1   **Genotype-specific transcriptional responses overshadow salinity effects in a marine diatom**

2   **sampled along the Baltic Sea salinity cline**

3

4

5   Eveline Pinseel[1*], Teofil Nakov[1], Koen Van den Berge[2,3,4], Kala M. Downey[1], Kathryn J. Judy[1], Olga

6   Kourtchenko[5], Anke Kremp[6], Elizabeth C. Ruck[1], Conny Sjöqvist[7], Mats Töpel[5], Anna Godhe[5] &

7   Andrew J. Alverson[1*]

8

9   [1] Department of Biological Sciences, University of Arkansas, Fayetteville, AR, USA

10   [2] Department of Statistics, University of California, Berkeley, CA, USA

11   [3] Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Ghent,

12   Belgium

13   [4] Bioinformatics Institute Ghent, Ghent University, Ghent, Belgium

14   [5] Department of Marine Sciences, University of Gothenburg, Gothenburg, Sweden

15   [6] Leibniz-Institute for Baltic Sea Research, Rostock, Germany

16   [7] Faculty of Science and Engineering, Åbo Akademi University, Turku, Finland

17

18   *Corresponding authors: eveline.pinseel@gmail.com, aja@uark.edu

19

20   The authors declare they have no conflict of interest.

ABSTRACT

The salinity gradient separating marine and freshwater environments represents a major ecological divide for microbiota, yet the mechanisms by which marine microbes have adapted to and ultimately diversified in freshwater environments are poorly understood. Here, we take advantage of a natural evolutionary experiment: the colonization of the brackish Baltic Sea by the ancestrally marine diatom *Skeletonema marinoi*. To understand how diatoms respond to low salinity, we characterized transcriptomic responses of *S. marinoi* grown in a common garden. Our experiment included eight genotypes from source populations spanning the Baltic Sea salinity cline. Changes in gene expression revealed a shared response to salinity across genotypes, where low salinities induced profound changes in cellular metabolism, including upregulation of carbon fixation and storage compound biosynthesis, and increased nutrient demand and oxidative stress. Nevertheless, the genotype effect overshadowed the salinity effect, as genotypes differed significantly in their response, both in the magnitude and direction of gene expression. Intraspecific differences included regulation of transcription and translation, nitrogen metabolism, cell signaling, and aerobic respiration. The high degree of intraspecific variation in gene expression observed here highlights an important but often overlooked source of biological variation associated with how diatoms respond and adapt to environmental change.

37 **INTRODUCTION**

38 The salinity gradient separating marine and freshwater environments represents one of the major

39 ecological divides structuring microbial diversity [1]. Differences in osmotic pressure impede marine–

40 freshwater transitions, and as a consequence, transitions are generally rare, occur on longer evolutionary

41 timescales [2, 3], and have repeatedly led to bursts of diversification in freshwater environments [4].

42 Comprehending the processes behind marine–freshwater habitat transitions is fundamental to both our

43 understanding of lineage diversification and habitat structuring on evolutionary time-scales [5], as well

44 as short term adaptive potential to climate change. The latter is rapidly altering marine environments:

45 melting ice caps, altered precipitation patterns, and changes in oceanic currents have led to freshening

46 of large regions as well as local changes in the seasonal or annual cycling of salinity regimes [6, 7]. Yet,

47 little is known about the cellular processes underlying acclimation and adaptation to low salinity

48 environments. Permanent establishment of ancestrally marine organisms in freshwaters depends on the

49 ability of the individual colonists to survive the initial hypo-osmotic stress, acclimate to low salinity,

50 and ultimately adapt to their new environment [8]. Consequently, such transitions are not likely to occur

51 instantly, but are rather thought to happen gradually [4, 5]. Euryhaline species that are able to tolerate

52 a wide range of salinities, or those inhabiting brackish environments, are most likely to succeed at

53 marine–freshwater transitions. Thus, to understand the cellular processes behind marine–freshwater

54 transitions, studies focusing on taxa that are most likely to colonize freshwater habitats are crucial.

55 Here, we take advantage of a natural evolutionary experiment currently underway: the

56 colonization of one of the world's largest brackish water bodies, the Baltic Sea, by the ancestrally

57 marine diatom *Skeletonema marinoi* (Fig. 1A). Geologically, the Baltic Sea is young, with sea ice from

58 the last glacial maximum having fully receded only 10 000 years ago and inundation of saline waters

59 from the adjacent North Sea occurring just 8 000 years ago [9]. Today, freshwater input from rivers and

60 precipitation, in combination with continued inflow of saline bottom-waters from the North Sea through

61 the Danish straits, results in a latitudinal and vertical salinity gradient ranging from near fresh to fully

62 marine conditions [9, 10] (Fig. 1A). This salinity gradient strongly structures biodiversity of aquatic

63 biota on both the species and population levels [11–13], including diatoms [14, 15]. This includes *S.*

64 *marinoi*, which is the dominant phytoplankton species and one of the main contributors to primary

3

65   production in the area [16, 17]. Paleoecological evidence showed that *S. marinoi* has been present in

66   the Baltic Sea since the marine inundation or shortly thereafter [18, 19]. Although *S. marinoi* is

67   ancestrally marine [20, 21], it can tolerate a wide range of salinities and is common along the entire

68   salinity gradient, from the North Sea coast to the upper reaches of the Baltic Sea [22]. Previous work

69   showed reduced gene flow between a high-salinity North Sea population of *S. marinoi* and a low-

70   salinity Baltic Sea population [22]. The Baltic population exhibited lower genetic diversity and optimal

71   growth at lower salinity, consistent with local adaptation [22]. Thus, *S. marinoi* presents an excellent

72   system for understanding the cellular mechanisms that govern adaptation of ancestrally marine diatoms

73   to low salinity environments.

74   We combined a laboratory common garden experiment with RNA-sequencing (RNA-seq) to

75   characterize the response of the euryhaline diatom *S. marinoi* to low salinity (Fig. 1). We collected eight

76   genotypes of *S. marinoi* along the Baltic Sea salinity cline, exposed them to a range of salinities, and

77   compared gene expression between high and low salinity treatments. Natural populations of *S. marinoi*

78   exhibit a broad range of variability in ecophysiological traits [23], so the inclusion of multiple genotypes

79   in our experiment allowed us to characterize potential variation in the salinity response as well, such as

80   which aspects of the response are shared and which ones differ among genotypes.

81

82   **MATERIAL & METHODS**

83

84   **Sample collection, culturing, and experimental design**

85   We collected sediment samples from eight locations across the Baltic Sea (Fig. 1A) and stored them in

86   the dark at 5 °C. We germinated *S. marinoi* resting cells into monoclonal cultures that were kept at their

87   native salinity (Table 1) [24]. Cells were grown at 12 °C on a 12:12 light:dark light regime (30 µmol

88   m^{-2} s^{-1} light intensity) for 12–26 months prior to the experiment. Strain identity was confirmed by

89   sequencing the LSU d1–d2 rDNA gene (see Supplementary Methods for details).

90   In our experiment, we grew eight different *S. marinoi* strains in triplicate for approximately two

91   weeks at three salinities (8, 16 and 24)—a design that included both biological (eight genotypes) and

92   technical replicates (three replicates per strain) (Fig. 1B). Further details about the experimental design,

93   RNA sequencing, and read processing are outlined in the Supplementary Methods. RNA-seq data are

94   available from the Sequence Read Archive (NCBI) under project number PRJNA772794.

95   We mapped RNA-seq reads against the reference genome of *S. marinoi* strain RO5AC v.1.1

96   (available from doi 10.5281/zenodo.5266588) with STAR [25], followed by gene-level read

97   quantification with HTSeq [26]. We performed statistical analyses in the R computing environment,

98   using edgeR [27] and stageR [28]. Gene ontology (GO) enrichment was done with TopGO (Over-

99   representation Analysis, ORA) [29] and CAMERA (Gene Set Enrichment Analysis, GSEA) [30]. We

100   obtained functional annotations for all genes using InterProScan [31], KofamKOALA [32], and

101   BLAST+ [33] searches against the Swissprot and Uniprot databases. We detected orthologs of *S.*

102   *marinoi* genes in other diatom genomes with OrthoFinder [34] and made protein targeting predictions

103   with MitoProt [35], HECTAR [36], SignalP [37], ASAFind [38], and TargetP [39]. Full details on the

104   bioinformatic pipeline, statistical models, and GO enrichment are outlined in the Supplementary

105   Methods, and the gene-level count data and R code used for the statistical analyses are available from

106   Zenodo (doi 10.5281/zenodo.5266588).

107

Hypothesis testing

108

109   We separately tested two sets of null hypotheses (Fig. 1C). The first set tested whether gene expression

110   was different across the salinity gradient for each genotype separately (Fig. 1C: individual genotypes)

111   and for all genotypes together (Fig. 1C: average of all genotypes), using a total of 27 contrasts (Fig.

112   1C). Compared to solely testing the average salinity effect, simultaneously accounting for the individual

113   genotypes increases the power to find differentially expressed (DE) genes, as the genotype effect

114   incorporates variability that would otherwise be unaccounted for. The second set of hypotheses tested

115   for an interaction effect between genotype and salinity, i.e., whether there are genotype-specific

116   responses to changes in the salinity (Fig. 1C). Here we defined 84 contrasts, testing each pairwise

117   combination of genotypes within all three salinity combinations (Fig. 1C). We tested the two sets of

118   hypotheses separately using a stage-wise testing procedure in stageR. This allowed us to control the

119   gene-level false discovery rate (FDR) at 5 % when testing multiple hypotheses simultaneously [28].

120

121   **RESULTS & DISCUSSION**

122

123   **General response of Baltic *S. marinoi* to low salinity environments**

124   The inclusion of multiple genotypes in our study design allowed us to characterize both the responses

125   of each individual genotype, and the average response across all genotypes (Fig. 1C). To obtain a

126   general overview of the response of *S. marinoi* to low salinity, we present the results of all three

127   contrasts (8-24, 8-16 and 16-24) together. Unless otherwise indicated, 'low salinity' refers to a lower

128   salinity (8 and/or 16) relative to a higher one (16 and/or 24), but does not distinguish between contrasts

129   unless necessary. This was done to gain a general overview of the response to low salinities across the

130   Baltic Sea salinity cline, and because GO enrichment suggested that many of the same processes are

131   enriched in the three salinity contrasts.

132   Growth reaction norms showed no substantial differences among genotypes and salinity

133   treatments, indicating that all genotypes grew well across a wide salinity range (Fig. 2). RNA-seq reads

134   of all genotypes mapped equally well to the reference genome. In the combined average- and genotype-

135   specific response (Fig. 1C), 7,905 of the 22,440 predicted genes in the *S. marinoi* genome were DE

136   using a 5 % FDR. The number of DE genes in the average response was higher than in any contrast of

137   the individual genotypes (Fig. 3A), with a total of 5,343 DE genes in at least one of three contrasts

138   tested in the average response. This is a result of combining all information from the different

139   genotypes, which resulted in 8x3 replicates for each salinity condition. As a consequence, the average

140   response allowed us to detect many more DE genes, including those with small effect-sizes, and shows

141   the benefit of including more than the standard three replicates used in many transcriptome studies.

142   The 8-24 contrasts consistently showed the most DE genes, and the least generally were found

143   in the 16-24 contrast (Fig. 3A). Thus, the largest drop in salinity (24 to 8), and the shift to lower salinity

144   (16 to 8), elicited the greatest transcriptomic responses. When including all DE genes, the number of

145   up- and downregulated genes was comparable within contrasts (Fig. 3A, Suppl. Fig. 1). However, when

146   only taking into account the top 100 genes based on P-value or logFC, substantially more genes were

147   upregulated in low salinities (Fig. 3B), indicating genes with the largest effect sizes or strongest

148   evidence for DE were more likely to be upregulated in low salinities. A similar pattern was seen in the

6

149 GO enrichment, where a CAMERA analysis found substantially more enriched GO terms that were

150 upregulated in low salinities (Fig. 3C). From these results it is clear that low salinities induce an

151 increased transcriptional response relative to high salinities. In the next paragraphs, we report on several

152 specific pathways that are up- or downregulated in low salinities. Unless otherwise noted, the data in

153 these sections represent the 8-24 contrast of the average response (Fig. 1C).

154

155 *Low salinities trigger profound changes in diatom metabolism*

156 *Skeletonema marinoi* exhibited profound changes in energy metabolism following exposure to low

157 salinity (Suppl. Figs 3-9). In lower salinities, *S. marinoi* experiences (i) significant upregulation of genes

158 from the photosynthetic electron transport chain, Calvin cycle, and chlorophyll biosynthesis (Fig. 4,

159 Suppl. Figs 2, 3A-B, 4A), and (ii) significant downregulation of genes involved in protein

160 ubiquitination, proteolysis, and aerobic respiration (Fig. 4, Suppl. Figs 2, 5A-B). The latter is evidenced

161 by downregulation of the mitochondrial electron transport chain and the TCA cycle, including

162 transcription factor *bZIP14*, which regulates the TCA cycle in diatoms [40] (Suppl. Fig. 5A-B). Thus,

163 the transcriptome suggests relatively more energy is acquired through light reactions, and there is less

164 intracellular recycling of proteins. This is in contrast to another euryhaline diatom, *Thalassiosira*

165 *weissflogii*, where carbon fixation was not impacted when cells were grown at different salinities [41].

166 Diatoms are known to accumulate storage compounds in unfavorable growth conditions [42,

167 43], and to modify the fatty acid and lipid composition of their membranes, which alters membrane

168 permeability and fluidity under salinity stress [44]. Here, we observed increased biosynthesis of storage

169 compounds, suggesting that low salinities are stressful to *S. marinoi*, despite their natural occurrence in

170 low-salinity regions of the Baltic Sea. Specifically, biosynthesis of two major storage compounds was

171 upregulated: the polysaccharide chrysolaminarin, i.e. β-1,3- and β-1,6-glucan polysaccharides [45], and

172 fatty acids that might be stored as triacylglycerols [46] (Fig. 4, Suppl. Figs 2, 6A). Four genes involved

173 herein were DE in all genotypes, underscoring that storage compound biosynthesis was a common

174 response in all investigated genotypes. In addition, further evidence from changes in cell membranes

175 came from upregulation of sulfolipids and a glycosyltransferase, which are involved in thylakoid

176 membranes and maintenance of membrane stability, respectively.

7

177   We found that tRNA-aminoacylation, translational elongation factors, and many ribosomal

178   proteins were upregulated, which together point to an increase in protein biosynthesis in low salinities

179   (Fig. 4, Suppl. Figs 2, 7). In parallel, protein refolding activity was upregulated (Fig. 4), suggesting

180   proteins need additional stress protection in low salinities. Several transcription factors were also

181   upregulated, including a putative heat stress transcription factor involved in DNA-binding of heat shock

182   promoter elements. Both nuclear division and microtubule-based movements were downregulated (Fig.

183   4, Suppl. Fig. 2), suggesting less mitosis in low salinities. Although our growth data measured from

184   relative chlorophyll *a* fluorescence do not support a decrease in cell division in low salinities (Fig. 2),

185   upregulation of chlorophyll biosynthesis in these conditions (Suppl. Fig. 4A) suggests that a decrease

186   in mitosis could have been masked by an increase in per-cell chlorophyll-content—an observation that

187   might have important consequences for similar experiments that use chlorophyll *a* as a proxy for

188   growth. In addition, two genes coding for an extracellular subtilisin-like serine protease were

189   upregulated in low salinities, as has also been observed in diatoms in response to other stressors such

190   as copper deficiency [47]. Finally, although activation of transposable elements has been linked to the

191   diatom stress response [48, 49], including *S. marinoi* [50], the majority of genes involved in transposon

192   activity (i.e., transposase, retrovirus-related Pol poly-protein) in the *S. marinoi* genome were not DE,

193   and if DE they tended to be downregulated in low salinities (Fig. 4, Suppl. Fig. 2).

194

195   ***Increased nutrient demand in low salinities***

196   Our data suggest that low salinities induce profound changes in nutrient demand, especially of nitrogen,

197   and potentially also highlight preferred nitrogen sources in *S. marinoi*. In fact, many transmembrane

198   transporters for essential nutrients such as nitrogen, phosphorus, molybdate, and sulphate were

199   upregulated in low salinities (Suppl. Figs 5D, 9). Specifically for nitrogen, *Nrt* nitrite/nitrate transporters

200   were highly upregulated, and to a lesser extent transporters for urea and ammonia (Suppl. Figs 5D, 9).

201   Most of the imported nitrogen is directed to the chloroplast, where nitrogen assimilation through

202   ferredoxin-nitrite reductase (*Fd-Nir*) and GSII-GOGAT$_{(Fd)}$ [51] was upregulated (Suppl. Fig. 5D). In

203   parallel, the anabolic part of the urea cycle was upregulated, including carbamoyl phosphate synthase

204   (*unCPS*) (Suppl. Figs 5C-D), suggestive of increased recycling of ammonia and biosynthesis of

8

205 arginino-succinate or arginine. Higher nutrient demands in low salinities likely reflect increased

206 biosynthesis of nitrogen-rich compounds such as amino acids and polyamines, essential for protein

207 biosynthesis and the stress response, respectively [52, 53]. In parallel, we detected a high number of

208 DE transporters for amino acids and polyamines, most of which were upregulated in low salinities

209 (Suppl. Fig. 9). Taken together, it appears that *S. marinoi* requires more nutrients in low-salinity

210 environments to support increased demand for proteins and other compounds. Consequently, *S. marinoi*

211 might struggle to maintain growth in low salinities under nutrient-deplete conditions, suggesting that

212 increasing nutrient loads in low-salinity regions of the Baltic Sea might benefit *S. marinoi* [54, 55]. In

213 contrast, silicic acid transporters were strongly downregulated in low salinities, and this downregulation

214 was most evident in the 16-24 salinity contrast (Suppl. Fig. 9). This could reflect decreased cell division

215 and/or changes in cell wall silicification.

216

217 *A diverse response to oxidative stress in low salinities*

218 Several lines of evidence suggest that *S. marinoi* experiences increased oxidative stress at low salinities,

219 as multiple mechanisms to deal with reactive oxygen species (ROS) were upregulated in low salinities.

220 This included genes involved in glutathione metabolism, ascorbate peroxidases, catalases, and

221 peroxiredoxin (Suppl. Fig. 6B). The xanthophyll cycle, which plays a critical role in protection from

222 oxidative stress due to excess light and ROS generated by other stressors [56], was also distinctly

223 upregulated (Suppl. Fig. 4B). Carotenoids for the xanthophyll cycle were produced primarily through

224 the non-mevalonate pathway, which was also upregulated in low salinities (Suppl. Fig. 4B).

225 We found upregulation of polyamine biosynthesis from ornithine via ornithine decarboxylase

226 (Suppl. Figs 5C-D). Polyamines play a complex role in abiotic stress responses, including salinity stress

227 in land plants, by increasing the activity of antioxidant enzymes, triggering the stress signal transduction

228 chain and serving an osmolyte function [53]. In diatoms, polyamines increase in response to both heat

229 and salinity stress [57, 58], and our data suggest a similar role in salinity acclimation.

230 Violaxanthin-de-epoxidase (xanthophyll cycle), and two genes involved in polyamine

231 biosynthesis were DE in all genotypes (Fig. 5), underscoring the highly conserved nature of the

232 oxidative stress response in *S. marinoi*. In addition, vitamin B6-binding activity was upregulated in low

9

233   salinities. Vitamin B6 is a vital cofactor of various enzymes, and plays a role in ROS deactivation in

234   land plants [59]. However, the response of *S. marinoi* to oxidative stress is complex, as several other

235   genes involved in ROS elimination were downregulated in low salinities (Suppl. Fig. 6B), including

236   superoxide dismutase (SOD), which is a first line of defense against ROS in land plants and macroalgae

237   [60, 61].

238

239   ***Osmotic stress response***

240   *Skeletonema marinoi* produces a variety of osmolytes. These are small organic molecules that help

241   mitigate hyperosmotic stress typical of marine environments [62–65]. Consequently, a decrease in

242   salinity should trigger a drop in osmolyte biosynthesis. Indeed, in low salinities, *S. marinoi* adjusted

243   intracellular osmolyte concentrations to hypoosmotic conditions by downregulating biosynthesis of the

244   osmolyte DMSP and upregulating breakdown of the osmolyte taurine via taurine dioxygenase (Suppl.

245   Fig. 6B). Although most genes involved in the diatom DMSP pathway remain unknown, *S. marinoi* has

246   a homolog of *TpMMT*, a methyltransferase that catalyzes a key reaction in DMSP biosynthesis in

247   diatoms [64]. This gene was strongly downregulated in low salinities (Suppl. Fig. 6B). Putative *BADH*

248   and *CDH* genes involved in the biosynthesis of the osmolyte betaine from choline [62] were present but

249   not DE. Betaine can also be synthesized from glycine by *TpGSDMT [62]*, and a possible homolog of

250   this gene was downregulated in all genotypes (Suppl. Fig. 6B). Genes involved in proline metabolism

251   were generally slightly upregulated (Suppl. Fig. 5D), suggesting proline is the preferred osmolyte in

252   low salinities.

253   Two genes that were strongly downregulated in low salinities in all genotypes could be involved

254   in the biosynthesis of a fifth osmolyte: ectoine (Fig. 5, Suppl. Fig. 8). These genes encode a bifunctional

255   aspartokinase/homoserine   dehydrogenase   (*Sm_g00011041*)   and   an   aspartate-semialdehyde

256   dehydrogenase (*Sm_g00011042*) (Fig. 5, Suppl. Fig. 8). Both are involved in the early steps of the

257   glycine/threonine/serine   pathway,   and   convert   aspartate   into   aspartate-semialdehyde   and/or

258   homoserine. The genome of *S. marinoi* contains several other homologs of both genes. When DE, these

259   homologs show opposite expression patterns to the aforementioned genes: they are upregulated in low

260   salinities, following the same expression patterns as other genes in this pathway (Suppl. Fig. 8). Peptide

261   targeting predictions showed that reactions in this pathway are compartmentalized across the

262   chloroplast, cytoplasm, and mitochondria, presumably allowing *S. marinoi* to run opposite reactions

263   simultaneously while avoiding futile cycles (Suppl. Fig. 8). Given their expression patterns and

264   compartmentalization, *Sm_g00011041* and *Sm_g00011042* are likely not involved in conventional

265   amino acid biosynthesis. Instead, one of their products, aspartate-semialdehyde, is a known precursor

266   for the osmolyte ectoine in bacteria (Suppl. Fig. 8) [66], and elevated levels of aspartate-semialdehyde

267   dehydrogenase have been detected in bacteria occupying high salinities [67]. Recently, marine diatoms

268   were found to both biosynthesize and import ectoine of bacterial origin [68]. Several *S. marinoi* genes

269   may be homologous to the bacterial ectoine genes *ectA*, *ectB*, and *ectC* that convert aspartate-

270   semialdehyde to ectoine. However, low sequence similarity (max. 47.8 %), and lack of downregulation

271   in low salinities, raises doubt about whether those genes are responsible for ectoine biosynthesis in *S.*

272   *marinoi*. Furthermore, all putative homologs received annotations different from ectoine-related genes

273   in Swissprot/Uniprot, suggesting they serve different functions from aforementioned ectoine genes. It

274   is possible that diatoms have other unknown genes involved in ectoine biosynthesis, or alternatively,

275   diatoms might provide ectoine precursors (e.g., aspartate-semialdehyde) to extracellular bacteria that

276   synthesize and return ectoine to the diatom. Such exchange of metabolites has previously been shown

277   to occur in diatom–bacteria interactions [69]. Our expression data are consistent with both scenarios

278   (Suppl. Fig. 8), and suggest ectoine might be an important osmolyte in diatoms.

279   Responses to osmotic stress also include shifts in cation import and export, such as sodium and

280   potassium [70]. Here, most identified potassium and sodium channels were either slightly upregulated

281   or not DE (Suppl. Fig. 9), and two detected aquaporins were either up- or downregulated in low

282   salinities. Six transporters for potassium, amino acids or unknown cations/solutes were DE in all

283   genotypes, often with large effect-sizes (Fig. 5), highlighting the conserved nature of this response in

284   *S. marinoi*.

285

286   **Genotype-specific data reveal intraspecific variation and a conserved core response to low salinity**

287   The genotype effect in our dataset overshadowed the salinity effect, indicating that genotypes differed

288   substantially in their response to low salinities. This was evidenced by a multidimensional scaling plot

289   and poisson-distance heatmap in which samples clustered primarily by genotype rather than salinity

290   (Fig. 6). Specifically, 1,791 genes were uniquely DE in only one genotype (Suppl. Fig. 10). The number

291   of uniquely DE genes ranged from 103 (genotype I) to 317 genes (genotype A) (Suppl. Fig. 10).

292   We defined a core response to low salinities by selecting genes that are DE in at least one

293   contrast of each genotype, which resulted in a set of 27 shared genes that are DE in each genotype (Figs

294   1C, 5). Obtaining this set of shared genes required subsetting the full set of DE genes (Fig. 1C), so the

295   5 % FDR for these 27 genes could not be guaranteed. However, the core response genes were

296   characterized by a combination of high logFC and low P-values (Suppl. Fig. 1), thus providing strong

297   evidence for DE in each genotype. These genes are also among the top DE genes in the full dataset (Fig.

298   1C: blue section): 13 overlapped with the top 25 DE genes as ranked by stageR's FDR-adjusted P-value

299   of the global null hypothesis (Padjscreen), 22 were part of the top 100 DE genes, and all were detected

300   within the top 225 DE genes. Core response genes upregulated in low salinities were involved in key

301   processes previously identified in the average response, including transport of amino acids and cations,

302   storage compound and polyamine biosynthesis, ROS deactivation, and transcription and translation

303   (Fig. 5). By contrast, core response genes that were downregulated in low salinities were involved in

304   proteolysis or the putative biosynthesis of the osmolyte ectoine (Fig. 5). Increasing the number of

305   technical replicates would likely enlarge the set of core response genes, as higher replicate numbers are

306   bound to improve detection of DE genes, especially those with small effect sizes [71]. Our set of 'core

307   response' genes is thus not exhaustive, but gives a first indication of which genes are likely to be part

308   of a conserved ancestral response to low salinity environments in *S. marinoi*.

309

310   **Interaction-effects reveal differences among genotypes in their response to low salinity**

311   A total of 3,857 genes showed interaction effects (Fig. 1C), i.e., significantly different expression

312   patterns between genotypes with a 5 % FDR. Of these, 2,820 differed between genotypes in the

313   magnitude of their response to low salinities, whereas far fewer (1,037) differed in the direction of their

314   response (Fig. 1C). However, 92 of the top 100 genes with interaction effects (ranked by stageR's FDR-

315   adjusted P-value of the global null hypothesis, Padjscreen) differed in the direction of their response

316   (Suppl. Fig. 11). Thus, although more genes overall showed differences in the magnitude of DE, those

317   with differences in direction of DE dominated the top set of interaction-effect genes.

318   The two classes of interaction-effect genes were enriched for different processes (Fig. 7, Suppl.

319   Fig. 12). Genes that differed significantly across genotypes in the magnitude of their response were

320   enriched for many of the same key processes identified in the average response, including

321   photosynthesis, glycolysis, and the biosynthesis of chlorophyll, carotenoids, and fatty acids (Fig. 7,

322   Suppl. Fig. 12). This suggests that there is considerable variation among *S. marinoi* genotypes in the

323   strength of the salinity response. By contrast, genes that differ significantly between genotypes in the

324   direction of their response were enriched for processes involving transcription regulation, peroxidase

325   activity, aerobic respiration, and nitrogen metabolism, including urea transmembrane transport (Fig. 7),

326   suggesting that genotypes differ, at least partially, in the invoked strategies to cope with low salinity.

327   The top 100 set of interaction-effect genes further suggested genotypes differ in the direction of their

328   response for cell wall and calcium-binding messenger proteins, as well as heat shock

329   proteins/chaperones (Suppl. Fig. 11), highlighting intraspecific differences in how salinity stress affects

330   protein function and cell-signalling pathways. For example, Ca^{2+}-signalling is involved in osmotic

331   sensing in diatoms [72], suggesting *S. marinoi* genotypes differ in how they perceive and respond to

332   osmotic stress.

333   Both classes of interaction-effect genes contained GO terms related to translational activity,

334   cell cycle progression, mitosis, and meiosis. In fact, two interaction-effect genes that show differences

335   in the direction of their response across genotypes coded for meiotic recombination protein *SPO11-2*

336   (Suppl. Fig. 11), which is thought to be expressed exclusively during meiosis [73, 74]. Sexual

337   reproduction in *S. marinoi* can be induced by shifts from low to higher salinity, e.g., 6 to 16 [73], but

338   our data suggest genotypes differ in the optimal salinity shift to induce meiosis. However, meiosis likely

339   never dominated during our experiment as growth rates were relatively stable over time (Fig. 2), and

340   gametes or auxospores were not observed. Although these observations could reflect intraspecific

341   differences in growth and sexual reproduction related to salinity, we cannot rule out that genotypes were

342   at different stages of their cell cycle when RNA was harvested.

13

343   The observed intraspecific variation in gene expression likely reflects an important mechanism

344   which allows *S. marinoi* to grow throughout the Baltic Sea. Diatom populations can harbour high levels

345   of both genotypic and phenotypic variation [23, 75–78]. Our finding of high levels of variation in gene

346   expression is a natural extension of this observation. Our study design did not allow testing whether the

347   observed intraspecific variation is related to the natural salinities at which the different genotypes occur,

348   as this would require sampling of multiple genotypes within populations. Nevertheless, visual

349   comparison of gene expression patterns did not show consistent differences across low- (D, F, I, J, K,

350   and P) and high-salinity (A and B) populations (e.g., Suppl. Figs 3-9), nor did those populations cluster

351   separately in the MDS-plot (Fig. 6A). This suggests that if signals of local adaptation along the Baltic

352   salinity cline [22] are due in part to differences in gene expression between high- and low-salinity

353   populations, then those differences are subtle. In any case, substantial intraspecific variation in gene

354   expression patterns in *S. marinoi* exists and is likely to be a crucial factor for survival, acclimation, and

355   adaptation to a dynamic environment such as the Baltic Sea, where in addition to a salinity gradient,

356   marked gradients and seasonal fluctuations in nutrients and temperature also occur [55]. The strong

357   degree of variation in gene expression within this population increases the chance that at least some

358   cells can survive rapidly fluctuating, potentially adverse, conditions in the short term. Assuming some

359   of this variation is heritable, variable gene expression can also enable long-term evolutionary adaptation

360   by providing targets for natural selection [8, 79]. In that sense, it may be that the observed variation in

361   gene expression is maintained by balancing selection.

362

363   **CONCLUSION**

364   Our study design, in which transcriptome data from eight genotypes were combined into one analysis,

365   allowed for a holistic view of the response of *S. marinoi* to low salinity conditions in the Baltic Sea, the

366   world's largest brackish water body. Transcriptome studies often include technical replicates of a single

367   genotype, but an increasing number of studies [49, 80] show that experiments without biological

368   replicates are unlikely to be generalizable, as different genotypes can exhibit markedly different patterns

369   of gene expression. Here, inclusion of both technical and biological replicates allowed us to pinpoint

370   both conserved and variable responses to low salinity.

Despite long-term presence of *S. marinoi* in brackish habitats in the Baltic Sea, our transcriptome data suggest that when Baltic *S. marinoi* is exposed to low salinities that mimic the natural Baltic salinity gradient, the diatom experiences elevated stress and increased energy and nutrient demands to maintain homeostasis. This suggests that *S. marinoi* has not fully adapted to low salinity regions of the Baltic Sea since it first colonized the Baltic some 9 000 years ago. Our analyses further revealed substantial levels of intraspecific variability in the response of *S. marinoi* to low salinities. Although our experimental design did not allow us to link this variability to the natural salinity of the source population, it highlights an important source of biological variation in natural populations of *S. marinoi* and presumably other diatoms as well. Variable gene expression could be an important mechanism at which diatoms respond and adapt to environmental change, and ultimately diversified in a wide range of marine, freshwater, and terrestrial habitats worldwide [4, 81].

ACKNOWLEDGEMENTS

COMPETING INTERESTS

The authors declare they have no conflict of interest.

REFERENCES

1. Lozupone CA, Knight R. Global patterns in bacterial diversity. *Proc Natl Acad Sci U S A* 2007; **104**: 11436–11440.

2. Logares R, Bråte J, Bertilsson S, Clasen JL, Shalchian-Tabrizi K, Rengefors K. Infrequent marine-freshwater transitions in the microbial world. *Trends Microbiol* 2009; **17**: 414–422.

3. Cavalier-Smith T. Megaphylogeny, cell body plans, adaptive zones: causes and timing of eukaryote basal radiations. *J Eukaryot Microbiol* 2009; **56**: 26–33.

4. Nakov T, Beaulieu JM, Alverson AJ. Diatoms diversify and turn over faster in freshwater than marine environments. *Evolution* 2019; **73**: 2497–2511.

5. Dittami SM, Heesch S, Olsen JL, Collén J. Transitions between marine and freshwater environments provide new clues about the origins of multicellular plants and algae. *J Phycol* 2017; **53**: 731–745.

6. Dickson B, Yashayaev I, Meincke J, Turrell B, Dye S, Holfort J. Rapid freshening of the deep North Atlantic Ocean over the past four decades. *Nature* 2002; **416**: 832–837.

7. Aretxabaleta AL, Smith KW, Kalra TS. Regime Changes in Global Sea Surface Salinity Trend. *J Mar Sci Eng* 2017; **5**: 57.

8. López-Maury L, Marguerat S, Bähler J. Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nat Rev Genet* 2008; **9**: 583–593.

9. Björck S. A review of the history of the Baltic Sea, 13.0-8.0 ka BP. *Quat Int* 1995; **27**: 19–40.

10. Krauss W. Baltic Sea Circulation. In: Steele JH (ed). *Encyclopedia of Ocean Sciences*. 2001. Academic Press, Oxford, pp 236–244.

11. Zettler ML, Schiedek D, Bobertz B. Benthic biodiversity indices versus salinity gradient in the southern Baltic Sea. *Mar Pollut Bull* 2007; **55**: 258–270.

12. Telesh I, Schubert H, Skarlato S. Life in the salinity gradient: Discovering mechanisms behind a new biodiversity pattern. *Estuar Coast Shelf Sci* 2013; **135**: 317–327.

13. Johannesson K, Le Moan A, Perini S, André C. A Darwinian Laboratory of Multiple Contact Zones. *Trends Ecol Evol* 2020; **35**: 1021–1036.

14. Busse S, Snoeijs P. Gradient responses of diatom communities in the Bothnian Bay, northern Baltic Sea. *Nova Hedwigia* 2002; **74**: 501–525.

15. Gasiūnaitė ZR, Cardoso AC, Heiskanen A-S, Henriksen P, Kauppila P, Olenina I, et al. Seasonality of coastal phytoplankton in the Baltic Sea: Influence of salinity and eutrophication. *Estuar Coast Shelf Sci* 2005; **65**:

429   239–252.

430   16. Jochem F. Distribution and importance of autotrophic ultraplankton in a boreal inshore area (Kiel Bight,

431   Western Baltic). *Mar Ecol Prog Ser* 1989; **53**: 153–168.

432   17. Lange CB, Hasle GR, Syvertsen EE. Seasonal cycle of diatoms in the Skagerrak, North Atlantic, with

433   emphasis on the period 1980–1990. *Sarsia* 1992; **77**: 173–187.

434   18. van Wirdum F, Andrén E, Wienholz D, Kotthoff U, Moros M, Fanget A-S, et al. Middle to Late Holocene

435   Variations in Salinity and Primary Productivity in the Central Baltic Sea: A Multiproxy Study From the

436   Landsort Deep. *Frontiers in Marine Science* 2019; **6**: 51.

437   19. Warnock J, Andrén E, Juggins S, Lewis J, Ryves DB, Andrén T, et al. A high-resolution diatom-based Middle

438   and Late Holocene environmental history of the Little Belt region, Baltic Sea. *Boreas* 2020; **49**: 1–16.

439   20. Ellegaard M, Godhe A, Härnström K, McQuoid M. The species concept in a marine diatom: LSU rDNA–

440   based phylogenetic differentiation in *Skeletonema marinoi/dohrnii (*Bacillariophyceae) is not reflected in

441   morphology. *Phycologia* 2008; **47**: 156–167.

442   21. Alverson AJ. Timing marine–freshwater transitions in the diatom order Thalassiosirales. *Paleobiology* 2014;

443   **40**: 91–101.

444   22. Sjöqvist C, Godhe A, Jonsson PR, Sundqvist L, Kremp A. Local adaptation and oceanographic connectivity

445   patterns explain genetic differentiation of a marine diatom across the North Sea-Baltic Sea salinity gradient.

446   *Mol Ecol* 2015; **24**: 2871–2885.

447   23. Kremp A, Godhe A, Egardt J, Dupont S, Suikkanen S, Casabianca S, et al. Intraspecific variability in the

448   response of bloom-forming marine microalgae to changed climate conditions. *Ecol Evol* 2012; **2**: 1195–1207.

449   24. Godhe A, Härnström K. Linking the planktonic and benthic habitat: genetic structure of the marine diatom

450   *Skeletonema marinoi*. *Mol Ecol* 2010; **19**: 4478–4490.

451   25. Dobin A, Gingeras TR. Mapping RNA-seq Reads with STAR. *Curr Protoc Bioinformatics* 2015; **51**: 11.14.1–

452   11.14.19.

453   26. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data.

454   *Bioinformatics* 2015; **31**: 166–169.

455   27. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis

456   of digital gene expression data. *Bioinformatics* 2010; **26**: 139–140.

457   28. Van den Berge K, Soneson C, Robinson MD, Clement L. stageR: a general stage-wise method for controlling

458   the gene-level false discovery rate in differential expression and differential transcript usage. *Genome Biol*

17

459   2017; **18**: 151.

460   29. Alexa A, Rahnenführer J. Gene set enrichment analysis with topGO. *Bioconductor Improv* 2009; **27**.

461   30. Wu D, Smyth GK. Camera: a competitive gene set test accounting for inter-gene correlation. *Nucleic Acids*
462   *Res* 2012; **40**: e133.

463   31. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein
464   function classification. *Bioinformatics* 2014; **30**: 1236–1240.

465   32. Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, et al. KofamKOALA: KEGG
466   Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* 2020; **36**: 2251–
467   2252.

468   33. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990;
469   **215**: 403–410.

470   34. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol*
471   2019; **20**: 238.

472   35. Claros MG. MitoProt, a Macintosh application for studying mitochondrial proteins. *Comput Appl Biosci* 1995;
473   **11**: 441–447.

474   36. Gschloessl B, Guermeur Y, Cock JM. HECTAR: a method to predict subcellular targeting in heterokonts.
475   *BMC Bioinformatics* 2008; **9**: 393.

476   37. Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. *J Mol*
477   *Biol* 2004; **340**: 783–795.

478   38. Gruber A, Rocap G, Kroth PG, Armbrust EV, Mock T. Plastid proteome prediction for diatoms and other
479   algae with secondary plastids of the red lineage. *Plant J* 2015; **81**: 519–528.

480   39. Almagro Armenteros JJ, Salvatore M, Emanuelsson O, Winther O, von Heijne G, Elofsson A, et al. Detecting
481   sequence signals in targeting peptides using deep learning. *Life Sci Alliance* 2019; **2**: e201900429.

482   40. Matthijs M, Fabris M, Obata T, Foubert I, Franco-Zorrilla JM, Solano R, et al. The transcription factor bZIP14
483   regulates the TCA cycle in the diatom *Phaeodactylum tricornutum*. *EMBO J* 2017; **36**: 1559–1576.

484   41. Bussard A, Corre E, Hubas C, Duvernois-Berthet E, Le Corguillé G, Jourdren L, et al. Physiological
485   adjustments and transcriptome reprogramming are involved in the acclimation to salinity gradients in diatoms.
486   *Environ Microbiol* 2017; **19**: 909–925.

487   42. Sayanova O, Mimouni V, Ulmann L, Morant-Manceau A, Pasquet V, Schoefs B, et al. Modulation of lipid
488   biosynthesis by stress in diatoms. *Philos Trans R Soc Lond B Biol Sci* 2017; **372**: 20160407.

18

489   43. Vårum KM, Myklestad S. Effects of light, salinity and nutrient limitation on the production of β-1,3-d-glucan

490   and exo-d-glucanase activity in Skeletonema costatum (Grev.) Cleve. *J Exp Mar Bio Ecol* 1984; **83**: 13–25.

491   44. Chen G-Q, Jiang Y, Chen F. Salt-induced alterations in lipid composition of diatom *Nitzschia laevis*

492   (Bacillariophyceae) under heterotrophic culture condition. *J Phycol* 2008; **44**: 1309–1314.

493   45. Kroth PG, Chiovitti A, Gruber A, Martin-Jezequel V, Mock T, Parker MS, et al. A model for carbohydrate

494   metabolism in the diatom *Phaeodactylum tricornutum* deduced from comparative whole genome analysis.

495   *PLoS One* 2008; **3**: e1426.

496   46. Zulu NN, Zienkiewicz K, Vollheyde K, Feussner I. Current trends to comprehend lipid metabolism in diatoms.

497   *Prog Lipid Res* 2018; **70**: 1–16.

498   47. Kong L, Price NM. Transcriptomes of an oceanic diatom reveal the initial and final stages of acclimation to

499   copper deficiency. *Environ Microbiol* 2021.

500   48. Maumus F, Allen AE, Mhiri C, Hu H, Jabbari K, Vardi A, et al. Potential impact of stress activated

501   retrotransposons on genome evolution in a marine diatom. *BMC Genomics* 2009; **10**: 624.

502   49. Pargana A, Musacchia F, Sanges R, Russo MT, Ferrante MI, Bowler C, et al. Intraspecific Diversity in the

503   Cold Stress Response of Transposable Elements in the Diatom *Leptocylindrus aporus*. *Genes* 2019; **11**: 9.

504   50. Amato A, Sabatino V, Nylund GM, Bergkvist J, Basu S, Andersson MX, et al. Grazer-induced transcriptomic

505   and metabolomic response of the chain-forming diatom *Skeletonema marinoi*. *ISME J* 2018; **12**: 1594–1604.

506   51. Smith SR, Dupont CL, McCarthy JK, Broddrick JT, Oborník M, Horák A, et al. Evolution and regulation of

507   nitrogen flux through compartmentalized metabolic networks in a marine diatom. *Nat Commun* 2019; **10**:

508   4552.

509   52. Mansour MMF. Nitrogen containing compounds and adaptation of plants to salinity stress. *Biol Plant* 2000;

510   **43**: 491–500.

511   53. Chen D, Shao Q, Yin L, Younis A, Zheng B. Polyamine Function in Plants: Metabolism, Regulation on

512   Development, and Roles in Abiotic Stress Responses. *Front Plant Sci* 2018; **9**: 1945.

513   54. Voss M, Dippner JW, Humborg C, Hürdler J, Korth F, Neumann T, et al. History and scenarios of future

514   development of Baltic Sea eutrophication. *Estuar Coast Shelf Sci* 2011; **92**: 307–322.

515   55. Savchuk OP. Large-scale nutrient dynamics in the Baltic sea, 1970–2016. *Front Mar Sci* 2018; **5**: 95.

516   56. Latowski D, Kuczyńska P, Strzałka K. Xanthophyll cycle--a mechanism protecting plants against oxidative

517   stress. *Redox Rep* 2011; **16**: 78–90.

518   57. Liu Q, Nishibori N, Imai I, Hollibaugh JT. Response of polyamine pools in marine phytoplankton to nutrient

519   limitation and variation in temperature and salinity. *Mar Ecol Prog Ser* 2016; **544**: 93–105.

520   58. Scoccianti V, Penna A, Penna N, Magnani M. Effect of heat stress on polyamine content and protein pattern

521   in *Skeletonema costatum*. *Mar Biol* 1995; **121**: 549–554.

522   59. Colinas M, Eisenhut M, Tohge T, Pesquera M, Fernie AR, Weber APM, et al. Balancing of B6 Vitamers Is

523   Essential for Plant Development and Metabolism in *Arabidopsis*. *Plant Cell* 2016; **28**: 439–453.

524   60. Alscher RG, Erturk N, Heath LS. Role of superoxide dismutases (SODs) in controlling oxidative stress in

525   plants. *J Exp Bot* 2002; **53**: 1331–1341.

526   61. Kumar M, Kumari P, Gupta V, Reddy CRK, Jha B. Biochemical responses of red alga *Gracilaria corticata*

527   (Gracilariales, Rhodophyta) to salinity induced oxidative stress. *J Exp Mar Bio Ecol* 2010; **391**: 27–34.

528   62. Kageyama H, Tanaka Y, Takabe T. Biosynthetic pathways of glycinebetaine in *Thalassiosira pseudonana*;

529   functional characterization of enzyme catalyzing three-step methylation of glycine. *Plant Physiol Biochem*

530   2018; **127**: 248–255.

531   63. Lyon BR, Bennett-Mintz JM, Lee PA, Janech MG, DiTullio GR. Role of dimethylsulfoniopropionate as an

532   osmoprotectant following gradual salinity shifts in the sea-ice diatom *Fragilariopsis cylindrus*. *Environ Chem*

533   2016; **13**: 181–194.

534   64. Kageyama H, Tanaka Y, Shibata A, Waditee-Sirisattha R, Takabe T. Dimethylsulfoniopropionate

535   biosynthesis in a diatom *Thalassiosira pseudonana*: Identification of a gene encoding MTHB-

536   methyltransferase. *Arch Biochem Biophys* 2018; **645**: 100–106.

537   65. Krell A, Funck D, Plettner I, John U, Dieckmann G. Regulation of proline metabolism under salt stress in the

538   psychrophilic diatom *Fragilariopsis cylindrus* (Bacillariophyceae). *J Phycol* 2007; **43**: 753–762.

539   66. Vargas C, Argandoña M, Reina-Bueno M, Rodríguez-Moya J, Fernández-Aunión C, Nieto JJ. Unravelling

540   the adaptation responses to osmotic and temperature stress in *Chromohalobacter salexigens*, a bacterium with

541   broad salinity tolerance. *Saline Systems* 2008; **4**: 14.

542   67. Khmelenina VN, Sakharovskii VG, Reshetnikov AS, Trotsenko YA. Synthesis of osmoprotectants by

543   halophilic and alkaliphilic methanotrophs. *Microbiology* . 2000. , **69**: 381–386

544   68. Fenizia S, Thume K, Wirgenings M, Pohnert G. Ectoine from Bacterial and Algal Origin Is a Compatible

545   Solute in Microalgae. *Mar Drugs* 2020; **18**: 42.

546   69. Amin SA, Hmelo LR, van Tol HM, Durham BP, Carlson LT, Heal KR, et al. Interaction and signalling

547   between a cosmopolitan phytoplankton and associated bacteria. *Nature* 2015; **522**: 98–101.

548   70. Latta LC, Weider LJ, Colbourne JK, Pfrender ME. The evolution of salinity tolerance in *Daphnia*: a functional

549   genomics approach. *Ecol Lett* 2012; **15**: 794–802.

550   71. Van den Berge K, Hembach KM, Soneson C, Tiberi S, Clement L, Love MI, et al. RNA Sequencing Data:

551   Hitchhiker's Guide to Expression Analysis. *Annu Rev Biomed Data Sci* 2019; **2**: 139–173.

552   72. Helliwell KE, Kleiner FH, Hardstaff H, Chrachri A, Gaikwad T, Salmon D, et al. Spatiotemporal patterns of

553   intracellular Ca2+ signalling govern hypo-osmotic stress resilience in marine diatoms. *New Phytol* 2021; **230**:

554   155–170.

555   73. Ferrante MI, Entrambasaguas L, Johansson M, Töpel M, Kremp A, Montresor M, et al. Exploring Molecular

556   Signs of Sex in the Marine Diatom *Skeletonema marinoi*. *Genes* 2019; **10**: 494.

557   74. Bilcke G, Van den Berge K, De Decker S, Bonneure E, Poulsen N, Bulankova P, et al. Mating type specific

558   transcriptomic response to sex inducing pheromone in the pennate diatom *Seminavis robusta*. *ISME J* 2021;

559   **15**: 562–576.

560   75. Ajani PA, Petrou K, Larsson ME, Nielsen DA, Burke J, Murray SA. Phenotypic trait variability as an

561   indication of adaptive capacity in a cosmopolitan marine diatom. *Environ Microbiol* 2021; **23**: 207–223.

562   76. Sjöqvist CO, Kremp A. Genetic diversity affects ecological performance and stress response of marine diatom

563   populations. *ISME J* 2016; **10**: 2755–2766.

564   77. Godhe A, Rynearson T. The role of intraspecific variation in the ecological and evolutionary success of

565   diatoms in changing environments. *Philos Trans R Soc Lond B Biol Sci* 2017; **372**: 20160399.

566   78. Bulankova P, Sekulić M, Jallet D, Nef C, van Oosterhout C, Delmont TO, et al. Mitotic recombination

567   between homologous chromosomes drives genomic diversity in diatoms. *Curr Biol* 2021; **31**: 3221–3232.e9.

568   79. Gomez-Mestre I, Jovani R. A heuristic model on the role of plasticity in adaptive evolution: plasticity

569   increases adaptation, population viability and genetic variation. *Proc Biol Sci* 2013; **280**: 20131869.

570   80. Nakov T, Judy KJ, Downey KM, Ruck EC, Alverson AJ. Transcriptional Response of Osmolyte Synthetic

571   Pathways and Membrane Transporters in a Euryhaline Diatom During Long-term Acclimation to a Salinity

572   Gradient. *J Phycol* 2020; **56**: 1712–1728.

573   81. Pinseel E, Janssens SB, Verleyen E, Vanormelingen P, Kohler TJ, Biersma EM, et al. Global radiation in a

574   rare biosphere soil diatom. *Nat Commun* 2020; **11**: 2382.

575   82. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO summarizes and visualizes long lists of gene ontology

576   terms. *PLoS One* 2011; **6**: e21800.

TABLES

Table 1. Details of the *S. marinoi* strains used in this study. The salinity values indicate the salinity of the natural sample from which the respective strains were isolated ('original salinity') and in which they were maintained prior to the experiment ('culture salinity'). GPS coordinates indicated with an asterisk (*) represent approximate sampling locations.

| Collection ID	Strain ID	Field ID	Country	GPS (N/E)	Collector	Isolation date	Culture medium	Original salinity	Culture salinity
AJA304	A.2.21b	A	Sweden	58.02868/11.13738 (*)	A.Godhe	2017-03-28	L1	15-33	24
AJA305	B.2.19b	B	Sweden	55.97744/12.69058	A.Godhe	2017-04-07	ASW	12-15	16
AJA332	D.1.27a	D	Sweden	58.33200/16.70583	A.Godhe, B.Andersson	2018-05-14	WC + salt	8-9	8
AJA328	F.1.2a	F	Sweden	63.65317/18.95200 (*)	A.Godhe	2018-03-15	WC + salt	~8	8
AJA311	I.3.11a	I	Finland	60.18000/25.50700	A.Kremp	2017-03-09	WC + salt	5-6	5
AJA313	J.3.42b	J	Finland	60.38964/27.37518 (*)	A.Kremp	2018-04-22	WC + salt	4-5	8
AJA318	K.3.3a	K	Estonia	57.81670/22.28330	S.Sildever	2018-05-23	WC + salt	~8	8
AJA333	P.2.6a	P	Poland	54.44778/18.57611	A.Witkowski	2018-05-16	WC + salt	5-7	5

FIGURES
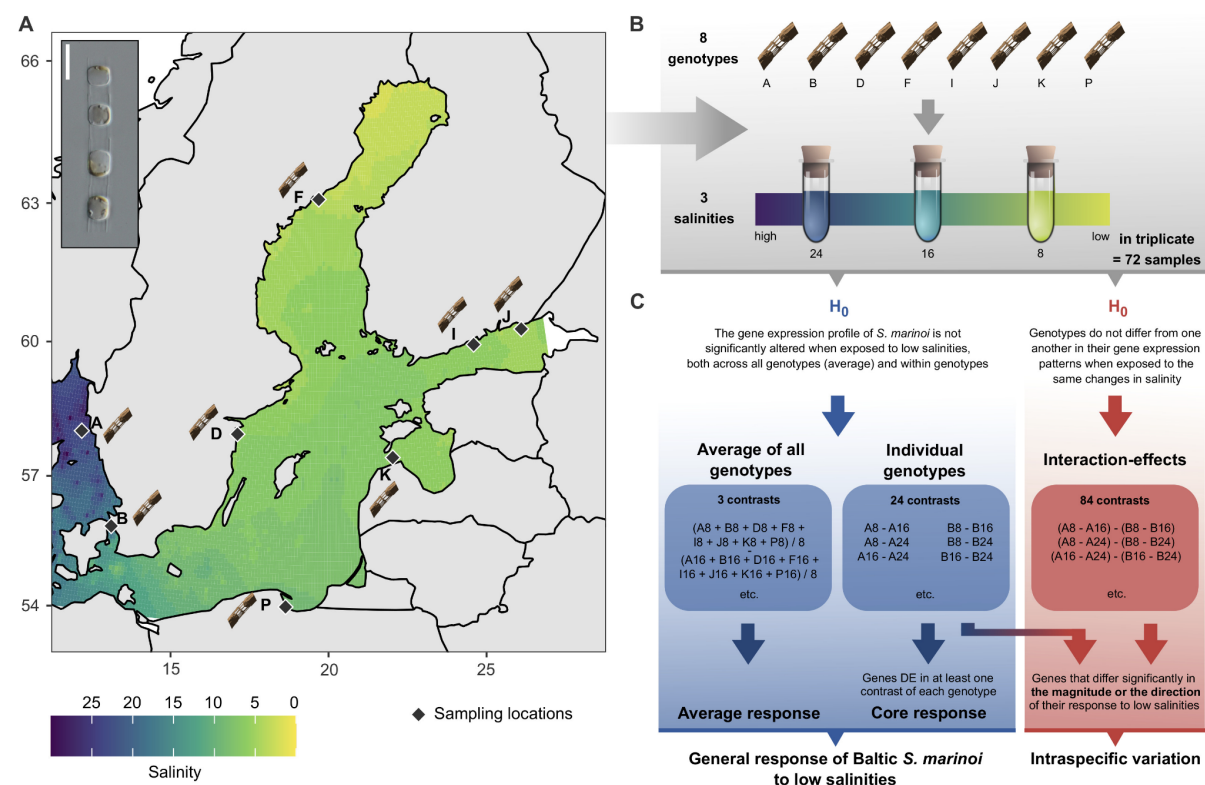
Fig. 1. Experimental design. a Field sampling. Natural salinity gradient in the Baltic Sea based on salinity measurements from surface samples (0-10 m depth) and interpolated across the Baltic Sea for the period 1990-2020. Salinity measurements were downloaded from ICES (ICES Dataset on Ocean Hydrography, 2020. ICES, Copenhagen) and Sharkweb (https://sharkweb.smhi.se/hamta-data/). Diamonds identify sampling locations for *S. marinoi*. The inset figure shows a light micrograph of a *S. marinoi* culture (scale bar = 10 μm). **b Laboratory experiment.** Experimental design of the laboratory experiment carried out in this study. Eight strains of *S. marinoi* were exposed to three salinity treatments (8, 16 and 24) in triplicate, resulting in 72 RNA-seq libraries. **c Statistical analyses.** Overview of the null hypotheses and contrasts tested in this study. Our experimental design allowed characterization of the general response of *S. marinoi* to low salinities as well as intraspecific variation. The lower blue arrows indicate which data were incorporated in the average and core responses, which together were used to define the general response of *S. marinoi*. Genes with significant interaction effects were subdivided in two categories using logFC values of the genotype-specific effects (blue-red gradient arrow), distinguishing genes that differed significantly in either the magnitude or direction of their response to low salinities. The first category includes genes that were DE in one genotype but not the others, or that were DE in multiple genotypes but with significant differences in logFC values in the same direction. Genes of the second category were significantly upregulated in some genotypes, whereas they were significantly downregulated in other genotypes.
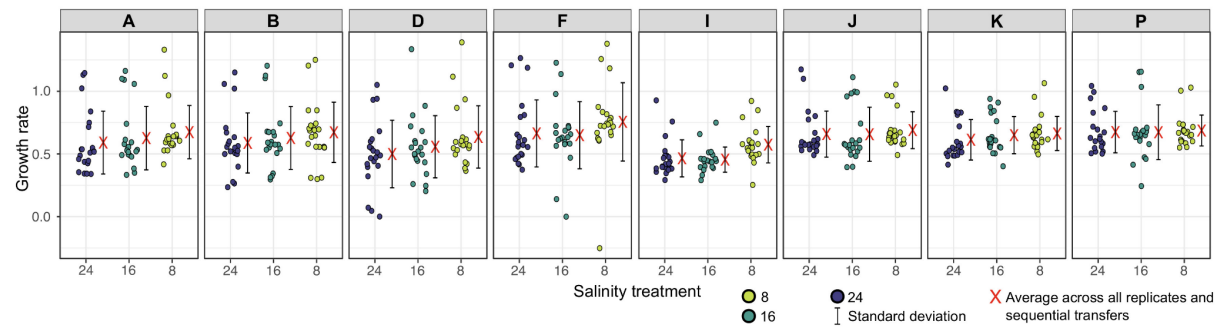
Fig. 2. Growth response of Baltic *S. marinoi* in low salinities. Growth rates of the eight *S. marinoi* genotypes examined in this study at three different salinities. The letters in the individual panels correspond with the sampling locations in Table 1 ('Field ID') and Fig. 1A. Each point represents a single estimate of the slope of the logarithm of in vivo relative fluorescence against time for each sequential transfer, using a horizontal jitter of points to avoid overplotting.
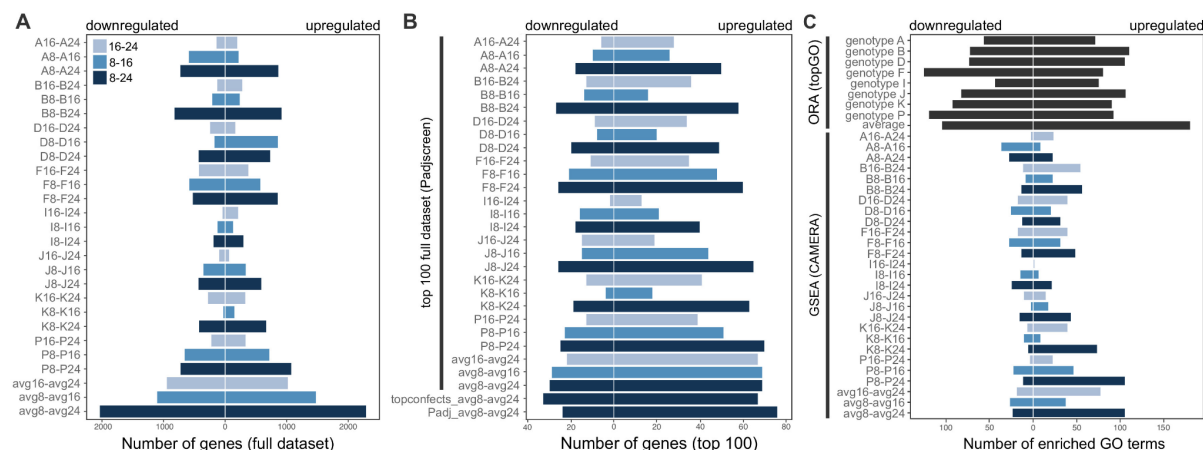
Fig. 3. Transcriptome response of Baltic *S. marinoi* to low salinities. a Number of DE genes at a 5 % FDR-level in the average and genotype-specific effects. The number of DE genes is indicated separately for each contrast, distinguishing between genes that are up- or downregulated. **b** Direction of DE in the top 100 genes of the average and genotype-specific effects as selected by P-value or logFC. For each contrast in the average and genotype-specific effects (vertical black bar), the direction of DE is indicated for the top 100 genes selected by stageR's FDR-adjusted P-value of the global null hypothesis (Padjscreen). Thus, although a gene can have a high P-value on a dataset-wide level, it is not necessarily DE in each individual contrast. In addition, we show the top 100 genes selected by logFC (topconfects) and the contrast-specific 5 % FDR-controlled P-value (Padj) for the 8-24 contrast of the average effects, as this contrast showed the greatest number of DE genes in **a**. **c** Number of enriched GO terms. Two types of GO enrichment are shown: Over-representation Analysis, ORA (in topGO) and Gene Set Enrichment Analysis, GSEA (in CAMERA). For ORA, we defined two sets of genes in each genotype/average response for which GO enrichment was performed separately: genes upregulated in low salinities, and genes downregulated in low salinities. Further details on how this selection was performed can be found in the Supplementary Methods. For GSEA, we performed GO enrichment on each contrast of the genotype-specific and average effects separately. Here, the number of up- and downregulated GO terms represents the output classification by CAMERA. The number of enriched GO terms includes Biological Process, Molecular Function and Cellular Component GO terms, prior to removal of redundant GO terms by REVIGO [82] (see Supplementary Methods).
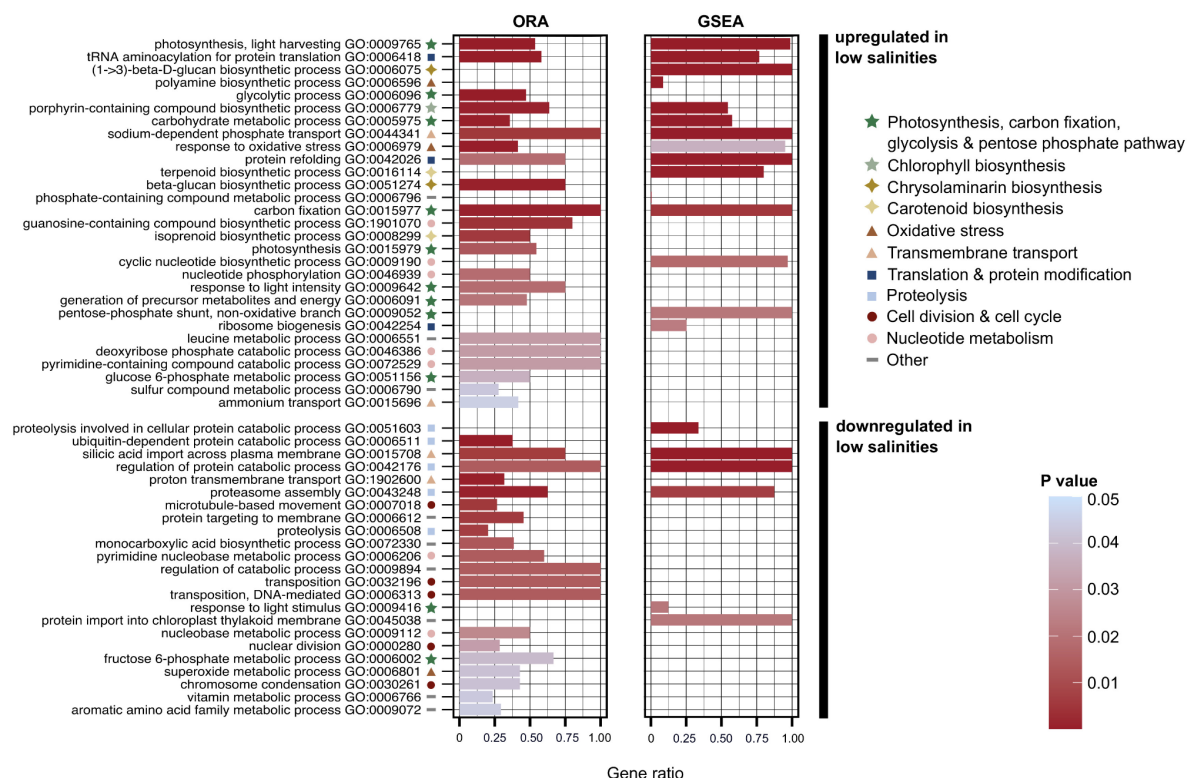
**Fig 4. GO enrichment on the average response of *S. marinoi* to low salinities: Biological Process**. The results of two types of GO enrichment analyses are shown: ORA (in topGO, Fisher's exact test, *elim* algorithm) and GSEA (in CAMERA), after removal of redundant terms by REVIGO. For ORA, we classified the total set of DE genes in the average response into two categories, distinguishing between genes that are up- or downregulated in low salinities, regardless of salinity contrast (see Supplementary Methods for more details). For CAMERA, we performed GSEA analyses on each individual contrast separately, showing only the 8-24 contrast in this figure. Barplot height indicates the proportion of genes that are DE with a given GO-term to the total number of genes with this GO-term in the genome of *S. marinoi*. The barplots are colored according to P-value. Within the set of up- and downregulated genes, the GO-terms are ranked from lowest to highest P-value, using the lowest of two P-values from ORA or GSEA. Symbols indicate major categories of cellular processes to which a GO-term belongs. Only Biological Process GO-terms are shown.
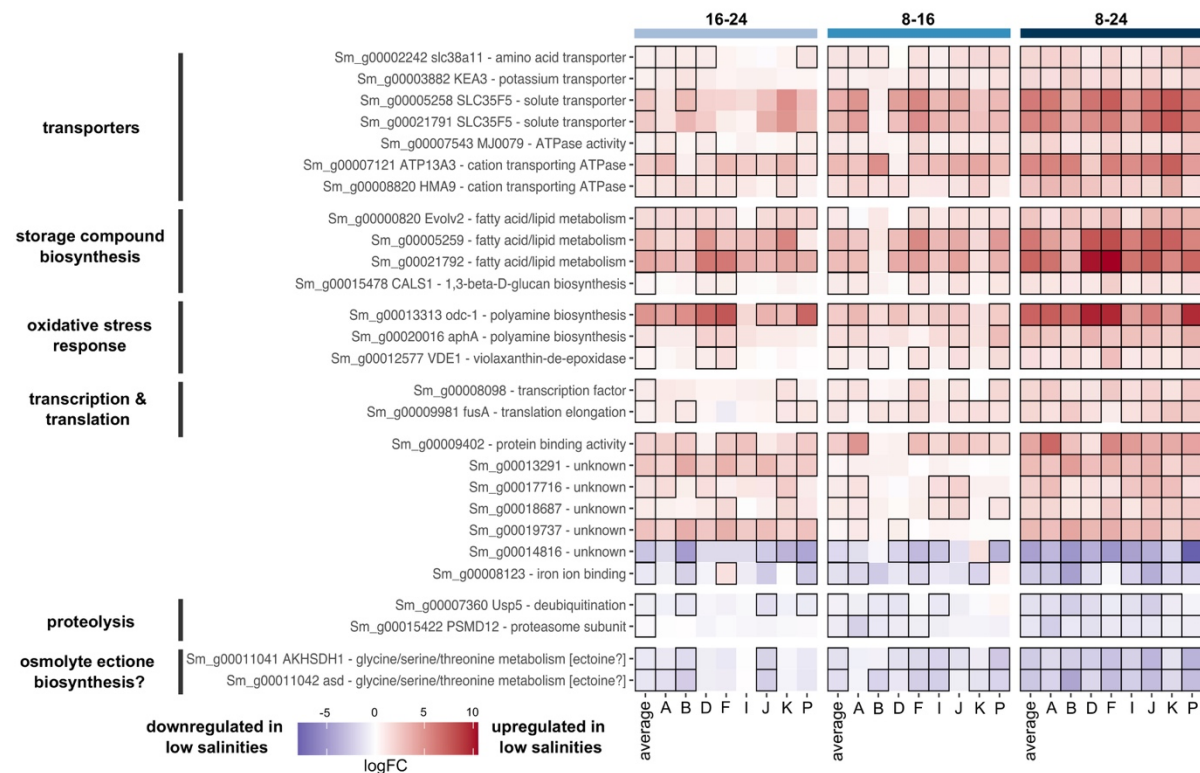
Fig. 5. Set of genes that are DE in at least one contrast of each genotype: the core response. The heatmap shows logFC values for the individual genotypes and average response of the 27 core response genes. The three salinity combinations are indicated on top of the figure. Contrasts that were significant are outlined in black. Rownames specify gene names and functional annotation. When DE, all genes are consistently up- or downregulated in low salinities in each genotype, with the exception of gene *Sm_g00008123*.
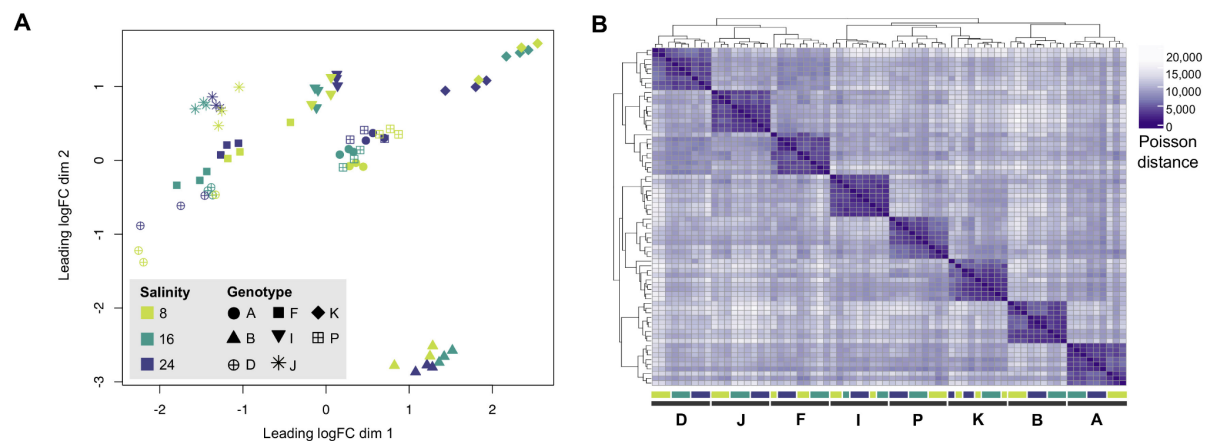
Fig. 6. Intraspecific variation in the response of Baltic *S. marinoi* to low salinities. a Multidimensional scaling (MDS) plot for the full dataset, showing that samples cluster primarily by genotype rather than salinity. Distances between the samples are based on logFC changes in the top 500 genes. **b** Poisson-distance heatmap of the full dataset. Colored bars below the heatmap indicate the position of samples belonging to different genotypes and salinities (Fig. 1 a), showing that samples of different genotypes cluster together.
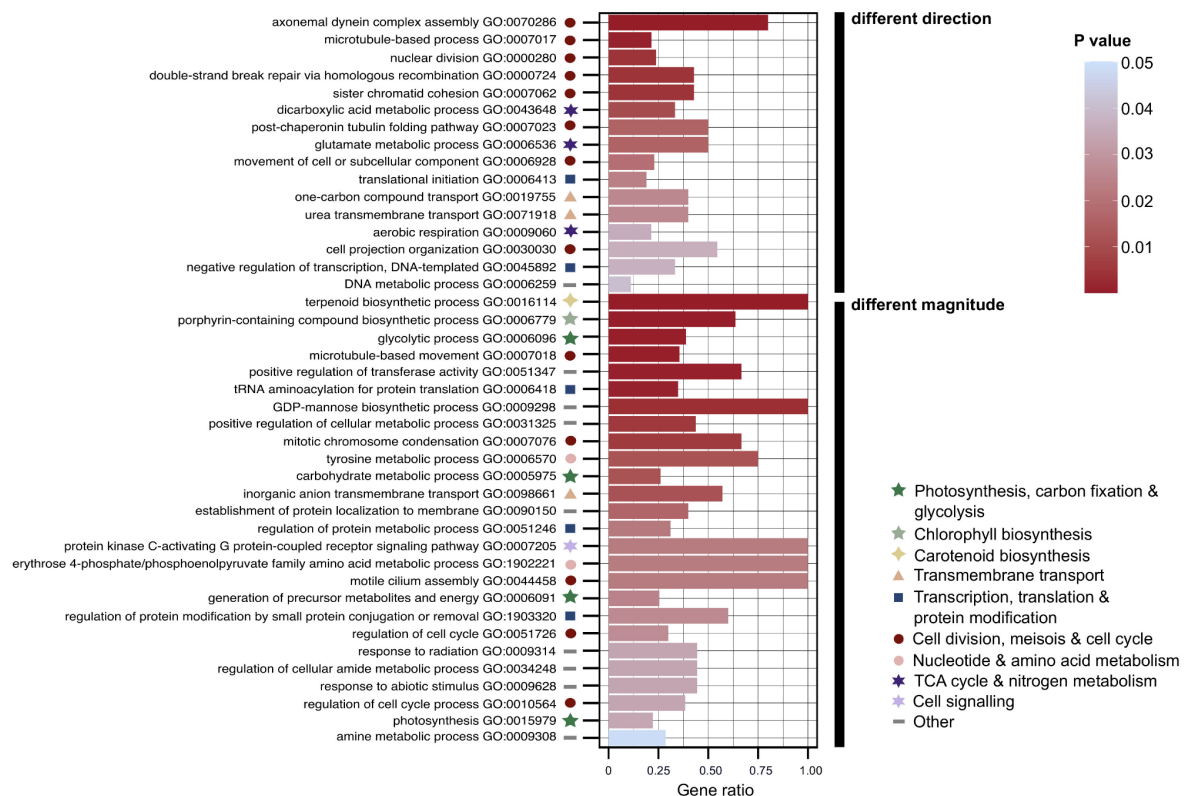
Fig. 7. GO enrichment of the interaction-effects: Biological Process. The barplot visualizes the significant GO terms retrieved by ORA (topGO, Fisher's exact test, *elim* algorithm) after removal of redundant GO terms by REVIGO. Two sets of GO enrichment were carried out which distinguished between genes that differ significantly between genotypes in the direction or magnitude of their response to low salinities. Barplot height indicates the proportion of genes that are DE with a given GO-term to the total number of genes with this GO-term in the genome of *S. marinoi*. The barplots are colored, and the GO terms ranked, according to P-value. Symbols indicate major categories of cellular processes to which a GO-term belongs. Only Biological Process GO-terms are shown.