Environmental fluctuations accelerate molecular evolution of thermal tolerance in a marine diatom

C.-Elisa Schaum[1,3*], A. Buckling[1], N. Smirnoff[2], D. J. Studholme[2], & G. Yvon-Durocher[1 *]

[1] Environment and Sustainability Institute, University of Exeter, Penryn Campus, Penryn, Cornwall TR10 9EZ, UK,

[2] Biosciences, College of Life and Environmental Sciences, Geoffrey Pope Building University of Exeter, Exeter, EX4 4QD, UK

[3] Institute for Hydrobiology and Fisheries, Section Oceanography, Hamburg University, 22767 Hamburg, Germany

*Correspondence to elisa.schaum@uni-hamburg.de or g.yvon-durocher@exeter.ac.uk

Words in summary: 150

Words in main text: 3878

Number of Figures (main text): 4

Number of Tables (main text): 2

Diatoms contribute roughly 20% of global primary production, but the factors determining their ability to adapt to global warming are unknown. Here we quantify the capacity for adaptation to warming in the marine diatom *Thalassiosira pseudonana*. We found that evolutionary rescue under severe warming was slow but adaptation to more realistic scenarios, where temperature increases were moderate or where they fluctuated between benign and severe conditions, was rapid. Adaption to warming was linked to major phenotypic changes in metabolism and elemental composition. Whole genome re-sequencing identified significant genetic divergence both among populations adapted to the different warming regimes and between the evolved and ancestral lineages. Consistent with the phenotypic changes, the most rapidly evolving genes were associated with transcriptional regulation, cellular response to oxidative stress and redox homeostasis. These results demonstrate that evolution of thermal tolerance in marine diatoms can be rapid, particularly in fluctuating environments, and is underpinned by major genomic and phenotypic divergence.

Introduction

Earth system models predict that global warming will result in significant declines in net primary production by marine phytoplankton throughout the 21st century (up to 20%)[1,2] driven by rising temperatures exceeding limits of thermal tolerance and increases in grazing and nutrient limitation in warmer, more stratified oceans [3-5]. Current models however, do not consider the potential for marine phytoplankton to rapidly adapt to environmental changes associated with global warming [6-8]. Such shortcomings have unknown consequences for projected changes in

41 global ocean primary production and arise because the mechanisms that facilitate or constrain the
42 capacity for rapid adaptation to warming in marine phytoplankton are entirely unknown.

43 Here, we address this fundamental knowledge gap by carrying out a 300-generation
44 selection experiment, with the model marine diatom *Thalassiosira pseudonana*, to assess the
45 potential for, and mechanisms that might facilitate rapid adaptation to warming in this globally
46 important phytoplankton [9-11]. A key hypothesis we test is that rapidly fluctuating temperatures –
47 an intrinsic feature of natural environments – will play a key role in adaptation [12-14]. Temporary
48 exposure to a benign environment resulting from temperature fluctuations could both accelerate
49 adaptation to severe conditions by increasing population size (a positive demographic effect) or
50 constrain adaptation, by relaxing selection for beneficial mutations that promote persistence in
51 the harsh environment (a negative population genetic effect)[12,13]. Our experiment was initiated
52 with a single clone (the ancestor), which had an upper thermal limit of 35ºC. The ancestor was
53 then distributed among 4 experimental treatments that represent a range of warming scenarios
54 (Fig. S1 for experimental set-up): (i) a control at 22ºC, which was the long-term culture
55 temperature; (ii) moderate warming at 26ºC; (iii) severe warming at 32ºC and (iv) a fluctuating
56 thermal regime, which cycled between 22 and 32ºC every 3-4 generations. The design of the
57 fluctuating treatment did not set out to mimic temperature variation in the natural environment,
58 but rather serves as a means to contrast evolutionary responses to warming under stable versus
59 fluctuating conditions, where temperatures vary predictably within a timescale that acclimation
60 can adjust the phenotype to cope with the change in environment [15]. To understand the
61 mechanisms that set the limits of thermal tolerance in marine diatoms and quantify the capacity
62 for adaptation beyond present limits, we re-sequenced the genomes, measured growth rate,
63 photosynthesis, respiration and elemental composition in the ancestor and evolved lineages after

64 approximately 300 generations of selection.

65 **Results & Discussion**

66 **Evolution of thermal tolerance**

67 Trajectories of population growth rate (μ, d^{-1}) over the course of the selection experiment
68 differed substantially between the selection regimes (Fig. 1; Tables S1, S2, S3). In the control
69 lineages, growth rates increased gradually, presumably reflecting continual laboratory
70 adaptation. However, trajectories of growth rate in the warming treatments were markedly
71 different from the control and one another. Lineages selected under severe warming (32ºC)
72 exhibited a characteristic pattern of evolutionary rescue (i.e. where evolution can reverse
73 population decline owing to environmental stress, and prevent otherwise inevitable extinction)
74 [14,16-18]. After an initial increase in the first 3 weeks of the experiment, growth rates (and
75 population densities, see Fig. S2) declined in the 32ºC environment and remained very low (0.24
76 ± 0.09) for more than one year (~ 100 generations). After approximately 100 generations, growth
77 rates rapidly increased and were statistically indistinguishable from those in the control
78 environment after 300 generations (0.63 ± 0.05 s.e.m. at 32ºC, 0.77± 0.11 s.e.m. at 22ºC). Under
79 moderate warming (26ºC), and in the regime that fluctuated between 22 and 32ºC growth rate
80 showed an immediate and sustained increase (2.1 and 1.9-fold faster than the ancestor
81 respectively). These results yield two important insights that are pertinent for understanding the
82 evolutionary dynamics of *T. pseudonana* in response to warming. First, adaptation to severe
83 warming was slow, with evolutionary rescue taking over a year to restore growth rates to levels
84 comparable with the ancestor in the control environment. Second, in the fluctuating environment
85 where populations experienced short bursts of exposure to 32ºC followed by periods in the

86 benign (22ºC) environment, adaptation to the severe environment was rapid. Consistent with our
87 hypothesis, lineages selected under the fluctuating regime maintained substantially larger
88 population sizes relative to those experiencing severe warming (Fig. S2 and Table S1, Table S3),
89 suggesting that temporary restoration of benign conditions increased the probability of fixing
90 beneficial mutations required for adaptation to the severe (32ºC) environment *via* a positive
91 demographic effect.

92 To assess whether adaptation to the various selection regimes changed the thermal
93 tolerance curves in the evolved lineages, we quantified growth rates across a temperature
94 gradient spanning 15ºC to 40ºC. Populations selected under the moderate (26ºC), severe (32ºC)
95 and fluctuating (22ºC/32ºC) warming treatments had higher optimal growth temperatures, T_{opt},
96 compared to the ancestor and the control (Fig. 1B; Tables S1, S4, and S5). Despite having a high
97 T_{opt}, the lineages selected under constant severe warming had the slowest growth rates at all
98 measurement temperatures except for 40ºC, the most extreme temperature, suggesting that these
99 populations were still under high-temperature stress in spite of the apparent evolutionary rescue
100 event and the passage of 300 generations of selection. In the fluctuating treatments, growth rate
101 at high temperatures was traded-off against slow growth at low temperature, suggesting that the
102 severe environment was the dominant driver of selection. By contrast, high temperature tolerance
103 in the moderate warming treatments appeared to incur no cost in performance at low
104 temperature. A key question therefore is: what mechanisms facilitated the rapid evolution of
105 increased thermal tolerance in the lineages selected under moderate and fluctuating warming?

106

107 **Increased metabolic efficiency facilitates evolution of thermal tolerance**

108 The fraction of photosynthesis that can be allocated to growth tends to decline with warming,
109 owing to the high temperature sensitivity of respiration relative to photosynthesis, suggesting
110 that the upper thermal tolerances of phytoplankton reflect metabolic constraints that limit the
111 efficiency of allocating metabolic energy to growth at high temperature[19]. To investigate whether
112 changes in metabolic traits could help explain the thermal tolerance curve of the ancestor and
113 evolutionary shifts observed over the selection experiment, we measured responses of gross
114 photosynthesis (*P*) and respiration (*R*) to acute gradients in temperature, spanning 4 to 45ºC, in
115 the ancestor and all evolved lineages. In contrast to previous work[19], the activation energies
116 characterising the temperature sensitivities of *P* and *R* in the ancestor were equivalent (e.g. E_a for
117 P = 1.08eV ± .0.18 s.e.m., and E_a for R = 1.07 ± 0.17 s.e.m.), i.e. increases in *P* and *R* up to the
118 optima were similar (Fig. 2; Tables S6, S7, S8). Optimum temperatures for *P* were however
119 lower than those for *R* (T_{opt} for *P*: 27 ºC ± 0.7ºC, (± s.e.m.); T_{opt} for R = 29 ºC ± 0.36ºC, (±
120 s.e.m.)), meaning that above T_{opt} for *P*, the carbon-use efficiency (CUE = 1 – *R*/*P*; i.e., the
121 *potential* carbon for allocation to growth) declined rapidly. The optimum temperature for growth
122 rate (T_{opt} = 28ºC; Fig. 1B) and subsequent decreases in growth at supra-optimal temperatures
123 coincided with declines in the CUE above T_{opt} for *P* in the ancestor, demonstrating that the
124 temperature dependence of CUE imposes a physiological constraint that shapes the thermal
125 tolerance curve. Thus, we hypothesize that shifts in traits that increase CUE at high temperatures
126 should have played an important role in facilitating adaptation to warming in the evolved
127 lineages of *T. pseudonana* (see also Table S9).

128 Following 300 generations of selection, we observed substantial shifts in the temperature
129 responses of *P* and *R*, both among treatments, as well as between the evolved lineages and the
130 ancestor (Fig. 2; Tables S6 –S9). For both *P* and *R*, estimates of T_{opt} in the warming treatments

131 were higher than the ancestor and control, and rates of P and R declined less abruptly at
132 temperatures exceeding T_{opt}. Consequently, CUEs remained high at hotter temperatures in the
133 warming treatments. The mass-specific rates of P and R at a reference temperature T_c = 18ºC (i.e.
134 $P(T_c)$ and $R(T_c)$, in Eq.6, which quantify the metabolic capacity per unit mass at a common
135 temperature) were significantly down-regulated under moderate, severe and fluctuating warming,
136 relative to the control treatment and the ancestor (Fig. 2 C, D, Tables S7 –S9). Rates of $R(T_c)$
137 also decreased more than those of $P(T_c)$, resulting in an overall increase in CUE in the lineages
138 selected under moderate and fluctuating warming (Fig. 2E). These results support our hypothesis
139 that fluctuating temperatures facilitate thermal adaptation and show that changes in the traits that
140 increase the CUE, played an important role in the evolution of elevated thermal tolerance in the
141 treatments subjected to warming. It is notable however, that the populations experiencing severe
142 warming had lower growth rates and lower CUEs than the lineages in the moderate and
143 fluctuating warming treatments (Fig. 1B, Fig. 2E, Table S1 and S9, also Fig. S5). This suggests
144 that while evolutionary rescue restored growth rates to levels comparable with the control, 300
145 generations of selection under constant, severe warming was insufficient to evolve CUEs that
146 facilitate growth rates comparable to populations that experienced only brief periods of severe
147 conditions.

148

149 **Contrasting effects of acclimation and evolution on cellular traits and elemental**
150 **composition**

151 Changes in temperature are known to alter optimal subcellular allocation of resources to various
152 macromolecular classes through acclimation (e.g. rapid changes in physiological traits not

153 attributable to genetic change)[20-26]. Indeed, algae exposed to warming often increase their
154 nitrogen-to-phosphorus ratios (N:P) by decreasing the density of P-rich ribosomes, relative to N-
155 rich proteins, owing to the increased efficiency of protein synthesis by ribosomes at higher
156 temperatures[23-26]. It is, however, unclear how these short-term changes in resource allocation are
157 linked to the shape of the thermal tolerance curve (e.g. the optimum or upper thermal limit). To
158 assess whether changes in subcellular allocation to various macromolecules could shed light on
159 the processes that set the limits of thermal tolerance in the ancestor, we quantified how key traits,
160 such as cell size, elemental composition and photophysiology, changed through acclimation
161 across the temperature gradient used to characterise the thermal tolerance curve of the ancestor
162 (Fig 1B). Consistent with expectations from the temperature-size rule[27], cell size declined
163 linearly with increasing temperature (Fig. 3A). In contrast, carbon (C), nitrogen (N), and
164 phosphorus (P), expressed per unit cell volume to account for changes in cell size, all increased
165 with temperature (Fig. 3 B-D). In line with the responses of N and P, protein and RNA content
166 also increased with temperature (Fig. 3 E, F). The N:P and C:P ratios increased with rising
167 temperature, but contrary to recent hypotheses[23-26], this was not driven by declines in P content,
168 but rather, by more pronounced increases in N and C relative to P (Fig. 3 H-J). Notably, at
169 temperatures exceeding the optimum for growth (T_{opt} = 28ºC), C, N, and P content, as well as
170 protein and RNA content increased markedly. These findings are consistent with a greater
171 demand for the transcription and translation of chaperones and other proteins required for the
172 repair and maintenance of heat-induced damage. In further support for this hypothesis, the
173 chlorophyll-to-carbon ratio showed a unimodal response to temperature, peaking at ca. 26ºC and
174 declining rapidly thereafter (Fig. 3G). Consequently, the increased N content observed at high
175 temperatures cannot be explained by greater allocation to photosynthetic machinery. Instead,

176 investment in proteins required for repair and maintenance provides a more parsimonious
177 explanation for the high N content in cells exposed to supra-optimal temperatures.

178 To assess whether evolutionary adjustments in cell size and elemental composition could
179 help to explain the evolution of elevated thermal tolerance, we quantified how these traits
180 changed in the evolved lineages after 300 generations in their respective selection regimes. Cell
181 size, C, N, and P content as well as protein and RNA content showed contrasting evolutionary
182 responses to temperature changes relative to the short-term, acclimation responses. In the
183 evolved lineages, cell size increased by a factor of almost 2 from 22ºC (control) to 32ºC (severe
184 warming) (Fig. 3A). This response contrasts with the linear decline in cell size observed as an
185 acclimation response to short-term increases in temperature in the ancestor. Likewise, C, N, and
186 P content as well as protein and RNA content, all declined between 22ºC and 32ºC in the evolved
187 lineages (Fig. 3 B-F). Consistent with the short-term temperature responses in the ancestor, the
188 N:P and C:P ratios both increased between 22ºC and 32ºC in the evolved lineages. However, in
189 spite of the similarity in the overall temperature response, the changes in N:P and C:P in the
190 evolved lineages, were driven by greater declines in P relative to N or C, consistent with
191 expectations that fewer ribosomes are required at higher temperature owing to their greater per
192 capita catalytic capacity[23-26] (Fig 3 H-J). This pattern contrasts directly with the short-term,
193 acclimation response in which increases in N:P and C:P were driven by greater allocation to N
194 and C with rising temperatures. Thus, while the increases in N and P with rising temperatures are
195 consistent with greater allocation to synthesising new proteins to help combat increased costs of
196 cellular maintenance in the short-term, declines in N and P content following evolutionary
197 adaptation may be indicative of re-allocation of metabolic energy back towards growth and
198 biomass synthesis over the long term. Consistent with this, the lineages that were selected under

199 constant, severe warming (i.e. 32ºC) evolved substantially higher chlorophyll-to-carbon ratios
200 relative to the short-term response of the ancestor when exposed to 32ºC, indicating that the
201 evolved populations were re-allocating energy away from repair/maintenance and towards light
202 harvesting machinery required for growth (Fig. 3G). In general, the opposing effects of
203 acclimation versus evolution on the direction of phenotypic trait responses to temperature change
204 are consistent with geographic patterns of counter gradient variation in which genetic influences
205 on a trait oppose environmental influences thereby minimising phenotypic change along the
206 gradient[28].

207

208 **Molecular evolution of thermal tolerance**

209 To investigate whether the observed changes in fitness and phenotypes in the evolved lineages
210 were also reflected by consistent patterns of molecular evolution, we re-sequenced the genomes
211 of the ancestor and the mixed (non-clonal) populations of the evolved lineages after 300
212 generations of selection. We identified candidate single-nucleotide variants (SNVs) in the
213 evolved populations compared to the ancestral clone that might have contributed to adaptation to
214 the warming treatments. Specifically, two types of genetic adaptation were of interest: (i)
215 mutations arising *de novo* after the t_0 time-point and becoming fixed in an evolved population
216 and (ii) alleles present in the ancestral clone, that become fixed in an evolved population.
217 Fixation of *de novo* mutations would result in an allele proportion rising from zero in the
218 ancestor to close to 1 after 300 generations of selection. Because the ancestor was a genetically
219 homogeneous clone, standing variation consists solely as heterozygosity; we would therefore
220 expect to observe alleles increasing in proportion from 0.5 in the ancestor to close to 1 after 300

221 generations. Therefore, we took two distinct approaches to surveying genetic variation: (i) we
222 identified SNVs that were undetectable in the ancestral population but were supported by 100%
223 of the aligned reads in an evolved population and (ii) we identified SNVs that plausibly had a
224 population proportion of 0.5 in the ancestral population (due to heterozygosity), and were
225 supported by 100% of the aligned reads in an evolved population.

226 We searched for non-silent SNVs affecting protein-coding genes that had become fixed
227 in at least one of the evolved populations; such variants might have contributed to the
228 populations' adaptation to the selection regime. SNVs were detected at 225,351 sites in the
229 genome. Of these, 7383 had reached fixation in at least one evolved population, including 1689
230 missense variants and 14 gain-of-stop-codon variants. In total, 1215 protein-coding genes were
231 affected by these fixed non-silent variants. Most fixed non-silent variants were present in the
232 ancestor at proportions between zero and one and were therefore probably heterozygous in the
233 ancestral clone (Fig. 4C), making the main mode of molecular evolution in these populations a
234 transition from heterozygosity to homozygosity. Only nine of the fixed variants were
235 undetectable in the ancestral clone, and are thus likely to be *de novo* mutations. None of these
236 nine *de novo* mutations were predicted to affect amino acid sequences of encoded proteins.

237 Using the SNVs that had reached fixation in at least one evolved population, we
238 quantified the genetic distance of each population from the ancestor as well as the genetic
239 divergence among populations using Analysis of Molecular Variance (AMOVA) based on
240 squared Euclidean distances and visualized this analysis using Principal Components Analysis
241 (PCA). Consistent with the fitness trajectories (Fig. 1A), all evolved populations showed
242 significant divergence from the ancestor and among one another (Fig. 4A; Table 1). The lineages

243 selected under moderate and fluctuating warming had the greatest distance from the ancestor on
244 the two major axes of (Fig. 4A; Fig. S6). Furthermore, the populations that experienced
245 fluctuations between severe and benign conditions were divergent from those experiencing
246 constant severe warming (Fig. 4A; Fig. S6), with significant variation in SNVs between
247 treatments (Table 1) . These results are consistent with the demographic data (Fig. 1A; Fig. S2),
248 indicating that the temporary restoration of benign conditions in the fluctuating regime may have
249 accelerated rates of molecular evolution by releasing cryptic genetic variation beneficial in the
250 stressful environment. Selection lines in the moderate warming treatment and the fluctuating
251 environment exhibited the greatest genetic divergence among replicate populations (Fig. 4A,
252 Table 1 - indicated by a larger σ^2_{AP}). Indeed, the genetic divergence among replicate populations
253 was significantly higher in the moderate warming treatment compared to those exposed to stable,
254 but severe warming (Fig. 4A, Table 1). The more conserved response in the severe environment
255 may have plausibly arisen from stronger directional selection owing to the exposure to extreme
256 conditions. The high genetic divergence among replicate populations in the fluctuating
257 environment is consistent with expectations that more complex environments lead to greater
258 diversity[29]. The above analyses were restricted to the subset of SNVs that had reached fixation in
259 at least one of the evolved populations and for which we therefore we had a high degree of
260 confidence that variants were either absent or heterozygous in the ancestor and rose to fixation in
261 the evolved lineage. When considering all SNVs, regardless of their level of fixation, (see
262 methods), the general patterns of genetic distance and divergence are qualitatively conserved
263 (see Fig. S7), indicating that the results are robust to any sequencing errors in estimating variants
264 at low frequencies.

265 Comparing principal component analyses (PCA) based on combining all the phenotypic
266 data, with the PCA from the non-synonymous SNVs, revealed striking similarities in the patterns
267 of divergence among treatments (Fig. 4A, B, Fig. S6, Fig. S8). These results suggest that the
268 observed changes in metabolic traits and elemental composition that were linked to the evolution
269 of thermal tolerance could also have a basis in some of the underlying genomic changes. To
270 investigate this, we identified the top 100 SNVs most strongly associated with the first and
271 second principal components as these are expected to be the most important SNVs for
272 characterising genomic variation among the lineages. We then identified the genes (see Table
273 S13) in which these SNVs fell and tested whether this gene set was enriched in any Gene
274 Ontology (GO)[30,31] terms related to biological processes, using the PCA loading score to weight
275 the relative importance of each gene. Genes that had gone to fixation in the evolved lineages and
276 were most strongly associated with the major axes of variation in the PCA were enriched in GO
277 terms related to transcriptional regulation, cellular response to oxidative stress and redox
278 homeostasis among others (Table 2). Notably, the genes that correspond to these enriched GO
279 terms were all most strongly associated with the moderate and fluctuating warming treatments in
280 the PCA.

281 We also found that several populations from the moderate and fluctuating warming
282 treatments displayed fixation of exceptionally high numbers of SNVs (Fig S12, S13). In these
283 lineages, the ancestral population's heterozygosity was lost at hundreds of single-nucleotide sites
284 in the genome. The mechanism(s) of heterozygosity loss among these populations is not known;
285 however, based on analysis of sequencing depth, we also observed substantial copy-number
286 variation (CNV) among populations across several large genomic regions (Figs. S9 - S11). These
287 findings are consistent with similar analyses of the *Chlamydomonas reinhardtii* genome in

288 laboratory and wild-type isolates[32]. This suggests that large-scale duplications and deletions
289 might be important mechanism for rapid evolution in algae, for example, when recurrent copy
290 number gains represent adaptive gain-of-function mutations[33] , dosage effects[34] and deletions
291 harbour potentially non-sense mutations, such as premature stop codons.

292 **Conclusions**

293 Our results shed new light on the adaptive potential and evolutionary dynamics of one of the
294 most abundant and widely distributed eukaryotic marine phytoplankton in response to warming.
295 We found that evolutionary rescue under severe warming was slow (> 1 year), but adaptation to
296 more realistic warming scenarios, where temperature increases were moderate or where they
297 fluctuated between benign and severe conditions, was rapid. The fluctuating environment
298 accelerated adaptation to severe warming because temporary restoration of benign conditions
299 increased population size and therefore the probability of fixing beneficial mutations required for
300 adaptation to the severe environment. Consistent with this demographic effect, the lineages
301 selected under fluctuating warming had the greatest genetic distance from the ancestor after 300
302 generations, indicating accelerated rates of molecular evolution. Because our experiments were
303 initiated with a single clone, they may in fact be conservative estimates of the evolutionary
304 potential of the highly diverse meta-population of *T. pseudonana* in the wild [5,35] where adaptation
305 will also be aided by standing genetic variation. Furthermore, our experiments were conducted
306 on a clone of the CCMP 1335 strain of *T. pseudonana*, which has been in laboratory culture for
307 many decades. Consequently, this strain has likely lost some of its genomic diversity owing to
308 long-term adaptation to simplified laboratory conditions, further increasing the likelihood that
309 observed evolutionary responses observed in our experiments are conservative estimates. Similar

experiments on freshly isolated strains and diverse phytoplankton taxa are urgently needed to provide improved estimates of the evolutionary potential of marine phytoplankton to adapt to global change. Our results also show how rapid, adaptive shifts in thermal tolerance are linked to major phenotypic changes in metabolism and elemental composition as well as significant molecular evolution, particularly in the populations evolved under moderate and fluctuating warming. Because the changes in carbon-use efficiency in the warm-adapted lineages arise due to differences the temperature sensitivities of photosynthesis and respiration, which are highly conserved metabolic pathways, our work raises the prospect that at the physiological-level, the mechanisms underpinning rapid thermal adaptation could be universal across the broad diversity of phytoplankton. Our findings could therefore help integrate rapid evolution into models of ocean biogeochemistry and improve projections of marine primary productivity over the 21st century[1,36].

References (1-36 main text, 37-65 methods)

1. Bopp, L. *et al.* Potential impact of climate change on marine export production. *Global Biogeochem. Cycles* **15,** 81–99 (2001).
2. Steinacher, M. *et al.* Projected 21st century decrease in marine productivity: a multi-model analysis. *Biogeosciences* **7,** 979–1005 (2010).
3. López-Urrutia, Á., Martin, E. S., Harris, R. P. & Irigoien, X. Scaling the metabolic balance of the oceans. *Proceedings of the National Academy of Sciences* **103,** 8739–8744 (2006).
4. Laufkötter, C. *et al.* Drivers and uncertainties of future global marine primary production in marine ecosystem models. *Biogeosciences* **12,** 6955–6984 (2015).
5. Regaudie-de-Gioux, A. & Duarte, C. M. Compensation irradiance for planktonic community metabolism in the ocean. *Global Biogeochem. Cycles* **24,** n/a–n/a (2010).
6. Reusch, T. B. H. & Boyd, P. W. Experimental evolution meets marine phytoplankton. *Evolution* **67,** 1849–1859 (2013).
7. Thomas, M. K., Kremer, C. T., Klausmeier, C. A. & Litchman, E. A Global Pattern

of Thermal Adaptation in Marine Phytoplankton. *Science* **338,** 1085–1088 (2012).
8. Schluter, L. *et al.* Adaptation of a globally important coccolithophore to ocean warming and acidification. *Nature Climate Change* **4,** 1024–1030 (2014).
9. Field, C. B. Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. *Science* **281,** 237–240 (1998).
10. Falkowski, P. G. Biogeochemical Controls and Feedbacks on Ocean Primary Production. *Science* **281,** 200–206 (1998).
11. Armbrust, E. V. The Genome of the Diatom Thalassiosira Pseudonana: Ecology, Evolution, and Metabolism. *Science* **306,** 79–86 (2004).
12. Hao, Y.-Q., Brockhurst, M. A., Petchey, O. L. & Zhang, Q.-G. Evolutionary rescue can be impeded by temporary environmental amelioration. *Ecol Lett* **18,** 892–898 (2015).
13. Ashander, J. & Chevin, L. M. Predicting evolutionary rescue via evolving plasticity in stochastic environments. *Proceedings of the Royal Society B: Biological Sciences* **283,** 20161690 (2016).
14. Bell, G. Evolutionary rescue and the limits of adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences* **368,** 20120080–20120080 (2012).
15. Botero, C. A., Weissing, F. J., Wright, J. & Rubenstein, D. R. Evolutionary tipping points in the capacity to adapt to environmental change. *Proceedings of the National Academy of Sciences* **112,** 184–189 (2015).
16. Lachapelle, J. & Bell, G. Evolutionary rescue of sexual and asexual populations in a deteriorating environment. *Evolution* **66,** 3508–3518 (2012).
17. Elena, S. F. & Lenski, R. E. Test of synergistic interactions among deleterious mutations in bacteria. *Nature* **390,** 395–398 (1997).
18. Bell, G. & Gonzalez, A. Adaptation and Evolutionary Rescue in Metapopulations Experiencing Environmental Deterioration. *Science* **332,** 1327–1330 (2011).
19. Padfield, D., Yvon-Durocher, G., Buckling, A., Jennings, S. & Yvon-Durocher, G. Rapid evolution of metabolic traits explains thermal adaptation in phytoplankton. *Ecol Lett* **19,** 133–142 (2015).
20. J Geider, R., Osborne, B. A. & Raven, J. A. Light dependence of growth and photosynthesis in phaeodactylum tricornutum (bacillariophyceae)1. *Journal of Phycology* **21,** 609–619 (1985).
21. Geider, R. J., Maclntyre, H. L. & Kana, T. M. A dynamic regulatory model of phytoplanktonic acclimation to light, nutrients, and temperature. *Limnol. Oceanogr.* **43,** 679–694 (1998).
22. Raven, J. A. & Geider, R. J. Temperature and algal growth. *New Phytol* **110,** 441–461 (1988).
23. Toseland, A. *et al.* The impact of temperature on marine phytoplankton resource allocation and metabolism. *Nature Climate Change* **3,** 979–984 (2013).
24. Daines, S. J., Clark, J. R. & Lenton, T. M. Multiple environmental controls on phytoplankton growth strategies determine adaptive responses of the N : P ratio. *Ecol Lett* **17,** 414–425 (2014).
25. Yvon-Durocher, G., Dossena, M., Trimmer, M., Woodward, G. & Allen, A. P. Temperature and the biogeography of algal stoichiometry. *Global Ecology and Biogeography* **24,** 562–570 (2015).

26. Geider, R. & La Roche, J. Redfield revisited: variability of C:N:P in marine microalgae and its biochemical basis. *European Journal of Phycology* **37,** 1–17 (2002).
27. Atkinson, D., Ciotti, B. J. & Montagnes, D. J. S. Protists decrease in size linearly with temperature: ca. 2.5% C-1. *Proceedings of the Royal Society B: Biological Sciences* **270,** 2605–2611 (2003).
28. Conover, D. O. & Schultz, E. T. Phenotypic similarity and the evolutionary significance of countergradient variation. *Trends in Ecology & Evolution* **10,** 248–252 (1995).
29. Replansky, T. & Bell, G. The relationship between environmental complexity, species diversity and productivity in a natural reconstructed yeast community. *Oikos* **118,** 233–239 (2009).
30. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nature Genetics* **25,** 25 EP –
31. Consortium, T. G. O. Expansion of the Gene Ontology knowledgebase and resources. *Nucl. Acids Res.* **45,** D331–D338 (2016).
32. Flowers, J. M. *et al.* Whole-Genome Resequencing Reveals Extensive Natural Variation in the Model Green Alga Chlamydomonas reinhardtii. *Plant Cell* **27,** 2353–2369 (2015).
33. Gresham, D. & Dunham, M. J. The enduring utility of continuous culturing in experimental evolution. *Genomics* **104,** 399–405 (2014).
34. Kondrashov, F. A. Gene duplication as a mechanism of genomic adaptation to a changing environment. *Proceedings of the Royal Society B: Biological Sciences* **279,** 5048–5057 (2012).
35. Armbrust, E. V. The life of diatoms in the world's oceans. *Nature* **459,** 185–192 (2009).
36. Dutkiewicz, S., Scott, J. R. & Follows, M. J. Winners and losers: Ecological and biogeochemical changes in a warming ocean. *Global Biogeochem. Cycles* **27,** 463–477 (2013).
37. Excoffier, L., Smouse, P. E. & Quattro, J. M. Analysis of Molecular Variance Inferred From Metric Distances Among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data . *Genetics* **131,** 479–491 (1992).
38. Boyd, P. W. *et al.* Marine Phytoplankton Temperature versus Growth Responses from Polar to Tropical Waters – Outcome of a Scientific Community-Wide Study. *PLoS ONE* **8,** e63091–17 (2013).
39. Thomas, M. K. *et al.* Temperature-nutrient interactions exacerbate sensitivity to warming in phytoplankton. *Global Change Biol* **23,** 3269–3280 (2017).
40. O'Donnell, D. R., Hamman, C. R. & Johnson, E. C. Temperature selection drives evolution of function-valued traits in a marine diatom. *bioRxiv* (2017). doi:10.1101/167817
41. Guillard, R. R. L. in Culture of Marine Invertebrate Animals 29–60 (Springer US, 1975). doi:10.1007/978-1-4615-8714-9_3
42. Schaum, C.-E. *et al.* Adaptation of phytoplankton to a decade of experimental warming linked to increased photosynthesis. *Nature Publishing Group* **1,** 1–7 (2017).
43. Amin, S. A. *et al.* Interaction and signalling between a cosmopolitan phytoplankton

and associated bacteria. *Nature* **522,** 98–101 (2015).
44. Amin, S. A., Parker, M. S. & Armbrust, E. V. Interactions between Diatoms and Bacteria. *Microbiology and Molecular Biology Reviews* **76,** 667–684 (2012).
45. Brickley, M. R., Lawrie, E., Weise, V., Hawes, C. & Cobb, A. H. Use of a potentiometric vital dye to determine the effect of the herbicide bromoxynil octanoate on mitochondrial bioenenergetics in Chlamydomonas reinhardtii. *Pest. Manag. Sci.* **68,** 580–586 (2011).
46. Guzmán, H. M., la Jara Valido, de, A., Duarte, L. C. & Presmanes, K. F. Estimate by means of flow cytometry of variation in composition of fatty acids from Tetraselmis suecica in response to culture conditions. *Aquacult Int* **18,** 189–199 (2009).
47. Holm-Hansen, O. & Riemann, B. Chlorophyll a Determination: Improvements in Methodology. *Oikos* **30,** 438 (1978).
48. Murphy, J. & Riley, J. P. A modified single solution method for the determination of phosphate in natural waters. *Analytica Chimica Acta* **27,** 31–36 (1962).
49. Grasshoff, K., Kremling, K. & Ehrhardt, M. *Methods of seawater analysis*. (2009).
50. Suggett, D. J., Moore, C. M., Hickman, A. E. & Geider, R. J. Interpretation of fast repetition rate (FRR) fluorescence: signatures of phytoplankton community structure versus physiological state. *Mar. Ecol. Prog. Ser.* **376,** 1–19 (2009).
51. Falkowski, P. G., Dubinsky, Z. & Wyman, K. Growth-irradiance re.lationships in phytoplankton. *Limnol. Oceanogr.* **30,** 311–321 (1985).
52. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997v2 (2013).
53. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25,** 1754–1760 (2009).
54. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly* **6,** 80–92 (2012).
55. Clopper, C.J. and Pearson, E.S. The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial. Biometrika **26**: 404. (1934)
56. Tibshirani, R., Hastie, T., Narasimhan, B. & Chu, G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proceedings of the National Academy of Sciences* **99,** 6567–6572 (2002).
57. Lê Cao, K.-A. *et al.* MixMC: A Multivariate Statistical Framework to Gain Insight into Microbial Communities. *PLoS ONE* **11,** e0160169–21 (2016).
58. NCBI Resource Coordinators. Database Resources of the National Center for Biotechnology Information. *Nucl. Acids Res.* **45,** D12–D17 (2017).
59. Huson, D. H., Mitra, S., Ruscheweyh, H. J., Weber, N. & Schuster, S. C. Integrative analysis of environmental sequences using MEGAN4. *Genome Research* **21,** 1552–1560 (2011).
60. Schoolfield, R. M., Sharpe, P. J. H. & Magnuson, C. E. Non-linear regression of biological temperature-dependent rate models based on absolute reaction-rate theory. *Journal of Theoretical Biology* **88,** 719–731 (1981).
61. Sharpe, P. J. & DeMichele, D. W. Reaction kinetics of poikilotherm development. *Journal of Theoretical Biology* **64,** 649–670 (1977).
62. Gillooly, J. F. Effects of Size and Temperature on Metabolic Rate. *Science* **293,** 2248–2251 (2001).

63. Brown, J. H., Gillooly, J. F., Allen, A. P., Van M Savage & West, G. B. Toward a metabolic theory of ecology. *Ecology* **85,** 1771–1789 (2004).
64. Van M Savage, Gillooly, J. F., Brown, J. H., West, G. B. & Charnov, E. L. Effects of Body Size and Temperature on Population Growth. *The American Naturalist* **163,** 429–441 (2015).
65. Allen, A. P. Global Biodiversity, Biochemical Kinetics, and the Energetic-Equivalence Rule. *Science* **297,** 1545–1548 (2002).

Acknowledgments. This study was funded by a Leverhulme Trust research grant (RPG-2013-335). Whole genome re-sequencing was carried out at Exeter Sequencing Service and Computational core facilities at the University of Exeter, where Dr. Karen Moore, Dr. Audrey Farbos, Paul O'Neill and Dr. Konrad Paszkiewicz lead the handling of the samples. Exeter Squencing Services are supported by Medical Research Council Clinical Infrastructure award (MR/M008924/1), Wellcome Trust Institutional Strategic Support Fund (WT097835MF), Wellcome Trust Multi User Equipment Award (WT101650MA) and BBSRC LOLA award (BB/K003240/1).

Author contributions. GYD conceived and designed the study, analysed data and wrote the manuscript. ES designed and carried out the experimental work, analysed data and wrote the manuscript. AB conceived and designed the study. DS carried out the bioinformatic analyses of the genomic data. NS contributed to the interpretation of the data analysis. All authors contributed to writing the manuscript.

Author information. The authors declare no competing financial interests. Correspondence and requests for materials (code and data) should be directed to ES or GYD.

505 **Data availability statement:** All data used for analyses and graphical output is made available
506 as part of the supporting information. All sequence data have been deposited in the Sequence
507 Read Archive (Leionen et al., DOI 10.1093/nar/gkq1019) with accession number SRP114919
508 and are accessible via BioProject PRJNA397360. Code is available from the authors upon
509 reasonable request.
510

511 **Tables**

512 **Table 1 | Results of the analysis of molecular variance (AMOVA).** To assess molecular
513 differentiation within treatments (i.e. among biological replicates in one treatment) and between
514 treatments (i.e. between the different selection environments), an AMOVA was fitted to the
515 frequency matrixes of SNVs that had reached fixation. It revealed that a significant proportion of
516 the variance in SNV frequencies at specific loci was attributable to differences between
517 treatments. Within individual treatments, there was significant variation between biological
518 replicates ('populations' in the table), with the largest amount of variation among populations,
519 ('AP'), in the 26ºC and FS treatments. σ^2 represents the variance of each hierarchical level (i.e.
520 within population 'WP', among populations 'AP' and between treatments 'BT'), and the % gives
521 the percentage of this variance calculated to a total of 100 %. Significance testing was carried out
522 through random permutation of the samples following the methods outlined in Ref. ([37]).
523

22ºC	σ^2	%	p Value
Within population (σ^2_{WP})	39.76	74.99	<0.05
Among population (σ^2_{AP})	13.26	25.01	<0.05
Total (σ^2_T)	53.02	100	
26ºC			
Within population (σ^2_{WP})	31.07	52.29	<0.01
Among population (σ^2_{AP})	28.35	47.71	<0.01
Total (σ^2_T)	59.42	100	
32ºC			
Within population (σ^2_{WP})	43.66	86.51	<0.05
Among population (σ^2_{AP})	6.81	13.49	0.051
Total (σ^2_T)	50.47	100	
FS			
Within population (σ^2_{WP})	31.03	50.04	<0.01
Among population (σ^2_{AP})	30.98	49.96	<0.01
Total (σ^2_T)	62.01	100	
All treatments			
Within population (σ^2_{WP})	28.97	49.69	<0.05
Among populations between treatments ($\sigma^2_{AP/BT}$)	8.82	15.12	<0.05
Between treatments (σ^2_{BT})	**20.522**	**35.19**	**<0.05**
Total (σ^2_T)	58.32	100	

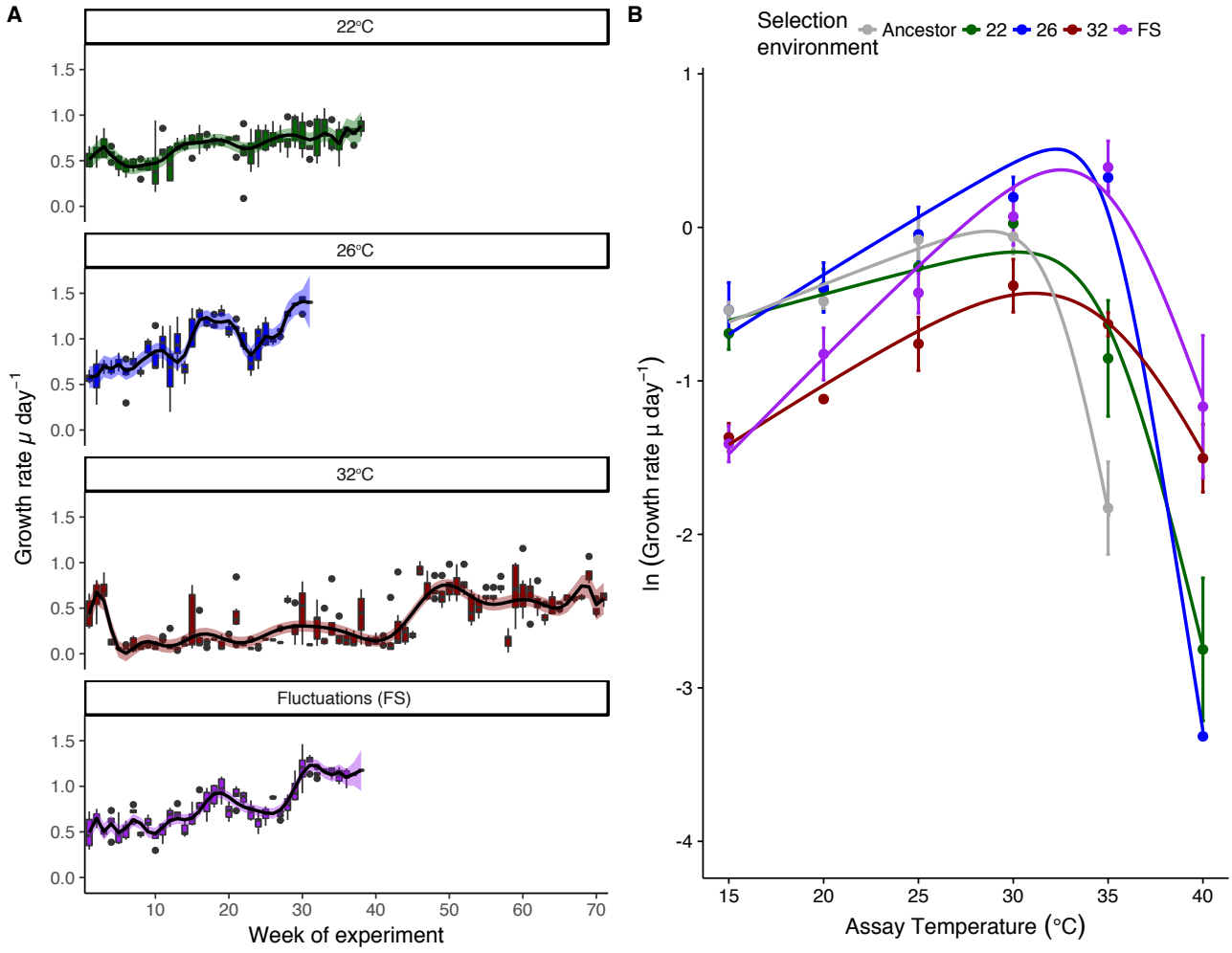
524

525 **Table 2 | Table of Gene Ontology (GO) terms relating to biological processes that were significantly enriched from the gene set**
526 **identified as being strongly associated with the major axes of variation in the principal components analysis.**
527 *GO.ID* is the Gene Ontology identifier retrieved from protists.ensembl.org, *term* is the biological process associated with the GO.ID,
528 and *Ann* is for the number of genes annotated to the GO.ID. *Sig* and *Exp* denote the number of significant and expected annotations for
529 the GO.ID category found in the dataset of genes that are associated with the top 100 PCA loadings for the two principal components
530 displayed in Fig 4A compared to the reference 'gene universe' made up from the entire *T. pseudonana* genome. *KS* is the p-value
531 output of a Kolmogorov-Smirnov test, which replaces Fisher's exact test when working with scores (see Methods) with $p < 0.05$
532 indicated by * for significant enrichment. Loading # indicates the position that the gene has in the PCA's top 100 loadings, and comp
533 indicates whether it is more strongly associated with the first or second principal component (see Figure 4A). The loading score is a
534 numerical value for the strength of the association with a component, where higher absolute values are indicative of a stronger
535 association. *Gene* gives the locus tag of the gene that was found to be strongly associated with the axis of variation. All genes are
536 coding for hypothetical proteins, and were assigned GOs through the most likely function that the protein may have given its amino
537 acid sequence. *Treat* denotes the treatment that SNVs in that gene were most likely to be associated with based on principal
538 component analysis.
539

GO.ID	Term	Ann	Sig	Exp	KS	Gene	Loading #	Comp	Treat	Loading score
GO:1902358	sulfate transmembrane transport	12	1	0.02	0.015 *	THAPS_23437 (hypothetical protein)	2	comp1	FS	0.0559
GO:0006357	regulation of transcription from RNA polymerase	44	1	0.03	0.015 *	THAPSDRAFT_7370 (hypothetical protein)	27	comp1	FS	0.0483
GO:0001522	pseudouridine synthesis	44	1	0.03	0.025 *	THAPS_23513 (hypothetical protein)	40	comp1	26	0.0472
GO:0034599	cellular response to oxidative stress	23	1	0.03	0.029 *	THAPSDRAFT_25121 (hypothetical protein)	22	comp2	FS	0.0570

GO:0008152	metabolic processes	77	1	0.03	0.035 *	THAPSDRAFT_23040 (hypothetical protein)	6	comp1	26	0.0533
GO:0045454	cell redox homeostasis	77	1	0.03	0.035 *	THAPSDRAFT_2086 (hypothetical protein)	31	comp1	26	0.0478
GO:0000103	sulfate assimilation	29	1	0.02	0.037 *	THAPSDRAFT_25121.2 (hypothetical protein)	22	comp2	FS	0.0570
GO:0016567	protein ubiquitination	3706	5	4.75	0.254	THAPS_23513,THAPSDRAFT_2086,THAPSDRAFT_7370 (hypothetical protein)	various	comp1 and comp2	FS and 26	various

543



544

545 **Figure 1| Growth rate trajectories and rapid evolutionary shifts in thermal tolerance**
546 **curves. (A)** Time-series of population growth rates reveal immediate and sustained increases in
547 growth rates in populations adapting to moderate warming and the environment that fluctuated
548 between 22 (benign conditions) and 32ºC (chronic conditions). In the 32ºC environment
549 populations performed very poorly for approximately 1 year (~ 100 generation), followed by an
550 evolutionary rescue event, which caused growth rates to increase to levels comparable with the
551 control treatment at 22ºC. Boxplots were created by binning growth rate estimates across
552 replicates on each week of the experiment until 300 generations had passed in each lineage.
553 Fitted lines are from the best fits of a GAMM (see Methods and Table S2). (**B**) Thermal response
554 curves of the ancestor and evolved lineages show marked shifts in thermal tolerance.
555 Populations selected under moderate, severe and fluctuating warming all evolved increased
556 tolerance of high temperatures relative to the ancestors and the control. Values are means and
557 error bars denote ± 1 s.e.m. Fitted lines are the fixed effects of a nonlinear mixed effects model.
558 Grey curves denote the ancestor, green, samples are the control at 22ºC, blue is 26ºC, red is
559 32ºC, and purple is the fluctuating environment.

560

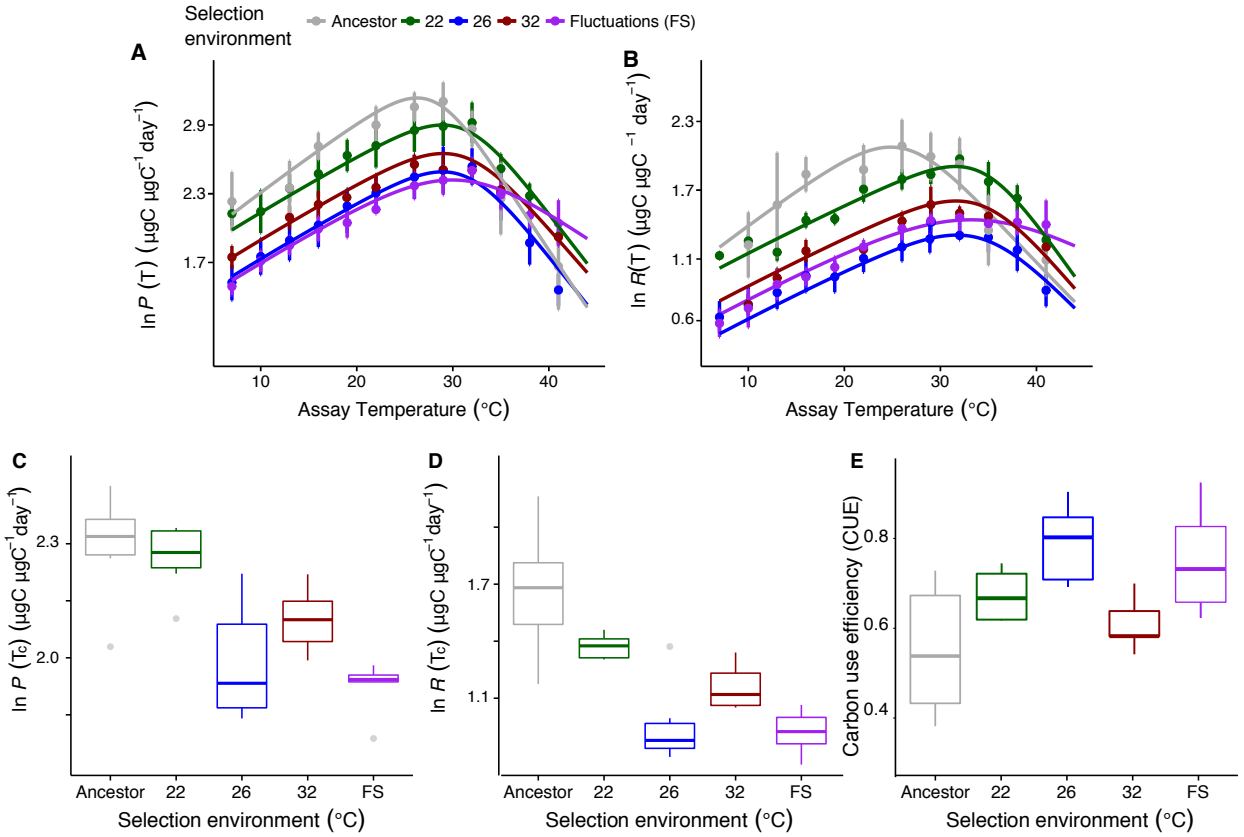
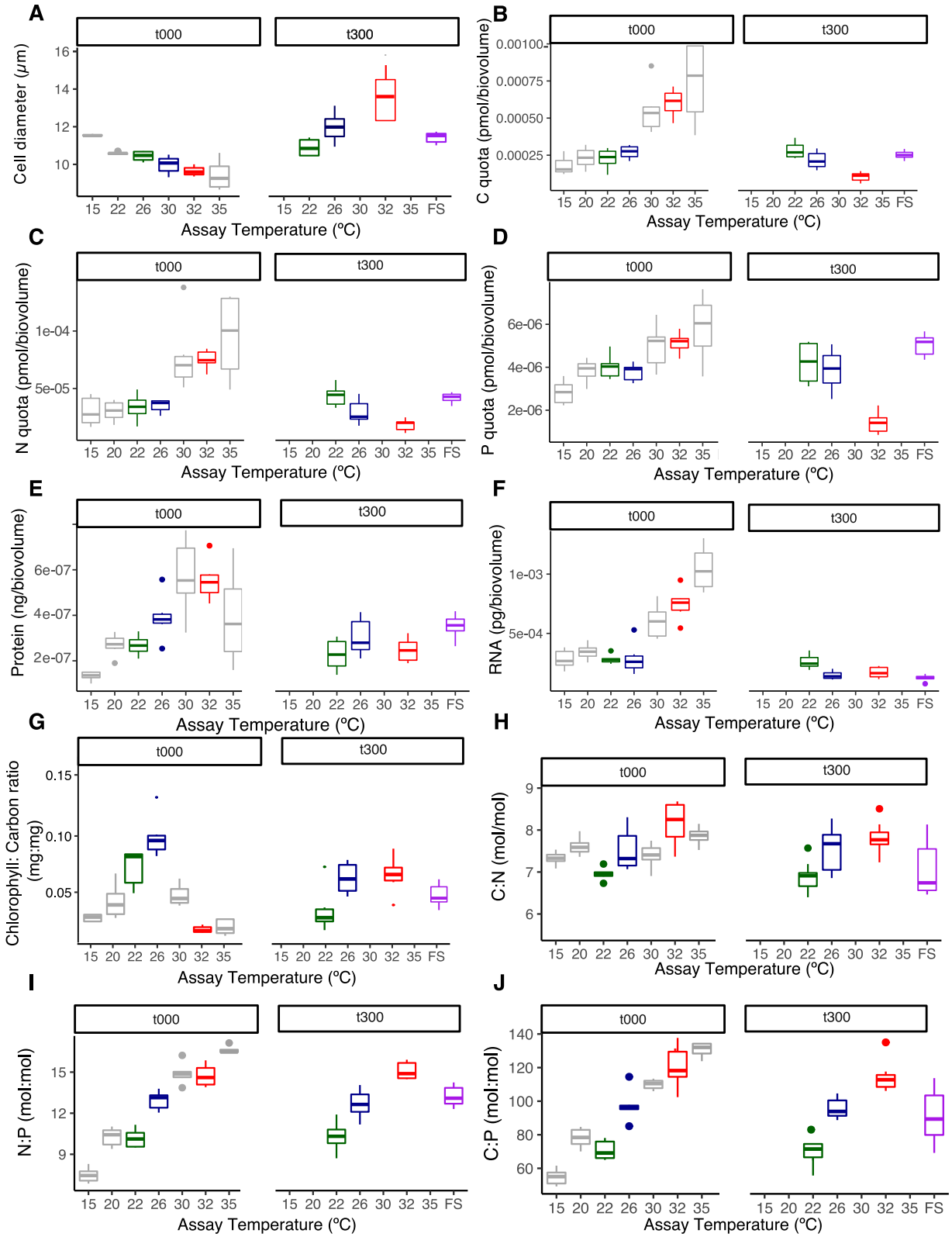
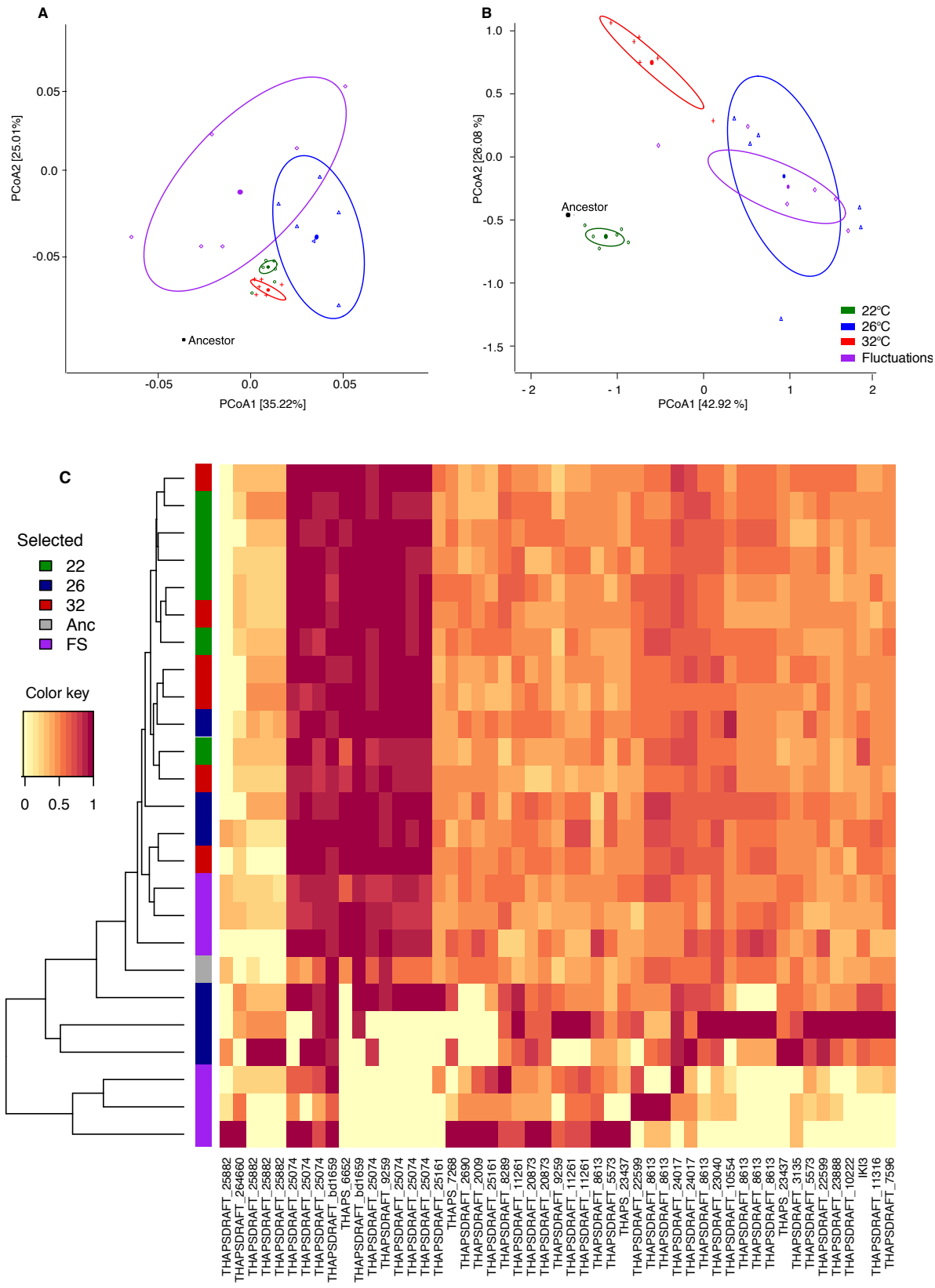


Figure 2| Rapid evolution of metabolic traits increase carbon-use efficiency. Thermal response curves for **(A)** gross photosynthesis rates (P), and (**B**) respiration (R) reveal substantial shifts in metabolic thermal traits among treatments, as well as between the evolved lineages and the ancestor. Data are presented as means with error bars denoting ± 1s.e.m, n = 6 per assay temperature, fitted lines are fixed effects from non-linear mixed effects models. Mass-specific rates of (**C**) gross photosynthesis and (**D**) respiration at a reference temperature ($P(T_c)$ and $R(T_c)$ respectively, see methods) were significantly down-regulated in the populations evolved under moderate, chronic and fluctuating warming. Furthermore, rates of $R(T_c)$ were down-regulated more than those of $P(T_c)$ in the moderate and fluctuating warming treatments meaning that **(E)** those lineages also had the highest carbon-use efficiency (CUE). Grey denotes the ancestor, green is the control at 22ºC, blue is 26ºC, red is 32ºC, and purple is the fluctuating environment. Carbon use efficiency for FS is displayed at 32ºC for easier comparison to the 32ºC-evolved samples, and did not differ significantly between assay temperatures of 22ºC, 26ºC, and 32ºC (see also Fig. S5).


A
t000
t300
Cell diameter (μm)
16
14
12
10
15 22 26 30 32 35
15 22 26 30 32 35 FS
Assay Temperature (°C)
B
t000
t300
C quota (pmol/biovolume)
0.00100
0.00075
0.00050
0.00025
15 20 22 26 30 32 35
15 20 22 26 30 32 35 FS
Assay Temperature (°C)
C
t000
t300
N quota (pmol/biovolume)
1e-04
5e-05
15 20 22 26 30 32 35
15 20 22 26 30 32 35 FS
Assay Temperature (°C)
D
t000
t300
P quota (pmol/biovolume)
6e-06
4e-06
2e-06
15 20 22 26 30 32 35
15 20 22 26 30 32 35 FS
Assay Temperature (°C)
E
t000
t300
Protein (ng/biovolume)
6e-07
4e-07
2e-07
15 20 22 26 30 32 35
15 20 22 26 30 32 35 FS
Assay Temperature (°C)
F
t000
t300
RNA (pg/biovolume)
1e-03
5e-04
15 20 22 26 30 32 35
15 20 22 26 30 32 35 FS
Assay Temperature (°C)
G
t000
t300
Chlorophyll: Carbon ratio (mg:mg)
0.15
0.10
0.05
15 20 22 26 30 32 35
15 20 22 26 30 32 35 FS
Assay Temperature (°C)
H
t000
t300
C:N (mol/mol)
9
8
7
6
15 20 22 26 30 32 35
15 20 22 26 30 32 35 FS
Assay Temperature (°C)
I
t000
t300
N:P (mol:mol)
15
12
9
15 20 22 26 30 32 35
15 20 22 26 30 32 35 FS
Assay Temperature (°C)
J
t000
t300
C:P (mol:mol)
140
120
100
80
60
15 20 22 26 30 32 35
15 20 22 26 30 32 35 FS
Assay Temperature (°C)

581
582 **Figure 3 | Evolutionary shifts in cell size, macromolecular composition, photosynthetic and**
583 **nutrient-use efficiency in the ancestor and evolved lineages.** (**A**) Cell size (here as µm
584 diameter) decreased with warming in the ancestor, but cells were larger at warmer temperatures
585 after 300 generations of evolution. Consequently, carbon **(B),** nitrogen **(C),** and phosphorus **(D)**
586 quotas per cell volume of ancestral samples remained relatively stable up to the ancestors
587 optimum temperature and then increased, whereas elemental quotas in evolved lineages
588 decreased with temperature (compare also Fig. S4 for per cell quotas). Protein **(E)** and (**F**) RNA
589 content per cell volume mirrored the changes in intracellular N and P content, respectively. The
590 chlorophyll-to-carbon ratio was also highest in the fastest growing lineages selected under
591 moderate warming (**G**) and the C:N **(H),** N:P **(I),** and C:P **(J)** ratios increased with increasing
592 selection temperature both in the short- and in the long-term, with values in the fluctuating
593 environment comparable to those in the moderate warming treatment See supporting information
594 for Φ_{PSII} (Fig. S3) and silicate content (Fig. S4). For all panels, $n = 6$. Boxplots are for the
595 ancestor (grey) and evolved lineages after 300 generations at 22 ºC (green), 26ºC (blue), 32ºC
596 (red) and the fluctuating environment (purple).


A
PCoA2 [25.01%]
PCoA1 [35.22%]
Ancestor
B
PCoA2 [26.08 %]
PCoA1 [42.92 %]
Ancestor
22°C
26°C
32°C
Fluctuations
C
Selected
22
26
32
Anc
FS
Color key
0
0.5
1
THAPSDRAFT_25882
THAPSDRAFT_264660
THAPSDRAFT_25882
THAPSDRAFT_25882
THAPSDRAFT_25882
THAPSDRAFT_25074
THAPSDRAFT_25074
THAPSDRAFT_25074
THAPSDRAFT_bd1659
THAPS_6652
THAPSDRAFT_bd1659
THAPSDRAFT_25074
THAPSDRAFT_9259
THAPSDRAFT_25074
THAPSDRAFT_25074
THAPSDRAFT_25074
THAPSDRAFT_25161
THAPS_7288
THAPSDRAFT_2690
THAPSDRAFT_2009
THAPSDRAFT_25161
THAPSDRAFT_8289
THAPSDRAFT_11261
THAPSDRAFT_20873
THAPSDRAFT_20873
THAPSDRAFT_9259
THAPSDRAFT_11261
THAPSDRAFT_11261
THAPSDRAFT_8613
THAPSDRAFT_5573
THAPS_23437
THAPSDRAFT_22599
THAPSDRAFT_8613
THAPSDRAFT_8613
THAPSDRAFT_24017
THAPSDRAFT_24017
THAPSDRAFT_8613
THAPSDRAFT_23040
THAPSDRAFT_10554
THAPSDRAFT_8613
THAPSDRAFT_8613
THAPSDRAFT_8613
THAPS_23437
THAPSDRAFT_3135
THAPSDRAFT_5573
THAPSDRAFT_22599
THAPSDRAFT_23888
THAPSDRAFT_10222
IKI3
THAPSDRAFT_11316
THAPSDRAFT_7596

599 **Figure 4 |Molecular and phenotypic evolution in *T. pseudonana*. (A)** Principal components
600 analysis (PCA) of evolved and ancestral lineages calculated from distance matrices based on
601 single nucleotide variants (SNVs) in protein-coding regions that had reached fixation after 300
602 generations of evolution. **(B)** PCA of phenotypic traits calculated from the change in each trait
603 measured at the treatment temperature relative to that expressed in the ancestor at the same
604 temperature. In panels A and B colours indicate the selection regime with black for the ancestor,
605 green for 22ºC, blue for 26ºC, red for 32ºC, and purple for the fluctuating regime. **(C)** Estimated
606 allele frequencies for 1703 single-nucleotide variants that putatively reached fixation in at least
607 one t300 population. The colour of each cell in the heat-map indicates estimated allele proportion
608 in the population, based on ratio of variant sequence reads versus total read depth at that genomic
609 site. Homozygously fixed alleles will appear as yellow or red (allele proportion zero or one,
610 respectively), heterozygously fixed alleles as orange. The side-bar indicates the selection regime
611 with colours as stated above.

612
613 **METHODS**

614 Experimental Design

615 In November 2014, a clone of the sequenced strain of *Thalassiosira pseudonana* (CCMP 1335)
616 [11] was obtained from the CCAP culture collection, where it had been re-isolated from the original
617 clone CCMP 1335 in 2009, and maintained in batch culture at 18-22ºC. This lineage is known to
618 exhibit a unimodal growth response to temperature between 16ºC and 31ºC, with optimal growth
619 temperatures around 27ºC [38-40]. The stock culture was made clonal by serial dilution in a 96-well
620 plate in March 2015 and a single clone was chosen at random to be the ancestor for all treatments
621 in the selection experiment. Determination of the thermal tolerance curve of the ancestor
622 revealed that growth rates peaked at 28ºC and that mortality was high above 35ºC (see section
623 below for details on characterizing the thermal tolerance curves). The ancestor was distributed
624 among 4 experimental treatments that represent a range of warming scenarios based on its limits
625 of thermal tolerance: (i) a control at 22ºC, which was the long-term culture temperature; (ii)
626 moderate warming at 26ºC; (iii) severe warming at 32ºC and (iv) a fluctuating thermal regime,
627 which cycled between 22 and 32ºC every 3-4 generations. Each treatment was replicated 6 times.

628 Lineages were grown in grown in f/2 Medium (Guillard's medium for diatoms[41]) with
629 artificial seawater, under a 12:12 light/dark cycle in INFORS HT shaking incubators. Salinity
630 was maintained at 32 (i.e. 32g NaCl L^{-1} in 39.5 g L^{-1} artificial seawater reagents) and pH
631 adjusted to 8.213 ± 0.291 (± s.e.m., averaged across all transfers) prior to each transfer. pH was
632 further measured at the end of each transfer (average across all transfers: 8.171 ± 0.252). The
633 light intensity of full-spectrum fluorescent tubes was 100 µmol quanta m^{-2} s^{-1}. As light intensity
634 varied throughout the incubators (minimum light intensity 70, maximum light intensity 110 µmol

635 quanta m^{-2} s^{-1}), culture vessels were moved randomly within the incubator every other day.
636 Lineages were maintained in semi-continuous batch culture, and transferred during the
637 exponential phase of growth to an inoculum of 100 cells mL^{-1}. Average cell densities were
638 calculated at each transfer from three technical replicates for each biological replicate.
639

640 Flow cytometry

641 Abundance and cell size (diameter assuming a spherical shape) were determined with a Accuri
642 C6 flow cytometer (BD Scientific). Cell size was detected using forward scatter (FSC) calibrated
643 using species and beads of known size against photographic images processed in ImageJ as
644 described in [42], with a conversion factor of FSC=108058 µm + 4.5665. Although *T. pseudonana*
645 cells are cylindrical rather than spherical, this is a useful approximation for relative changes in
646 cell size. Red fluorescence (FL3) was used for chlorophyll fluorescence in order to distinguish
647 autotrophs from bacteria, but was not used here to approximate chlorophyll content, which was
648 instead estimated through acetone extraction (see below). Co-occuring bacteria, e.g. *Rhodobacter*
649 spp. and *Marinobacter* spp. are known to be crucial for nutrient recycling and signaling molecule
650 metabolism in *T. pseudonana* [43,44]. We therefore made no attempt to make the cultures axenic as
651 this would have incurred a substantial fitness cost. We quantified the associated bacteria by
652 regularly filtering the cultures between 0.2µm to 2µm, staining with SYBR gold and
653 enumerating bacterial densities and cell size via flow cytometry. Co-occurring bacteria
654 contributed less than 1% of total biomass throughout, with no systematic variation between time-
655 points or treatments.

656 Flow cytometry was further used in combination with a rhodamine stain to measure a
657 proxy for mitochondrial potential – the rate of ATP production by ATP synthase is proportional

658 to mitochondrial potential, which in turn depends on the activity of respiratory proteins
659 embedded in the mitochondrial membrane. Since each unit area of membrane has several ATP
660 synthase molecules and the rate at which each of these produces ATP depends on the
661 mitochondrial potential, the overall rate of ATP production per mitochondria depends on both
662 the area of mitochondrial membrane and the mitochondrial potential. Mitochondrial area is a
663 good approximation of the area of mitochondrial membrane, and changes in mitochondrial
664 potential can be measured in live cells using potentiometric dye. Previous work [45] has shown that
665 the styrl dye rhodamine 123 (R123) can be used in live cells of phytoplankton to measure
666 mitochondrial potential in real time. For staining, a 10 mg mL^{-1} stock solution in DMSO was
667 prepared and diluted to a 1μM working solution on the day of measurements. Of this, 20 μL
668 were added to 200 μL of sample and left to incubate in the dark for 45 minutes. To ensure that
669 longer incubation as is inevitable on a 96 well plate did not effect FL1 fluorescence as measured
670 in the accuri c6, samples were randomized across each 96-well plates with the first and the last
671 well containing the same sample. We found no effect of prolonged incubation after an incubation
672 period of 45 minutes.

673 A Nile Red stain was performed to measure relative lipid contents. Here, the dye was
674 used as a proxy to determine relative quantities of intracellular polar and neutral lipids[46] in
675 evolved samples, and in samples from short-term assays. The stock dye was obtained as a
676 powder. The working solution was diluted to 0.1 mg L^{-1}, and 10 μL were added to each 200 μL
677 sample on a 9-well plate (final concentration 15 mM) and left to incubate in the dark for 30
678 minutes, as pilot trials had shown that after this, fluorescence levels were stable long enough for
679 the time taken to measure one 96 well plate. Samples were randomized on the plate as described
680 above for R123 stains. As Nile Red excites in the same wavelength as chlorophyll (FL3) and

681 chlorophyll derivatives (FL2), samples were measured before and after adding the dye, and the
682 chlorophyll fluorescence subtracted from the fluorescence obtained after staining the sample.
683 While these traits are not discussed in detail here, they are part of the distance matrices used to
684 generate Fig. 4B (see also Fig. S8 and Table S12).

685

686 Growth rate trajectories

687 At the beginning and at the end of each transfer *T. pseudonana* cells were counted on the flow
688 cytometer as described above and used to estimate specific growth rates (μ d^{-1})

689 $$\mu = \frac{\ln(N_{t1}) - \ln(N_{t0})}{\Delta t} \quad (1)$$

690 where N_{t1} is the density of cells at the end of the transfer, N_{t0} is the inoculation density, and Δt is
691 the time passed in days.

692

693 Thermal tolerance curves

694 To characterize the thermal tolerance curves of the ancestor and each of the evolved lineages, an
695 inoculum of 100 cells per mL from the middle of the logarithmic phase of growth was
696 transferred into fresh media at 15ºC, 20ºC, 25ºC, 30ºC, 32ºC, 35ºC, and 40ºC. Cell count was
697 determined daily on a flow cytometer and populations were transferred to fresh media once at
698 each temperature during the middle of the logarithmic phase of growth, before being left to reach
699 stationary phase. Growth rates were estimated from the abundance data using a logistic growth
700 function:

701 $$N(t) = \frac{K}{1+\frac{K-N_{t0}}{N_{t0}}e^{rt}} \quad (2)$$

702 where t is time, K is carrying capacity of the population (cell mL^{-1}), r is the maximum growth
703 rate (d^{-1}) and N_{t0} is the cell count at t_0 (cell mL^{-1}).

704

705 Intracellular macromolecular composition

706 Cellular carbon (C), nitrogen (N), phosphorus (P) and chlorophyll content (total chlorophyll
707 calculated from chlorophyll *a* and chlorophyll *c* content) were quantified in the ancestor and each
708 of the evolved lineages. We investigated the effects of temperature on macromolecular
709 composition both over short-term acclimation and long-term evolutionary adaptation. Short-term
710 thermal acclimation was investigated by exposing the evolved lineages to the 15 to 45ºC thermal
711 gradient described above for quantifying the thermal tolerance curves for growth rate. Samples
712 from the 40ºC and 45ºC environment were omitted from further analyses due to high mortalities
713 that limited the amount of biomass available for further tests. Lineages were given 1 transfer to
714 acclimate to a given assay temperature in the gradient and then harvested during the logarithmic
715 phase of growth in the second transfer. Long-term evolutionary adaptation to the selection
716 regimes was quantified by harvesting evolved lineages directly from the experiment during the
717 logarithmic phase of growth

718 The remaining samples were prepared for C, N, P and chlorophyll measurement by
719 spinning down 50 mL of culture in a centrifuge at 4ºC and 3500RPM for 30 minutes. The pellet
720 was then transferred to a 1 mL Eppendorf tube, spun again for 15 minutes at 3500RPM and the
721 remaining supernatant decanted. Chlorophyll content was determined on a spectrophotometer
722 (Jenway 7351) after[47], with extraction in acetone, and absorption spectra measured in 10 nm

723 increments from 500 to 700nm, which span the full range of emission peaks for chlorophyll a
724 and chlorophyll c.

725 Pellets for C and N determination were freeze dried for 24 hours, transferred to zinc
726 capsules, weighed and analyzed using a SerCon Isotope Ratio Mass Spectrometer (CF-IRMS)
727 system (continuous flow mode). Intracellular P content was determined was determined via a
728 colourimetric reaction on a Seal Analytics AA3 segmented flow auto-analyzer. Freeze dried
729 pellets were washed in 0.17M Na_2SO_4, transferred to scintillation vials and re-suspended in 4 ml
730 0.017M $MnSO_4$. The samples were transferred to an autoclave (1h, 121ºC), shaken vigorously
731 and centrifuged at 2500 r.p.m for 30 minutes, the pellet discarded, and the supernatant brought to
732 10 ml with MillliQ purified water. The samples were immediately analyzed on the AA3 using
733 the colourimetric molybdate/antimony method [48]. Using these data we calculated the molar
734 stoichiometric ratios, C:N, C:P, N:P and the Chl:C ratio.

735 Silica content, although not a main trait under investigation here, was determined through
736 preparing pellets as described above. Then, 4ml of 0.2M NaOH were added to the pellet in
737 polyethylene tubes and the pellets vortexed briefly. The samples were then transferred to a
738 heating block at 90ºC for ca. 1h. 2 ml of 1M HCl were added, the samples vortexed again and
739 then spun at 3500 rpm for 1h. The supernatant was passed through an 0.5µm filter, pipetted into
740 fresh tubes, topped up to 10ml with MiliQ water and used immediately on the segmented flow
741 auto-analyser, where the colourimetric procedure followed that described in [49].

742

743 Photochemistry

744 We characterised a range of photochemical parameters in the ancestor and each of the evolved
745 lineages using fast repetition rate fluorometry (FastPro8, FRRf3, Fast Ocean System Chelsea

746 Technology Group). 500 µL of dilute sample (with a cell count of less than 1000 cells mL^{-1})
747 were added to 5 ml of fresh culture medium. Samples were then pre-incubated in the dark at the
748 assay temperature for 15 minutes in a water-bath, and another 10 minutes in the fluorometer to
749 make sure that samples were fully dark acclimated and all reaction centers closed. All
750 measurements were made at the selection temperature for the 22, 26, and 32ºC lineages, while
751 for the lineages in the fluctuating regime measurements were carried out at both 22 and 32ºC.
752 Photochemical traits were measured in response to rapid flashes at increasing light intensities
753 from 0 to1600 µmol m^{-2} s^{-1}. Flash frequency and rate followed standard protocols for
754 phytoplankton [50], with 100 flashes of 1.1 µs at 1 µs intervals. Peak emission wavelengths of the
755 LEDs used for excitations were at 450 nm, 530 nm, and 624 nm. Φ_{PSII} was particularly relevant
756 to our study as it is commonly used to describe the light responses of photosynthetic efficiency.
757 Φ_{PSII} values are used as an indication of the proportion of the total light absorbed that is used in
758 photochemical reactions in PSII (Fig. S3). Additionally, we determined rP, NPQ, and *C* (see Fig.
759 4B and SI Table 13 for PCA on phenotypic traits). These describe the relative rate of
760 photosynthesis in response to irradiance and are obtained as an estimate of electron transport
761 through PSII (rP), the cell's ability to maintain photochemical function at high light intensities
762 (NPQ, non-photochemical quenching) and the proportion of PSII reaction centers in a closed
763 state (*C*). As measurements were carried out following a light response curve, we were then able
764 to measure these functions both at saturating light intensity and at the light intensity that the
765 samples were grown at in the incubators.
766

767 Thermal responses of photosynthesis and respiration

768 Measurements of photosynthesis and respiration were made on the ancestor and all evolved

769 lineages after 300 generations of selection when in the middle of the logarithmic phase of
770 population growth. Net photosynthesis (*NP*) was measured as O_2 evolution at increasing light
771 intensities in intervals of 50 $\mu mol^{-1}\ m^{-2}\ s^{-1}$ up to 300 $\mu mol^{-1}\ m^{-2}\ s^{-1}$, and then in intervals of 100
772 $\mu mol^{-1}\ m^{-2}\ s^{-1}$ up to 1000 $\mu mol^{-1}\ m^{-2}\ s^{-1}$, followed by 200 µmol steps up to 2000 $\mu mol^{-1}\ m^{-2}\ s^{-1}$.
773 The maximum rate of light saturated photosynthesis was determined by fitting the *NP* data to a
774 dynamic model of photoinhibition via non-linear least squares regression using the methods
775 outlined in [19]:

776 $$NP(I) = \frac{NP_{max} I}{\frac{NP_{max}}{\alpha I_{opt}^2} I^2 + \left(1 - 2\frac{NP_{max}}{\alpha I_{opt}}\right) I + \frac{NP_{max}}{\alpha}} - R \qquad (3)$$

777 where $NP(I)$ is the rate of photosynthesis at light intensity I, NP_{max} is the maximum rate of net
778 photosynthesis at the optimal light intensity, I_{opt}, and α controls the rate at which $NP(I)$ increases
779 up to NP_{max} and R is the rate of respiration (i.e. the rate of O_2 flux when $I = 0$). Gross
780 photosynthesis (P) was estimated as $P = NP_{max} + R$. Measurements of O_2 flux were measured in
781 a Clark-type oxygen electrode (Hansatech Ltd, King's Lynn UK Chlorolab2). Aliquots (50 mL)
782 of the populations were concentrated through centrifugation to a density of approximately $8*10^4$
783 cells mL^{-1} and acclimatised to the assay temperature for 15 minutes in the dark before measuring
784 metabolic rates. The thermal responses of *P* and *R* were quantified by measuring rates over a
785 temperature gradient from 4ºC to 45ºC (in 3ºC increments, with additional measurements at the
786 selection temperatures). Rates of *P* and *R* were expressed in units of µg C per µg C using the
787 following equation [51]

788 $$b(\mu gC\ \mu gC^{-1}) = \frac{b(\mu mol\ O_2\ mL^{-1}\ d^{-1} * 32 * M * (\frac{12}{44})}{\mu mol\ C\ cell^{-1} * cells\ \ mL^{-1}} \qquad (4)$$

789 where b is the metabolic rate (either P or R), 32 * M * (12/44) is used to convert µmol O_2 into µg

790 C, and the factor *M,* is the assimilation quotient of CO_2:O_2. In our study, NO_3^- was the source of
791 nitrogen, thus assuming a set of balanced growth equations [51]

792 $$nCO_2 + (n+1)H_2O + HNO_3 \rightarrow (CH_2O)_n\, NH_3 + (n+2)O_2 \qquad (5)$$

793 if the C:N ratio is calculated in moles the assimilation ratio of CO_2:O_2 will be n/n+2 (see[51]). The
794 estimated *M* values are given in Table S1 and ranged from ~ 0.77 to ~ 0.79. Alongside the
795 biomass measurements for phytoplankton, we also periodically measured bacterial biomass (see
796 methods detailed above), which typically comprised <1% of algal biomass. Consequently,
797 bacterial respiration did not contribute significantly to metabolic rates, and when 2µm filtered
798 samples were run for metabolic rate measurements, the slopes were indistinguishable from water
799 blanks.

800

801 <u>DNA extraction for Whole Genome Re-Sequencing</u>

802 At the beginning (t0) and after 300 generations (t300) of selection in each environment, 250ml of
803 each sample in exponential growth were spun to a pellet in 50ml batches at 4ºC and 3500 RPM
804 (rounds per minute). The supernatant was discarded and the pellet stored at -20ºC until further
805 use. Prior to DNA-extraction, samples were put through three cycles of thawing and refreezing,
806 as this had shown to increase yield of diatom DNA relative to bacterial DNA in pilot studies.
807 DNA was extracted following a standard cTAB protocol in chloroform:isoamyl and isopropanol
808 (Murray and Thompson, 1980) with a proteinase K step to digest proteins, and an RNAse step to
809 digest RNA. The resulting DNA pellet was quality controlled on 1% agarose gels and
810 fluorometrically on a Qbit system. The extracted genomic DNA was then frozen in TE buffer at -
811 20ºC until samples were processed at the Exeter Sequencing Services (University of Exeter,

812 UK). Whole genome re-sequencing (Illumina) was carried out on a HiSeq 2500 platform with
813 paired-end reads. Library preparation was carried out at the Exeter Sequencing Services using
814 Netflex. The resulting reads were assembled against the existing *T. pseudonana* genome [11].
815
816 Alignment of sequence reads against the reference genome
817 A prerequisite of variant calling from genomic shotgun sequencing data is the alignment
818 of sequence reads against a reference sequence[46], in this case the genome sequence of *T.*
819 *pseudonana* (GenBank: GCA_000149405.2). Prior to alignment, reads were trimmed and filtered
820 to remove low-confidence base-calls by applying TrimGalore with command line options "-q
821 30" and "--paired". Alignment was performed using BWA-mem version 0.7.12-r1039 [52,53]. As a
822 precaution against artefacts arising from non-uniquely mapping reads we used command line
823 option "-q 1" when using *samtools view* [53] to convert the alignments from SAM to BAM format;
824 the use of this option has the effect of excluding from the alignment all sequence reads that map
825 equally well to more than one unique site on the reference genome sequence.
826
827 Calling single nucleotide variants (SNVs)
828 We identified candidate single-nucleotide variants in BWA-mem alignments of genomic
829 sequence reads against the reference genome sequence using the SAMtools/BCFtools pipeline
830 version 1.3.1 [53]. Candidate SNVs were identified using *bcftools* 1.3.1 with the following
831 command-lines:
832 "samtools mpileup -u -f genome.fasta alignment.bam > alignment.bcf" and
833 "bcftools call -m -v –Ov alignment.bcf > alignment.vcf".
834

835 The candidate variants were then filtered using the following command line: “bcftools filter --
836 SnpGap 10 --include '(REF="A" | REF="C" | REF="G" | REF="T") & %QUAL>=35 &
837 MIN(IDV)>=2 & MIN(DP)>=5 & INDEL=0' alignment.vcf > alignment.filtered.vcf”. This
838 filtering step eliminates indels with low-confidence single-nucleotide variant calls. It also
839 eliminates candidate SNVs within 10 base pairs of an indel, since alignment artefacts are
840 relatively common in the close vicinity of indels. Finally, annotation was added to each filtered
841 VCF file using snpEff [54] version 4.3t.

842 The motivation for variant calling was to identify candidate genetic changes in the
843 evolved populations (compared to the ancestral clone) that might contribute to thermal
844 adaptation. Specifically, two types of genetic adaptation were of interest: (i) mutations arising *de*
845 *novo* after the t0 time-point and becoming fixed in an evolved population and (ii) alleles present
846 in the ancestral clone, that become fixed in an evolved population. In both cases, fixation could
847 arise through either drift or selection. Fixation of *de novo* mutations would result in allele
848 proportion rising from zero at t0 to close to 1 by 300 generations. Because the ancestral
849 population was a genetically homogeneous clone, standing variation consists solely as
850 heterozygosity, and we expect to observe alleles increasing in proportion from 0.5 at t0 to close
851 to one by 300 generations. Therefore, we took two distinct approaches to surveying variation: (i)
852 identifying SNVs that were undetectable in the ancestral population (at a sequencing depth of at
853 least 10 x) but were supported by 100% of the aligned reads in an evolved population (at depth
854 of at least 5 x) and (ii) identifying SNVs that plausibly had a population proportion of 0.5 in the
855 ancestral population (due to heterozygosity), at depth of at least 10 x, and were supported by
856 100% of the aligned reads in an evolved population, at depth of at least 5 x (see also Table S14
857 for average sequencing depths).

858 The output of SNV-calling was a matrix of estimated allele frequencies at each SNV site
859 for each sample. All SNVs included in the matrices had passed reliability filtering, as described
860 above. Following the approach of a recent study[50], the estimated allele frequency was defined as
861 the ratio of sequence reads matching the alternative allele versus total number of reads aligned at
862 that site. In other words, the sample proportion is used as an estimate of the proportion:

863 $$f_{pm} \approx \frac{A_{pm}}{D_{pm}}$$

864 where for mutation m and population p, A_{pm} and D_{pm} are respectively the count of aligned reads
865 matching the alternate allele and the total counts (*i.e.* depth). Inevitably, this estimator of f_{pm} will
866 be confounded by sampling error. Assuming that each sequence read is drawn randomly from the
867 population, the sample proportion will follow a binomial distribution:

868 $$A_{pm} \sim \text{Binomial}(D_{pm}, f_{pm})$$

869 Given the sample size (in this case, D_{pm}) and the population proportion (f_{pm}), it is possible
870 to calculate the 95% confidence interval of the estimate of f_{pm}[55]. With small sample sizes, the
871 confidence intervals are wide. For example, with D_{pm} (depth) of 10 x and f_{pm} (allele frequency)
872 of 0.5, the 95% confidence interval is 0.5 ± 0.31. For depth 50 x it is 0.5 ± 0.14 and for depth
873 100 x it is 0.5 ± 0.10 (see also Fig. S7). Therefore, in this study, we refrain from making
874 inferences based on small changes in allele proportions among populations, but rather restrict
875 this to scenarios where estimated proportion changes from zero to one or from 0.5 ± 0.2 to one.
876 Notwithstanding the problems of precisely quantifying allele frequency, it is
877 straightforward to identify with confidence those SNVs that were absent in the ancestral
878 population and at very high abundance in one or more of the evolved populations after 300
879 generations. A variant m was considered absent in the ancestral t0 population if D_{t0m} >= 10 and
880 A_{t0m} = 0. Since this ancestral population was clonal, the possible allele frequencies that D_{t0m} can

take are zero or unity or, if heterozygous, 0.5. Assuming a binomial distribution, the probability of an allele with population frequency of 0.5 being not represented among a sample of 10 aligned sequence reads is $0.5^{10} = 0.00098$. Similarly, if in a t300 population p, $D_{pm} >= 5$ and $A_{pm}/D_{pm} = 1$, then it can be confidently inferred that variant m is indeed present in population p. Assuming a high sequencing error rate of 1%, the probability of falsely calling the presence of a variant on this basis is $0.01^5 = 1e^{-10}$; in practice, Illumina error rates are closer to 0.1% than 1% making the probability of a false positive even smaller. Thus, we include in our analyses any candidate variant m in a population p if all the following criteria are satisfied:

i. $D_{t0m} >= 10$
ii. $A_{t0m} = 0$
iii. $D_{pm} >= 5$
iv. $A_{pm}/D_{pm} = 1$
v. Variant p is called by the *bcftools* and passes filters in at least one population

The matrices containing SNVs that had passed reliability filtering, had become fixed after 300 generations, and were within protein-coding regions, were used for the downstream statistical analyses implemented in R using the packages vegan (2.4-3), mixOmics (6.1-3), ade4 (1.7-10), topGO (2.26-0,with Bioconductor 3.4) and biomaRt (2.3-0 with Bioconductor 3.4), and phangorn (2.2-0).

The matrices were first used for principal component analyses based on Euclidean distance matrices. From these, we pulled the top 100 loadings for the two major axes of variation, and associated them to a treatment through the plotLoadings functions within the mixOmics package (6.1-3), which performs discriminant PCA and then represent the loading weight of each selected

904 variable on each component, i.e. the highest absolute value indicating the strongest association of
905 a gene with SNVs to a treatment group. This allows us to find treatment associated groups in
906 which the changes in a gene or SNVs are maximal [56,57]. From this, we can infer a number of
907 genes, were SNVs are most likely associated with a treatment. For these, we performed an
908 enrichment analysis for gene ontology using the R package topGO (2.26-0,with Bioconductor
909 3.4) and biomaRt (2.3-0 with Bioconductor 3.4). We created a 'gene universe' as a reference
910 data set, using the entire *T. pseudonana* genome, and matching all gene locus tags with the
911 appropriate GO term for biological processes through protists.ensemble.org. Our test data frame
912 contained the genes that had been drawn from the top 100 loadings of the two major principle
913 component axes, the associated GO terms, and the PCA loadings. Based on these we performed a
914 Kolmogorov-Smirnov like test which computes enrichment taking into account not only the GO
915 hierarchies, but also the PCA loading scores. This allow us to test the over-representation of Go
916 terms within the group of genes found to be most important in the PCA (table 2).
917

918 To test for separation of samples by treatment, and within treatment variation, the Euclidean
919 distance matrices were used to perform AMOVAs (Analyses of molecular variation) following
920 [37]. Briefly, the Euclidean distance matrix of SNVs is used to apportion molecular variance
921 according to hierarchical subdivision. In our case, this is variation within populations (biological
922 replicates), between populations (biological replicates within the same treatments), and between
923 treatments. To disentangle whether i) there were differences in the amount of variation between
924 populations in different treatments and ii) treatments were significantly genetically divergent
925 from each other, AMOVAs were performed on populations from each treatment separately
926 (yielding a measure of variation that can be compared across treatments) and for all populations

927 in all treatments combined (yielding a measure of variation, or distance, between populations
928 from specific treatments). The significance of the different components of variance was tested
929 using a permutation approach[37], which removes the assumption of normality that underlies
930 conventional ANOVAS, but which is often not appropriate for molecular data.
931
932 All trees for were built within the phangorn package through neighbor-joining, and annotation of
933 trees was carried out in the FigTree software (FIGTREE 1.4.3
934 http://tree.bio.ed.ac.uk/software/figtree/).
935
936
937
938 Assessment of the abundance of other taxa
939 The diatom cultures used in this study were not axenic and therefore the genomic sequence data
940 contained both sequence reads originating from the diatom and also from other organisms such
941 as bacteria that were present in the cultures. To estimate the relative abundances of taxa
942 represented in the data, we used BLASTN (version 2.5.0+) to align 10,000 sequence reads from
943 each sample against the NCBI's non-redundant Nucleotide (also known as "NT") database [58]
944 downloaded on December 9th, 2016, using command-line options "-num_alignments 20 -evalue
945 1e-10". Matches were assigned to taxa using MEGAN (version 5.11.3) [59]. Similar to analysis of
946 the diatom data, we combined the allele frequencies into a single matrix, which was used for the
947 downstream statistical analyses implemented in R using the 'vegan' (2.4-3) and 'ecodist' (2.01)
948 packages. Distance matrices using the Bray-Curtis index were created from these, on which we
949 ran PERMANOVAs to test for separation of samples by treatment. Pairwise contrasts between

950 treatments were examined in PERMDISP followed by TukeyHSD post-hoc tests. See Fig. S14
951 and Table S16 for a comparison of bacterial composition across treatments.
952
953 Other Statistical Analyses
954 *Fitness trajectories*
955 The resultant time series of specific growth rates were analyzed using a generalised additive
956 mixed effects model (GAMM) to assess whether the fitness trajectories differed between the
957 selection regimes. We used GAMMs to account for the hierarchical nature of our experimental
958 data. For example, our experimental design yielded replicate fitness trajectories in each selection
959 treatment. This hierarchical structure meant that measurements were non-independent – e.g.
960 measurements from the same replicate will be auto correlated. We account for this by treating
961 replicate as a random effect on the intercept of the model, which models deviations among
962 replicates from the fixed effects as normally distributed with a mean of zero. The most complex
963 models included an effect of treatment (e.g. '22', '26', '32', '22-32') on the intercept (which
964 characterises the median value of the response variable) and also allowed the shape of the time
965 series, which was modeled using a cubic regression spline, to vary among treatments. Treatment
966 effects on the shape and intercept of the seasonal phenology were modeled as fixed effects in the
967 GAMMs. Model selection entailed fitting a range of models to the data, starting with the full
968 model and then a series of reduced models with interaction terms and main effects removed to
969 test hypotheses about the potential differences in the fitness trajectories among treatments. For
970 multi-model selection, we computed small sample-size corrected AIC scores (AICc) and then
971 compared between models by calculating delta AICc values and AIC weights using the 'MuMIn'
972 package. GAMMs were fitted to the data using the 'gamm4' package and were conducted in R

973 (v.3.23) (see Fig. S15 for trajectories per biological replicate).

974

975 *Thermal responses of growth and metabolism*

976 The thermal responses for growth, photosynthesis and respiration were quantified using a
977 modified version of the Sharpe-Schoolfield equation (see [60,61] for the original equations), which
978 assumes that the rate of growth or metabolism is limited by single enzyme catalyzed reaction

979 $$\ln\big(b(T)\big) = E_a\left(\frac{1}{kT_c} - \frac{1}{kT}\right) + \ln\big(b(T_c)\big) - \ln\left(1 + e^{E_h\left(\frac{1}{kT_h} - \frac{1}{kT}\right)}\right) \quad (6)$$

980 where $b(T)$, is the rate of metabolism (in µg C µg C^{-1} d^{-1}) or growth (d^{-1}), k is Boltzmann's
981 constant (8.62×10^{-5} eV K^{-1}), E_a is the activation energy (in eV) for the metabolic process,
982 indicative of the steepness of the slope leading to a thermal optimum, T is temperature in Kelvin
983 (K), E_h characterises temperature-induced inactivation of enzyme kinetics above T_h where half
984 the enzymes are rendered non-functional and $b(T_c)$ is the rate of metabolism at an arbitrary
985 reference temperature, here T_c = 18ºC, where no low or high temperature inactivation is
986 experienced. Equation (6) yields a maximum metabolic rate at an optimum temperature, where
987 metabolic rates are fastest:

988 $$T_{opt} = \frac{E_h T_h}{E_h + kT_h \ln\left(\frac{E_h}{E_a} - 1\right)} \quad (7)$$

989 Equation (6) differs from the Sharpe-Schoolfield equations in several ways. First, we
990 exclude parameters from Eq. (6) used to characterise low-temperature inactivation due to
991 insufficient data to quantify this phenomenon in our analysis. Second, rather than characterise

992 temperature effects below T_{opt} using the Eyring (1935) relation, $\left(\frac{T}{T_c}\right) e^{E_a(\frac{1}{kT_c}-\frac{1}{kT})}$, we instead use

993 the simpler Boltzmann factor, $e^{E_a(\frac{1}{kT_c}-\frac{1}{kT})}$. This simplification enables an explicit solution for T_{opt}

994 (Eq. 7) and facilitates more direct comparison with previous work on the temperature

995 dependence of metabolism using metabolic theory [62-65].

996 The parameters $b(T_c)$, E_a, E_h, T_h, and T_{opt}, represent traits that characterise a unimodal

997 thermal response curve, and we expect them to differ between selection regimes owing to

998 thermal adaptation. To test this hypothesis, we fitted the rate data to Eq. (6) using non-linear

999 mixed effects models in the 'nlme' package in R. We analyzed the photosynthesis, respiration

1000 and growth data in separate mixed effects models. We also analyzed the thermal responses for

1001 ancestor separately from the data for the evolved lineages following the selection experiment.

1002 Models included random effects on each of the parameters of Eq. (6) by replicate, and for the

1003 analysis with evolved lineages, 'selection environment' as a fixed four level factor on each

1004 parameter. For the analysis of the metabolism data for the ancestor we included both

1005 photosynthesis and respiration data together, with 'flux' as a two-level fixed factor on each of the

1006 parameters in Eq. (6), to establish whether the thermal responses differed between these fluxes

1007 and likely physiological constraints on growth prior to the selection experiment. Model fitting

1008 and selection started with the most complex possible model, including fixed and random effects

1009 on all parameters. Model selection then proceeded by first removing treatment effects on the

1010 parameters individually, then in pairs of two and in pairs of three, and finally, by removing the

1011 treatment-effect all four parameters. For multi-model selection we computed small sample-size

1012 corrected AIC scores (AICc) and then compared between models by calculating delta AICc

1013 values and AIC weights using the 'MuMIn' package. When candidate models deviated from the

1014 most parsimonious model (that with the lowest AICc score) by less than two AICc units,

1015 parameters were averaged across those candidate models. The relative importance of the fixed
1016 factors in the averaged model was determined using the sum of their relative weights (see Table
1017 S5 and S6).

1018

1019 *Carbon-use efficiency (CUE)*

1020 Carbon use efficiency, as the *potential* for carbon allocation to growth, was calculated from the
1021 gross photosynthesis (P) and respiration (R) data as CUE = $1 - R/P$. For statistical analysis, we
1022 used CUE calculated at the temperature of the selection environment. In the fluctuating
1023 treatment, we used data for photosynthesis and respiration at 32ºC to aid comparison with the
1024 populations experiencing 32ºC throughout. We fitted a linear mixed effects model to these data,
1025 with 'selection regime' as a fixed effect and biological replicate as a random effect. Model
1026 selection proceeded as described above. See also Fig S5.

1027

1028 *Macromolecular composition and photosynthetic efficiency*

1029 The C, N and P content, the C:N, C:P, N:P, Chl:C ratio, size and silicate content, as well as Φ_{PSII}
1030 at high and low light, were each analyzed using a separate mixed effects model, where 'selection
1031 regime' was a fixed effect and replicate was a random effect on the intercept (see Table S10 and
1032 S11 for details). To determine quantify the effects temperature on the elemental composition of
1033 the ancestor, a separate mixed effects model was run for ancestral samples only, with 'assay
1034 temperature' as a fixed factor and replicate as a random effect (see also Table S11 for details).
1035 For both datasets, model selection was as described above, with the best model being chosen
1036 based on the lowest AICc and high AICc weight.

1037

1038 We quantified the light response curve of Φ_{PSII} using an exponential decay model

1039 $$\Phi_{PSII} = a * \exp(I * b) \quad (8)$$

1040 where I is the irradiance, a is a normalization constant, and b is the rate constant which

1041 characterises how rapidly Φ_{PSII} declines with increasing I. Eq. (8) was fitted to the Φ_{PSII} data

1042 using a non-linear mixed model including random effects by biological replicate on each

1043 parameter and 'selection environment' as a fixed factor. Model output was then used to calculate

1044 Φ_{PSII} at the light intensity that samples would have experienced in the incubators (i.e. ~ 100µmol

1045 quanta m^{-2} s^{-1}). Model selection then proceeded as described above for P and R data. Detailed

1046 model output and model selection data are available in Table S9 where nomenclature and

1047 acronyms are as given here (see also Fig. S3).

1048

1049 On the phenotype data, we ran PERMANOVAs (Permutational Multivariate Analysis of

1050 Variance) to test for separation of samples by treatment. Here, the test is based on the distance

1051 matrix containing all phenotypic data of the evolved lineages from all treatments relative to the

1052 ancestor at that same temperature. Subsequently, contrasts between treatments were examined

1053 through a test of homogeneity of dispersion (permdist in R), followed by TukeyHSD post-hoc

1054 tests.

1055