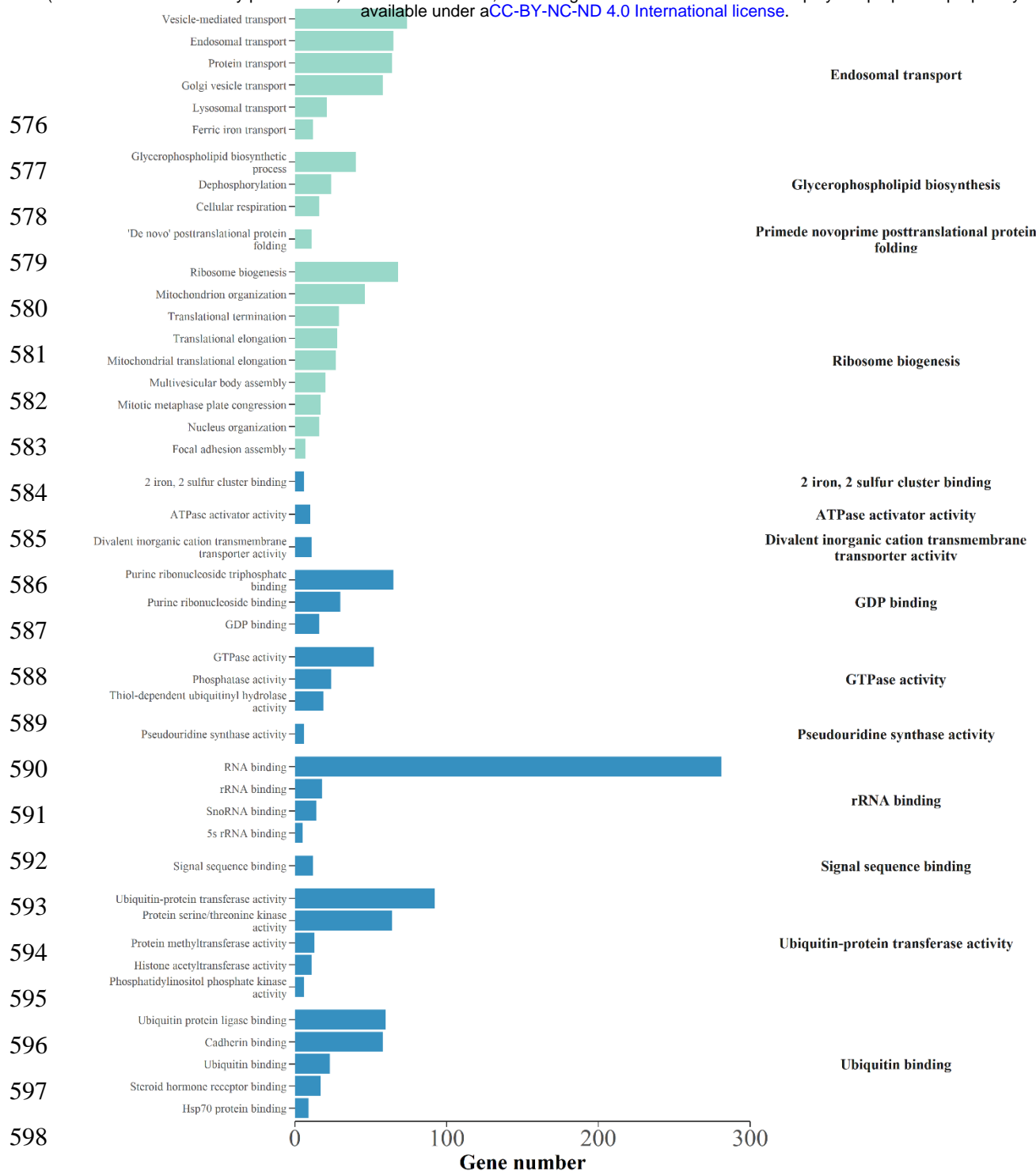
541

542

543

Figure S15: Summary of enriched gene ontology (GO) terms associated with biological processes (green) and molecular functions (blue) in transcripts that were downregulated following 12 h of TFM (3-trifluoromethyl-4'-nitrophenol; 22.06 mg L^{-1} nominal) exposure in the liver of bluegill (*Lepomis macrochirus*). Transcripts were considered differentially regulated at a false discovery rate < 0.05. Only GO terms from the functional analysis with an adjusted p < 0.05 and at least four transcripts were considered significantly enriched. REVIGO was used to summarize GO terms to reduce redundancy and group according to similarity (right labels).
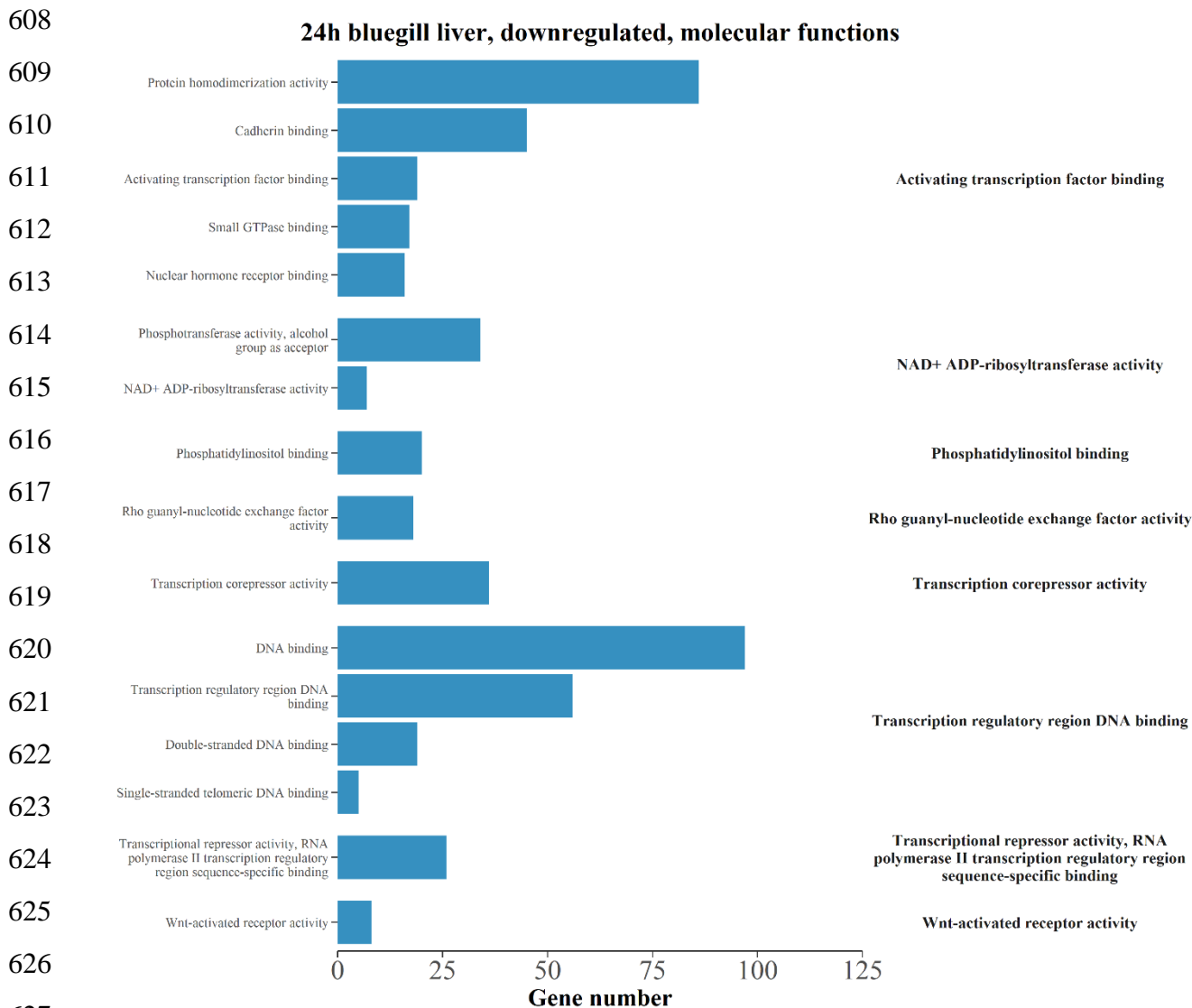
24 h bluegill liver, upregulated

Figure S16: Summary of enriched gene ontology (GO) terms associated with biological processes (green) and molecular functions (blue) in transcripts that were upregulated following 24 h of TFM (3-trifluoromethyl-4'-nitrophenol; 22.06 mg L^{-1} nominal) exposure in the liver of bluegill (*Lepomis macrochirus*). Transcripts were considered differentially regulated at a false discovery rate < 0.05. Only GO terms from the functional analysis with an adjusted p < 0.05 and at least four transcripts were considered significantly enriched. REVIGO was used to summarize GO terms to reduce redundancy and group according to similarity (right labels).

Figure S17: Summary of enriched gene ontology (GO) terms associated with molecular functions in transcripts that were downregulated following 24 h of TFM (3-trifluoromethyl-4'-nitrophenol; 22.06 mg L^{-1} nominal) exposure in the liver of bluegill (*Lepomis macrochirus*). Transcripts were considered differentially regulated at a false discovery rate < 0.05. Only GO terms from the functional analysis with an adjusted p < 0.05 and at least four transcripts were considered significantly enriched. REVIGO was used to summarize GO terms to reduce redundancy and group according to similarity (right labels).