Quantifying the selection regime in a natural *Chironomus riparius* population

Markus Pfenninger[1,2,3] & Quentin Foucault[1,2]

[1]Dept. Molecular Ecology, Senckenberg Biodiversity and Climate Research Centre, Senckenberganlage 25, 60325 Frankfurt am Main, Germany

[2]Institute for Molecular and Organismic Evolution, Johannes Gutenberg University, Johann-Joachim-Becher-Weg 7, 55128 Mainz, Germany

[3] LOEWE Centre for Translational Biodiversity Genomics, Senckenberg Biodiversity and Climate Research Centre, Senckenberganlage 25, 60325 Frankfurt am Main, Germany

Corresponding author: Markus Pfenninger

Markus.Pfenninger@senckenberg.de

Summary

While the evolutionary fitness of natural populations is affected by a multitude of environmental factors, theory predicts that selective responses are in principle limited. However, we lack empirical knowledge on the magnitude of different selection pressures natural populations adaptively track. Here, we developed a framework to investigate the quantitative and qualitative complexity of the effectively acting selection regime using population genomic time series data. We applied the approach to a natural population of the multivoltine midge *Chironomus riparius*. Using six seasonal samples over three years from the same natural population, we could show with fitness experiments that the population continuously evolved in response to a highly variable environment. Analyses of genome-wide allele-frequencies revealed that tens of thousands of haplotypes responded at least once to selection during the monitored period. Clustering the temporal haplotype frequency trajectories revealed 46 different patterns, i.e. selection pressures. Some of these co-varied with measured environmental variables known to be selective factors for the species. Our results demonstrate that 1) adaptive tracking of multiple fluctuating selection pressures occurs in natural populations, 2) the estimated minimum number of simultaneously acting selective pressures is quite high but appears to be limited and 3) changes in intensity and direction of selective responses can be frequent. This shows that adaptation in natural populations can be rapid, pervasive and complex

Introduction

Evolution in natural populations is expected to occur when selection acts on traits with a heritable basis of phenotypic variation. There is a general agreement that selection in the wild is intense [1,2]. It is also variable in space and time [3-5], even though there is some debate whether changes in the direction of selection are frequent or not [6-8]. Recent theoretical and empirical work has shown that selection in natural population can lead to very rapid adaptation [9], in particular of polygenic traits [10]. If the rate of environmental change is not too fast and the population characteristics allows for effective selection, adaptation from standing genetic variation to track moving phenotypic optima is theoretically possible nearly in real-time [11].

However, adaptation is not without limits. Theoretical approaches suggest that if many arbitrary selection pressures are acting simultaneously (multivariate selection [12]), the ability of populations to

respond to these may be severely reduced [13,14]. The number of loci in a genome that can react to different selection pressures is principally limited by the number of independently inherited genome sections due to physical linkage [15]. In addition, trade-offs and pleiotropy between traits constrain evolution further [16,17]. Calculations thus suggest that the number of genetically truly independent traits is in the order of a few dozen, despite abundant genetic variation and strong selection [14,18].

A better understanding of the selective limits in natural populations is not only a central concern of population genetics and macroevolution [13], it also has wide-ranging consequences for conservation [19] and the potential for adaptation of biodiversity to global change [20]. To our knowledge an attempts to empirically test the hypothesis of limited selective response capacity, e.g. by comprehensively quantifying the different selective responses of a natural population, is lacking. This may be due to principal difficulties in identifying and measuring all the relevant, yet independent phenotypic traits in natural populations [14,21].

Here we approach this problem from a different angle. Experimental evolution has shown that population genomic time series of evolving populations are a powerful tool to identify selected haplotypes throughout the genome and follow their temporal trajectory [22]. We extended this approach to natural populations. The rationale is as follows: if a selection pressure changes in intensity and/or direction over time, allele frequencies of loci responding to this pressure - and sites physically linked to them - should shift towards the new optimum equilibrium frequency [10], just as allele frequencies respond along selective clines in space [23]. For quantitative traits with a broad genetic basis, subtle adjustments of allele frequencies at many loci usually suffice and not necessarily their fixation [24].

In environments with fluctuating selection, physically unlinked haplotypes responding to the same selection pressure should therefore show temporally correlated allele frequency changes, following the same evolutionary trajectory [3,25]. Given sufficiently long time series and a variable selection regime, the number of differentiable allele frequency trajectory patterns over time should therefore give a minimum estimate of the selective forces effectively acting on this population.

We applied this approach to a natural population of the non-biting midge *Chironomus riparius*. The species is widely distributed in the temperate northern hemisphere [26]. *C. riparius* spends most of its life-cycle as larva in freshwater habitats like small streams, ditches and puddles [27]. The species is amenable for manipulation, in particular fitness measurements, in the laboratory [28] and comprehensive genomic resources are available [29,30]. Many environmental factors are known to act as selection pressures on this species (e.g. organic load [31], conductivity [27], nitrogen [32], temperature [33], anthropogenic substances [34]). Its multivoltinity (about 10 generations per year in the studied population [35]) together with the high effective and demographic population size (> 1,000,000 [36]) and the very high number of offspring per breeding pair (400-800) allows for rapid adaptation [37,38]. The natural and anthropogenic variability of environmental parameters in small streams situated in the intensely used landscape of Central Europe likely creates a temporally highly variable multivariate selective regime, i.e. a setting that allows differentiating between the evolutionary reactions to different selective forces.

Using six seasonal samples over three years from the same natural population, we could show with fitness experiments that the population continuously evolved in response to a highly variable environment. Population genomic time series analyses revealed that the minimum number of

different selection pressures that the population effectively follows is quite high but limited. Changes in intensity and direction of selective responses were frequent.

Results

Changes in environmental parameters over time

Measured physico-chemical water parameters varied substantially through time (Suppl. Fig.2). The UPGMA dendrogram on these environmental parameters, which represented the original distance matrix reasonably well (cophenetic distance 0.82), showed that the environmental conditions changed continuously over the monitored period (Suppl. Fig. 3). The measured environmental regime never returned close to any previous state. This was mirrored in the poor temporal autocorrelation of the parameters, except for seasonal temperature changes (Suppl. Fig. 4).

Fitness changes over time evidenced continuous evolution

A two-way ANOVA showed that the natural population showed highly significantly different fitness, measured as population growth rate, among the sampling dates when exposed to three different experimental temperature conditions ($F = 9.509$, $p = 5.07 \times 10^{-6}$, Figure 1). Also the interaction between sampling date and three experimental temperatures (14°C, 20°C and 26°C) had a significant influence on the measured fitness ($F = 3.631$, $p = 0.0017$, Suppl. Table 1). As expected, the experimental temperature had the largest fitness effect ($F = 269.6$, $p = 1 \times 10^{-30}$, Suppl. Table 1).

Large parts of the genome showed evidence of selective evolution

The genome-wide mean estimate of distance-dependent linkage disequilibrium (LD) dropped to an r^2 of 0.5 within 10 base pairs and to the genome-wide background level within 120 bp (Suppl. Fig. 5). Consequently, the expected haplotype structure was extremely short.

Allele frequency estimates for all sampling dates were obtained for 22,693,348 SNP positions in the genome. Neutral drift simulations, taking observed sampling variance of the poolSeq approach into account, showed that the chosen selection detection threshold of 2.0 -log10p for an allele frequency change (AFC) between consecutive sampling times or between the first and the last sampling implied a false discovery rate of $q < 0.001$. This corresponded to an AFC of at least 0.15. Under the condition of at least one selected AFC between any two consecutive sampling times or the first and the last sampling time, identified a set of 367,446 allele frequency trajectories (AFT) with signs of selection.

AFTs were combined to haplotypes by a correlation coefficient clustering threshold of 0.75 among consecutive SNPs along the genome. This resulted in 281,153 haplotypes of which 35,574 consisted of more than a single SNP. Of the 100 haplotypes randomly chosen for validation, 98 showed the expected linkage between the marker SNPs of the haplotype (Suppl. Fig. 6). In two cases, both with more than 2 haplotype marker SNPs, one of the marginal SNPs was only slightly linked ($r < 0.6$). All haplotypes checked were distinct from their respective neighbours. Only haplotypes with more than a single SNP were retained for subsequent analyses. In these, the average number of SNPs per haplotype was 3.30 +/- 2.64 (mean +/- s.d.). The mean length of inferred haplotypes was 893.1 +/- 1790.3 bp. They were on average 3466.0 +/- 4293.5 bp apart. Of all 13,490 annotated genes in the genome, 5793 (= 42.9%) harboured at least one selected haplotype, some more.

The mean selection coefficient over all retained haplotypes was 0.30 with a standard deviation of 0.68. The latter is equivalent to the intensity of selection over time, and its much larger value relative to the mean indicated large changes in magnitude and frequent reversals in the direction of selection [3].

Clustering of allele frequency trajectories

To avoid using an arbitrary clustering threshold, we used a Monte-Carlo simulation approach. Clustering of simulated random AFTs, drawn from the observed AFC distribution, indicated that a clustering threshold of 0.75 yielded informative clusters without undue lumping (Suppl. Fig. 7). Applying this threshold, the 35,574 AFTs were grouped into 46 distinct cluster (Fig. 2). The number of AFT cluster members was highly different. The largest cluster (AFT16) contained 17,914 individual haplotype trajectories, the smallest merely three (AFT28), the mean was 816.8 (Suppl. Table 2).

GO term analysis of AFTs and correlation of AFTs to environmental parameters

For 44 out 46 AFT clusters, at least one significantly enriched GO term was found (Suppl. Table 2). Most enriched terms for the individual AFT clusters were on a very high hierarchical level. Consequently, they were rather uninformative or affected only a small minority of the genes with overlap to a selected haplotype. However, some hinted at a specific biological process or supported the observed association of the AFT cluster to a measured physico-chemical variable. Ten allele trajectories were highly correlated to measured water parameters in the period before sampling. Taking only $r > 0.9$ into account, three AFTs (AFT19 $r = 0.922$, AFT18 $r = 0.901$, AFT17 $r = 0.926$) were correlated to changes in NO_3. Two AFTs followed variation on Biological Oxygen Demand (BOD) closely (AFT20 $r = 0.902$, AFT21 $r = 0.916$). Two other AFTs were associated to NO_2 (AFT39 $r = 0.933$, AFT40 $r = 0.913$), AFT3 was correlated to PO_4 ($r = 0.915$), AFT 38 to temperature ($r = 0.961$), while the strongest association was found for AFT 34 and oxygen ($r = 0.984$, Figure 3).

The enriched terms for AFT4 (GO terms *0051704 multiorganism process*, *0006952 defense response*) suggested reaction to a biological interaction. The largest cluster AFT16 contained hundreds of genes involved in osmoregulation, reflected in the enrichment for GO terms *0055085 transmembrane transport* and *0006811 ion transport*. Despite affecting only about 20 genes, AFT36 showed an enrichment of terms associated with *porphyrin-containing compound biosynthetic process (0006779)* and *pigment biosynthetic process (046148)*. Genes affected by AFT40 haplotypes were enriched for signal transduction processes (*0007186 G protein-coupled receptor signaling pathway*), perhaps in conjunction with hormones (*1901362 organic cyclic compound biosynthetic process*, *0008202 steroid metabolic process*). AFT3 was enriched for term (*0006811 ion transport*). AFT39, with *calcium (transmembrane) transport* (terms *0006813* and *0071805*), which is co-annotated to *nitric-oxide synthase inhibitor activity*. There was an enrichment of the term *0005975 carbohydrate metabolic process* in AFT21. AFT22 was enriched *for 0007688 sensory perception of smell* and *0006979 response to oxidative stress*.

Discussion

Natural selection is thought to be a major force in the evolution of natural populations, yet we have a poor understanding of the quantitative and qualitative complexity and dynamics of the multivariate selective regime acting upon them. Most studies focus on a single or a few selection pressures, even though the environment is highly complex. Using a new approach, we investigated the selective regime effectively acting on a natural population of the non-biting midge *C. riparius* by surveying the temporal changes on the ultimate level of selection, the changes of allele frequencies. Below, we discuss the relevance of the findings for our understanding of evolution in natural populations and relate them to previous work.

Complex and unpredictable multidimensional fluctuating selective regime in natural populations

The largely uncorrelated temporal changes as well as the observed range of the measured physico-chemical water parameters suggested that the local environment presented a highly fluctuating, multidimensional selective regime for *C. riparius*. For most of these parameters, the fitness relevance and thus their selective potential for the species has already been shown [27,31-33,39]. Even though some

172 parameters showed seasonal changes, these were either overlain by long-term trends or were minor
173 in comparison to the observed jumps that did not seem to follow any regular pattern. In
174 consequence, the selective regime as a whole defined by this physico-chemical parameter set was
175 likely at no point in time identical or even similar to any other period. However, only a small choice of
176 the potentially selectively relevant biotic and abiotic parameters was measured. For example,
177 predation [40] and other intra- and interspecific ecological interactions are known to play an important
178 role for the fitness of *C. riparius* [41,42]. It is therefore reasonable to assume that an unknown, but much
179 larger number of selective factors is simultaneously, yet in strength and direction more or less
180 independently and irregularly, acting on this population. The local habitat is therefore probably
181 conforming to the multi-scale theory of the environment [3].

182 *Continuous fitness evolution*

183 The life-cycle tests showed significant differences in population growth rate as integrative fitness
184 measure among sampling dates, thus concurring to selective dynamics patterns generally known
185 from natural populations [6]. We assumed that the seasonal water temperature is the most important
186 environmental parameter for selection. The initial expectation was thus that, if seasonal adaptation
187 to temperature exists, seasonal samples should perform relatively better at experimental
188 temperatures that resembled the natural temperature regime experienced during the period before
189 the sampling took place. However, it was recently shown with the same population that the
190 laboratory conditions as such and not applied temperature regime constituted a by far more
191 important and complex selection pressure [38]. The experiment therefore rather tested the fitness
192 performance of the natural population, adapted to a complex natural selection regime, in another
193 complex experimental selection regime. The observed differences in fitness among sampling dates
194 therefore evidenced more or less continuous evolutionary changes in the field population that
195 affected fitness in the constant experimental conditions.

196 *Rapid and pervasive selective genomic evolution*

197 The analysis of population genomic time series showed that tens of thousands loci showed signs of at
198 least one selection-driven changes in allele frequencies within the less than 30 generations over the
199 period monitored. Rapid and substantial changes in allele frequencies were observed throughout the
200 genome. As the chosen approach accounted for drift and sampling variance [43], and immigration was
201 unlikely to change allele frequencies noticeably [44], the observed significant changes are likely due to
202 selection alone. This supports the increasing observations of rapid adaptive change in natural
203 populations (e.g. [45,46]). The observation that the standard deviation of the selection coefficients, an
204 equivalent to selection intensity, was much larger than the respective mean (0.68 vs. 0.30) proved
205 substantial changes in magnitude and frequent reversals in the direction of selection. The acting
206 selection regime thus corresponded to the multi-scale model of fluctuating selection advocated by
207 Bell [3]. In classical population genetic predictions, adaptive tracking as observed here is viewed as
208 unlikely to evolve, because fluctuating selection should rapidly eliminate genetic variation [47]. Several
209 recent models, however, have called these predictions into question. It was shown that in a
210 multivoltine species, adaptive tracking is more likely to evolve than phenotypic plasticity or bet-
211 hedging [48]. Dominance shifts of adaptive alleles [49] and boom-bust cycles [50] can additionally contribute
212 to the maintenance of genetic variation and are probably common in species satisfying these
213 conditions [51].

214 The propensity to rapid adaptation with simultaneous preservation of genetic variation in *C. riparius*
215 is thus probably due to an interplay of genomic, life cycle and demographic traits of the species. On
216 the genomic level, the low observed mean LD resulted in a highly fragmented haplotype structure,
217 comparable in scale to other insects like *Drosophila* [52]. A short range haplotype structure is a

218 precondition for the observed short distance between haplotypes responding to selection. Since only
219 haplotypes with more than a single SNP were considered, this may even present an underestimate,
220 because ancient selected polymorphisms may very well be completely separated from their
221 background by recombination. In some instances, even different parts of the same gene followed
222 different evolutionary trajectories. The observed fine-scaled selective response to independently
223 fluctuating environmental parameters requires that many different recombinational variants are
224 permanently present in the population. This may be one reason for the gross overproduction of
225 offspring in the species [44]. A large offspring number implies a respectively large number of
226 independent recombinational events, creating a wide variety of new genotypes from which to select.
227 Together with the large, fluctuating demographic and effective population size, the preconditions for
228 rapid, fine-scaled and variable evolutionary reaction to fluctuating changes in the selective regime
229 are therefore given.

230 *The number of temporal evolutionary trajectories is limited but dynamic*

231 In comparison to the huge number of potentially independent loci in terms of physical linkage
232 reacting to selection, the number of statistically distinguishable evolutionary trajectories was quite
233 limited (46). Despite expected trade-offs and pleiotropy between loci [17], this argued that rather the
234 number of acting selection pressures was limited than the capacity of the population to follow them.
235 That the majority of genes was not affected by selection supported the assumption of unused
236 adaptive potential, but could also be due to small effect size loci that did not reach statistical
237 significance in the present analysis.

238 While the order of magnitude of estimated effectively acting selection pressures is certainly correct,
239 the exact value may not be. On the one hand, the statistically derived clustering threshold was
240 possibly too rigid and therefore distinguished too many clusters, as indicated by the correlation of
241 some very similar AFT clusters (AFT21, AFT22 and AFT17, AFT18, AFT19) to the same environmental
242 variables (Fig. 3). On the other hand, in longer time series other selection pressures could have
243 elicited additional statistically detectable AFCs or split co-varying but distinctive selection pressures
244 into distinguishable trajectories, thus increasing the estimate of acting selective agents. A hint in this
245 direction was the increase from distinguishable 36 AFCs in a preliminary analysis after five temporal
246 samplings (data not shown) to 46 after six. It was therefore possible that the number of selection
247 pressures that the population can effectively follow independently within few generations may be
248 finally limited, but could be different than estimated here. However, general conclusions will require
249 the application of the framework presented here to more populations and different species.

250 Changes in intensity and direction of AFTs were observed frequently. Since these changes could be
251 attributed to the action of selection, this was contrary to expectations derived from the evolution of
252 phenotypic traits [6]. It is difficult to distinguish whether this discrepancy is due to the nature of the
253 measured traits (mainly morphological and life-history phenotypic traits vs. genomic traces of
254 selection), the relatively short monitoring periods (usually a few generations vs. ~30 generations
255 here) or differences in the dynamics of the temporal selection regime (annual or longer living
256 organisms vs. multivoltine species here). The number of haplotypes that showed a continuous
257 increase or decline in allele frequency over the monitored period were relatively limited (AFT4, 724
258 loci), suggesting that adaptation proceeds mainly by short time scale tracking of varying selection
259 pressures and not by slow fixation of alleles.

260 As expected from studies on *Drosophila*, AFT patterns reflecting seasonal variation were identified.
261 AFT38, highly correlated to temperature and AFT40 showed respective patterns, accounting for
262 about 1.2% of all selected haplotypes. This is in the range found for seasonal fluctuations of allele
263 frequencies in *Drosophila* [53,54]. However, compared to the high variability of the remaining AFT

264 patterns, the seasonal variations in temperature, precipitation, resource availability etc. were
265 surprisingly of relatively minor importance.

266 *Environmental associations and functional gene enrichments of AFTs*

267 The number of AFTs in each cluster was highly different, ranging from a few to many thousands. On
268 average, many independent loci were therefore involved in the same adaptive response, which is
269 typical for polygenic traits [55]. However, enrichment analyses in the GO category Biological Process for
270 genes overlapping with haplotypes from respective AFT clusters rarely yielded clearly interpretable
271 patterns. This was probably as much due to the patchy functional annotation of genes in the
272 *C. riparius* genome [30] as to the scant knowledge of the biological processes in which individual genes
273 are involved in different organism [56] s. However, in some instances, hypotheses on the acting
274 selection pressures could be gained from the enrichment analyses. In particular AFT4 suggested a
275 biological interaction [57], perhaps in defence to parasites. Immune genes adaptively tracking the
276 abundance of potential pathogens over seasons were also found in wild *D. melanogaster*
277 populations. Another interesting set of enriched genes associated with porphyrin containing pigment
278 metabolism was found in AFT36. The most important pigments in Chironomids are hemoglobins,
279 giving the characteristic red colour to these midges (also called "bloodworms"). They play an
280 important role in hypoxia/anoxia [58] and are negatively influenced in their function by nitrite [59]. It was
281 thus not surprising that the highest correlations of this AFT were to oxygen ($r = 0.74$) and NO_2 ($r =$
282 0.78), respectively. Less than a quarter of the AFT patterns were highly (> 0.9) correlated with the
283 few measured environmental parameters, indicating that only a small proportion of selectively
284 relevant parameters were measured. Some correlations suggestively matched the inferred functions
285 from the enrichment analysis. AFT39, enriched for genes involved in *nitric-oxide synthase inhibitor*
286 *activity,* was highly correlated to NO_2, underlining the importance of nitrogen metabolism for the
287 species. Another match was found for AFT21, strongly associated with biological oxygen demand,
288 which reflects the availability of organic matter in the stream, an important food source for the
289 midges [39]. Among the enriched GO terms, *carbohydrate metabolism* was found, insinuating an
290 adaptive reaction to the food availability.

291 *Conclusion*

292 The introduced approach of correlating the temporal trajectory of selected haplotypes provided a
293 first glimpse on the extent and dynamics of the multivariate selection regime acting on natural
294 populations and their evolutionary potential. It showed empirically that adaptive tracking of
295 environmental changes is possible at least for *C. riparius*. Species following such an evolutionary
296 regime should be able to cope also with long term trends as imposed by global change. The next
297 logical step will be to increase the temporal sampling and analyse the spatial dimension. This
298 promises to obtain statistically reliable associations whose functional causalities may then be
299 experimentally validated (but see [38]). Beyond these important scientific questions, the presented
300 framework opens the possibility to use population genomic time series for applied functional
301 environmental monitoring. While currently rather the presence or absence of species is used to
302 evaluate the state of environments, the monitoring of selective changes associated with certain
303 anthropogenic or environmental factors could serve as an early warning system before any
304 community change would be visible. Such knowledge may soon be of crucial importance to direct
305 conservation and mitigation strategies in a rapidly changing world.

306 Material and Methods

307 *Sampling*

308 In February 2016, than 500 living red *thummi*-type chironomid larvae were sampled with a sieve at a
309 single site situated in a small river (Hasselbach, Hessen, Germany 50.167562°N, 9.083542°E)
310 following the protocol of [28]. The specimen's specific identity was ascertained by DNA-barcoding of a
311 mitochondrial (COI) and a nuclear locus (L44). Eighty thus identified *C. riparius* were pooled and
312 subjected to PoolSeq sequencing (see below). The sampling was likewise repeated in September
313 2016, assuming that the two annual samples thus were subjected to the selective extremes of the
314 winter and summer season, respectively. This sampling scheme was repeated for the two following
315 years, 2017 and 2018, resulting in a time series of six population pools (NP0-NP5). From February to
316 September, the population produced about seven generations, in the period from autumn to early
317 spring about three. The entire sampling period thus covered about 27 generations (Suppl. Fig. 1). We
318 used additional *C. riparius* individuals from the field, sampled at the same time as the individuals for
319 genomic analysis except for September 2018, to perform standardised life cycle fitness tests at three
320 temperatures (14°C, 20°C and 26°C).

321 *Environmental parameters*

322 Mean monthly water parameters for the period 2009-2018 were obtained from the Hasselbach
323 water treatment plant of the Abwasserverband Freigericht. The values were continuously measured
324 automatically in the river past the outflow, less than 50 m from the sampling site. The following
325 parameters were recorded: water temperature [°C], conductivity [mS/cm], oxygen, biological oxygen
326 demand, chemical oxygen demand, NH_4, NO_3, NO_2, N_{tot} and PO_4, all [mg/l]. To assess the similarity of
327 the entirety of the parameters over time, the monthly means were summarised in a seasonal mean
328 for the period before sampling, *i.e.* from March to September and from October to February. The
329 standardised values (zero means, standard deviation of one) were clustered with UPGMA, using their
330 pairwise correlations as similarity matrix. Temporal autocorrelation was calculated in Past 3
331 {Hammer, 2001 #28016}.

332 *Fitness test of seasonal samples*

333 The life cycle fitness test involved recoding of several biological parameters, including mortality and
334 emergence time (EmT_{50}), the latter being a proxy for speed of development and, hence, the
335 generation length. From these parameters, the population growth rate as inclusive population fitness
336 measure was calculated. A detailed description of the life cycle fitness test can be found in Foucault
337 *et al.* (2018a).

338 *Genomic analyses of pooled samples*

339 DNA was extracted for each pool from the field or the experiment using the Quiagen blood and
340 tissue extraction kit on pooled samples of 80 larval head capsules, respectively. Integrity and quality
341 of extracted DNA was controlled using electrophoresis, and the DNA concentration for each samples
342 measured with a Qubit fluorimeter (Invitrogen).

343 Whole genome pool-sequencing was carried out on an Illumina MySeq with 250bp paired end reads.
344 Reads were trimmed using the wrapper tool autotrim [36] that integrates trimmomatic [60] for trimming
345 and fastQC [61] for quality control. The trimmed reads were then mapped on the latest *C. riparius*
346 reference genome [30] using the BWA mem algorithm [62]. Low quality reads were subsequently filtered
347 and SNPs were initially called using samtools [63]. The pipelines Popoolation and Popoolation2 [64,65]
348 were used to call SNPs, remove indels, and to estimate genetic diversity as Watterson's theta each
349 pool. Allele frequencies for all SNPs with coverage between 15x and 70x were estimated with the R
350 library PoolSeq [43].

351 *Inference of Linkage Disequilibrium*

352 The expected length of segregating haplotypes depends on the recombination rate and their age. The
353 former can be approximated by an estimate of linkage disequilibrium (LD) [52]. We estimated the
354 extent of LD and its decay using the maximum likelihood method of Feder et al. [52] on one of the pools
355 (NP0). This method exploits the haplotype information contained in (paired) reads and has been
356 shown to yield accurate estimates if mean LD decay to background levels is shorter than the read
357 length [52].

358 *Identification of allele frequency trajectories with signs of selection*

359 Selected SNP loci were identified by their allele-frequency change (AFC) larger than expected by
360 neutral drift. First, Fisher's exact tests for the AFCs among all consecutive time points (NP0 > NP1
361 etc.) and among the first and last sampling were calculated. This test identifies AFC larger than
362 expected by chance, but does not account for genetic drift expected. Therefore, neutral simulations
363 were used to compute false discovery rate q-values < 0.001 with parameters (number of SNPs,
364 starting allele frequencies in the ancestral population, sequence coverage, number of generations)
365 matching those of the respective sample. As effective population size, we used 15,000, which
366 constitutes a very conservative estimate (see [36]). All calculations and simulations were performed
367 with the R-library poolSeq [43]. A SNP locus was considered showing signs of selection if at least one
368 sampling interval or between the first and the last sample period had an AFC significantly larger than
369 expected. The allele frequency trajectory (AFT) of such SNPs was recorded, resulting in 367,446 AFTs.

370 *Clustering of spatially coherent SNPs along the scaffolds into haplotypes*

371 The extent of haplotypes on scaffolds was determined following the rationale of [22] that physically
372 linked variants following the same evolutionary trajectory should show correlated AFTs through time.
373 However, contrary to an experimental evolution situation with a constant selection pressure, it could
374 not be expected that the observed allele frequency shifts were always in the same direction. Walking
375 along a scaffold, an AFT was added to a haplotype if the absolute correlation coefficient to any of the
376 SNPs already in the haplotype equalled or exceeded the given threshold. This was repeated until the
377 next SNP did not meet the criterion. The number of ATF in the haplotype, its start and end and the
378 mean AFT were recorded and a new haplotype was started. The analysis was carried out with a
379 custom Python script.

380 *Validation of haplotypes*

381 Inferred haplotypes were validated with 30 whole genome resequenced individuals (25X coverage,
382 150 bp paired end reads) from all sampling periods. The reads from these individuals were quality
383 checked, mapped against the reference genome as described above and the genotypes called with
384 the GATK pipeline. Consensus sequences of one hundred randomly chosen haplotypes with more
385 than 2 SNPs, including one neighbouring haplotype were extracted from the individual VCF-files,
386 aligned and phased with FastPhase in DNAsp {Librado, 2009 #27798}. The resulting phased
387 alignments were manually checked for linkage of the previously identified target haplotype's typical
388 alleles and linkage equilibrium to the neighbouring haplotypes.

389 *Calculation of selection coefficients for haplotypes*

390 The selection coefficient *s* was calculated from the change in mean allele frequencies between
391 samplings for all selected haplotypes by the formula given in [66] Assuming codominance (i.e. h = 0.5),
392 the selection coefficient *s* can be calculated as

$$s = 2 * (\ln\left(\frac{p_t}{1-p_t}\right) - \ln\left(\frac{p_0}{1-p_0}\right)) * t^{-1}$$

where p_0 is the allele frequency of the rising allele in the ancestral population and p_t its frequency in generation t.

Hierarchical clustering of allele frequency trajectories

The matrix of absolute pairwise correlation coefficients among mean AFT per haplotype was transformed into a distance matrix (1 – correlation coefficient), hierarchically clustered with UPGMA and the resulting dendrogram plotted. The number of clusters was determined by applying a clustering threshold of 0.25, corresponding to a minimum similarity of 0.75 of all AFT within a cluster (for choice of clustering threshold, see below).

Determination of the clustering threshold for haplotypes with similar temporal trajectories

To determine a clustering threshold sensibly balancing between arbitrary lumping of actually different trajectories and over-splitting of trajectories differing rather by stochastic fluctuations, we applied a randomisation approach. Starting with a distance threshold of 0.05, we iteratively increased the threshold by 0.05. In each iteration, we randomly drew from the empirical AFC distribution of all SNPs in the analysed set (i.e. at least one significant AFC among consecutive sampling periods or between the first and last sampling period, see above), 1,000 times the empirically found number of trajectories, satisfying the same conditions. For each of the 1,000 replicates, the trajectories were hierarchically clustered as described above and the number of resulting clusters recorded. The simulated distribution was then compared to the empirically observed number of clusters, applying the same threshold. The smallest distance threshold for which the observed number of clusters fell below the 0.05 quantile of the randomly simulated distribution was deemed appropriate for the analysis. This was the case for a distance threshold of 0.25; i.e. the trajectories within each cluster showed a correlation of at least 0.75. The analysis was performed with a custom Python script.

GO-term enrichment analysis

The annotated gene overlapping with a haplotype was identified as putative selection target. The resulting gene lists for all member haplotypes of each mean AFT pattern were tested for over-representation of GO terms in the category “biological function” in the R-library TopGO with a Fisher’s exact test [67].

Correlation of environmental parameters and mean AFTs

The seasonal mean environmental parameters (see above) before the first and the respective period between two sampling times were correlated to the means of all AFT patterns. Given the obviously lacking statistical power, this proceeding should be rather viewed as an attempt to generate hypotheses.

Author Contributions

Conceptualization MP, Methodology Development MP, Software Programming MP,Validation MP, Formal Analysis MP, Investigation QF, MP, Resources MP, Data Curation MP, Writing – Original Draft Preparation MP, Writing – Review & Editing Preparation QF, MP, Visualization Preparation MP, Supervision MP, Project Administration Management MP, Funding Acquisition MP.

Acknowledgements

We thank Andreas Wieser for help with sampling and the Abwasserverband Freigericht for generously providing the environmental Hasselbach data. The work was funded by DFG (grant PF390/8-1).

Declaration of interests

The authors declare no conflict of interest.

Data and Code Availability

Sequencing data is publicy available on ENA (project ERP115516, samples ERS4040036-ERS4040041). Core scripts and sync-files were made available on DRYAD (doi.org/10.5061/dryad.fxpnvx0nd).

References

1 Hoekstra, H. E. *et al.* Strength and tempo of directional selection in the wild. *Proceedings of the National Academy of Sciences* **98**, 9157-9160 (2001).

2 Kingsolver, J. G. *et al.* The strength of phenotypic selection in natural populations. *The American Naturalist* **157**, 245-261 (2001).

3 Bell, G. Fluctuating selection: the perpetual renewal of adaptation in variable environments. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 87-97 (2010).

4 Price, T. D., Grant, P. R., Gibbs, H. L. & Boag, P. T. Recurrent patterns of natural selection in a population of Darwin's finches. *Nature* **309**, 787-789 (1984).

5 Siepielski, A. M., DiBattista, J. D., Evans, J. A. & Carlson, S. M. Differences in the temporal dynamics of phenotypic selection among fitness components in the wild. *Proceedings of the Royal Society B: Biological Sciences* **278**, 1572-1580 (2011).

6 Kingsolver, J. G., Diamond, S. E., Siepielski, A. M. & Carlson, S. M. Synthetic analyses of phenotypic selection in natural populations: lessons, limitations and future directions. *Evolutionary Ecology* **26**, 1101-1118 (2012).

7 Kingsolver, J. G. & Pfennig, D. W. Patterns and power of phenotypic selection in nature. *Bioscience* **57**, 561-572 (2007).

8 Siepielski, A. M., DiBattista, J. D. & Carlson, S. M. It's about time: the temporal dynamics of phenotypic selection in the wild. *Ecology Letters* **12**, 1261-1276 (2009).

9 Messer, P. W. & Petrov, D. A. Population genomics of rapid adaptation by soft selective sweeps. *Trends in Ecology and Evolution* **28**, 659-669 (2013).

10 Stephan, W. Signatures of positive selection: from selective sweeps at individual loci to subtle allele frequency changes in polygenic adaptation. *Molecular Ecology* **25**, 79-88 (2016).

11 Matuszewski, S., Hermisson, J. & Kopp, M. Catch me if you can: adaptation from standing genetic variation to a moving phenotypic optimum. *Genetics* **200**, 1255-1274 (2015).

12 Kopp, M. & Matuszewski, S. Rapid evolution of quantitative traits: theoretical perspectives. *Evolutionary Applications* **7**, 169-191 (2014).

13 Barton, N. & Partridge, L. Limits to natural selection. *Bioessays* **22**, 1075-1084 (2000).

14 Walsh, B. & Blows, M. W. Abundant genetic variation+ strong selection= multivariate genetic constraints: a geometric view of adaptation. *Annual Review of Ecology, Evolution, and Systematics* **40**, 41-59 (2009).

15 Betancourt, A. J. & Presgraves, D. C. Linkage limits the power of natural selection in Drosophila. *Proceedings of the National Academy of Sciences* **99**, 13616-13620 (2002).

16 Gomulkiewicz, R. & Houle, D. Demographic and genetic constraints on evolution. *The American Naturalist* **174**, E218-E229 (2009).

17 Li, Y., Petrov, D. A. & Sherlock, G. Single nucleotide mapping of trait space reveals Pareto fronts that constrain adaptation. *Nature Ecology & Evolution*, 1-13 (2019).

18 Johnson, T. & Barton, N. Theoretical models of selection and mutation on quantitative traits. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360**, 1411-1425 (2005).

19 Lynch, M. A quantitative-genetic perspective on conservation issues. *Conservation Genetics: case histories from nature*, 471-501 (1995).
20 Waldvogel, A. M. *et al.* Evolutionary genomics can improve prediction of species' responses to climate change. *Evolution Letters* **4**, 4-18 (2020).
21 Pigliucci, M. Genetic variance–covariance matrices: a critique of the evolutionary quantitative genetics research program. *Biology and Philosophy* **21**, 1-23 (2006).
22 Franssen, S. U., Barton, N. H. & Schlötterer, C. Reconstruction of haplotype-blocks selected during experimental evolution. *Molecular Biology and Evolution*, **34**, 174-184 (2016).
23 Barton, N. H. Clines in polygenic traits. *Genetics Research* **74**, 223-236 (1999).
24 Barghi, N. *et al.* Genetic redundancy fuels polygenic adaptation in *Drosophila*. *PLoS Biology* **17**, e3000128 (2019).
25 Messer, P. W., Ellner, S. P. & Hairston Jr, N. G. Can population genetics adapt to rapid evolution? *Trends in Genetics* **32**, 408-418 (2016).
26 Armitage, P. D. in *The Chironomidae: The biology and ecology of non-biting midges* (eds P. D. Armitage, P. S. Cranston, & L. C. V. Pinder) 194-224 (Chapman & Hall, 1995).
27 Pfenninger, M. & Nowak, C. Reproductive isolation and ecological niche partition among larvae of the morphologically cryptic sister species *Chironomus riparius* and *C. piger*. *PLoS One* **3** (2008).
28 Foucault, Q., Wieser, A., Waldvogel, A. M. & Pfenninger, M. Establishing laboratory cultures and performing ecological and evolutionary experiments with the emerging model species Chironomus riparius. *Journal of Applied Entomology* **143**, 584-592 (2019).
29 Oppold, A. M. & Pfenninger, M. Direct estimation of the spontaneous mutation rate by short-term mutation accumulation lines in *Chironomus riparius*. *Evolution Letters* **1**, 86-92 (2017).
30 Schmidt, H. *et al.* A high-quality genome assembly from short and long reads for the non-biting midge *Chironomus riparius* (Diptera). *G3: Genes, Genomes, Genetics* **10**, 1151-1157 (2020).
31 Kraak, M. H., Stuijfzand, S. C. & Admiraal, W. Interaction between organic matter and toxicants in polluted river water: beneficial and inhibiting effects on the midge *Chironomus riparius*. (2000).
32 Nemec, S., Heß, M., Nowak, C. & Pfenninger, M. Experimental evidence for niche segregation in a sister species pair of non-biting midges. *Hydrobiologia* **691**, 203-212 (2012).
33 Nemec, S., Patel, S., Nowak, C. & Pfenninger, M. Evolutionary determinants of population differences in population growth rate× habitat temperature interactions in *Chironomus riparius*. *Oecologia* **172**, 585-594 (2013).
34 Nowak, C. *et al.* Rapid genetic erosion in pollutant-exposed experimental chironomid populations. *Environmental Pollution* **157**, 881-886 (2009).
35 Oppold, A.-M. *et al.* Support for the evolutionary speed hypothesis from intraspecific population genetic data in the non-biting midge *Chironomus riparius*. *Proceedings of the Royal Society B: Biological Sciences* **283**, 20152413 (2016).
36 Waldvogel, A. M. *et al.* The genomic footprint of climate adaptation in *Chironomus riparius*. *Molecular Ecology* **27**, 1439-1456 (2018).
37 Foucault, Q., Wieser, A., Waldvogel, A. M., Feldmeyer, B. & Pfenninger, M. Rapid adaptation to high temperatures in *Chironomus riparius*. *Ecology and Evolution* **8**, 12780-12789 (2018).
38 Pfenninger, M. & Foucault, Q. Genomic processes underlying rapid adaptation of a natural *Chironomus riparius* population to unintendedly applied experimental selection pressures. *Molecular Ecology* **29**, 536-548 (2020).
39 Ristola, T., Pellinen, J., Ruokolainen, M., Kostamo, A. & Kukkonen, J. V. Effect of sediment type, feeding level, and larval density on growth and development of a midge (Chironomus riparius). *Environmental Toxicology and Chemistry: An International Journal* **18**, 756-764 (1999).
40 Correia, A. M., Bandeira, N. & Anastácio, P. M. Predator–prey interactions of Procambarus clarkii with aquatic macroinvertebrates in single and multiple prey systems. *Acta Oecologica* **28**, 337-343 (2005).

41 Hooper, H. L., Sibly, R. M., Hutchinson, T. H. & Maund, S. J. The influence of larval density, food availability and habitat longevity on the life history and population growth rate of the midge *Chironomus riparius*. *Oikos* **102**, 515-524 (2003).

42 van de Bund, W. J. & Davids, C. Complex relations between freshwater macro-and meiobenthos: interactions between *Chironomus riparius* and *Chydorus piger*. *Freshwater Biology* **29**, 1-6 (1993).

43 Taus, T., Futschik, A. & Schlötterer, C. Quantifying selection with pool-seq time series data. *Molecular Biology and Evolution* **34**, 3023-3034 (2017).

44 Foucault, Q. *et al.* An experimental assessment of reproductive isolation and its consequences for seasonal hybridization dynamics. *Biol. J. Linnean Soc.* **126**, 327-337 (2019).

45 Dayan, D. I. *et al.* Population genomics of rapid evolution in natural populations: polygenic selection in response to power station thermal effluents. *BMC Evolutionary Biology* **19**, 61 (2019).

46 Torda, G. *et al.* Rapid adaptive responses to climate change in corals. *Nature Climate Change* **7**, 627-636 (2017).

47 Hedrick, P. W. Genetic polymorphism in heterogeneous environments: the age of genomics. *Annual Reviews in Ecology Evolution and Systematics* **37**, 67-93 (2006).

48 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).

49 Wittmann, M. J., Bergland, A. O., Feldman, M. W., Schmidt, P. S. & Petrov, D. A. Seasonally fluctuating selection can maintain polymorphism at many loci via segregation lift. *Proceedings of the National Academy of Sciences* **114**, E9932-E9941 (2017).

50 Bertram, J. & Masel, J. Different mechanisms drive the maintenance of polymorphism at loci subject to strong versus weak fluctuating selection. *Evolution* **73**, 883-896 (2019).

51 Kain, J. S. *et al.* Variability in thermal and phototactic preferences in *Drosophila* may reflect an adaptive bet-hedging strategy. *Evolution* **69**, 3171-3185 (2015).

52 Feder, A. F., Petrov, D. A. & Bergland, A. O. LDx: estimation of linkage disequilibrium from high-throughput pooled resequencing data. *PloS One* **7** (2012).

53 Bergland, A. O., Behrman, E. L., O'Brien, K. R., Schmidt, P. S. & Petrov, D. A. Genomic evidence of rapid and stable adaptive oscillations over seasonal time scales in Drosophila. *PLoS Genetics* **10** (2014).

54 Machado, H. *et al.* Broad geographic sampling reveals predictable and pervasive seasonal adaptation in Drosophila. *BioRxiv*, 337543 (2018).

55 Jain, K. & Stephan, W. Modes of rapid polygenic adaptation. *Molecular Biology and Evolution* **34**, 3169-3175 (2017).

56 Leale, G., Bayá, A. E., Milone, D. H., Granitto, P. M. & Stegmayer, G. Inferring unknown biological function by integration of GO annotations and gene expression data. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* **15**, 168-180 (2016).

57 Behrman, E. L. *et al.* Rapid seasonal evolution in innate immunity of wild *Drosophila melanogaster*. *Proceedings of the Royal Society B: Biological Sciences* **285**, 20172599 (2018).

58 Grazioli, V., Rossaro, B., Parenti, P., Giacchini, R. & Lencioni, V. Hypoxia and anoxia effects on alcohol dehydrogenase activity and hemoglobin content in *Chironomus riparius* Meigen, 1804. *J. Limnol* **75**, 347-354 (2016).

59 Ha, M.-H. & Choi, J. Effects of environmental contaminants on hemoglobin of larvae of aquatic midge, *Chironomus riparius* (Diptera: Chironomidae): A potential biomarker for ecotoxicity monitoring. *Chemosphere* **71**, 1928-1936 (2008).

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

61 Andrews, S. (Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom, 2010).

62 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics* **25**, 1754-1760 (2009).

586 63 Li, H. *et al.* The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics* **25**,
587 2078-2079 (2009).
588 64 Kofler, R. *et al.* PoPoolation: A Toolbox for Population Genetic Analysis of Next Generation
589 Sequencing Data from Pooled Individuals. *PLoS One* **6**, e15925 (2011).
590 65 Kofler, R., Pandey, R. V. & Schlötterer, C. PoPoolation2: identifying differentiation between
591 populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics* **27**, 3435-
592 3436 (2011).
593 66 Crow, J. F. & Kimura, M. *An introduction to population genetics theory*. (Harper and Row,
594 1970).
595 67 Alexa, A. & Rahnenführer, J. Gene set enrichment analysis with topGO. *Bioconductor*
596 *Improvment* **27** (2009).

597

598

Figures

Figure 1. Population growth rate as fitness measure of natural population when exposed to different experimental temperatures. The differences between sampling dates are significant (F = 9.509, p = 5.07 x 10^{-6}) as well as the interaction between sampling date and experimental temperature (F = 3.631, p = 0.0017).

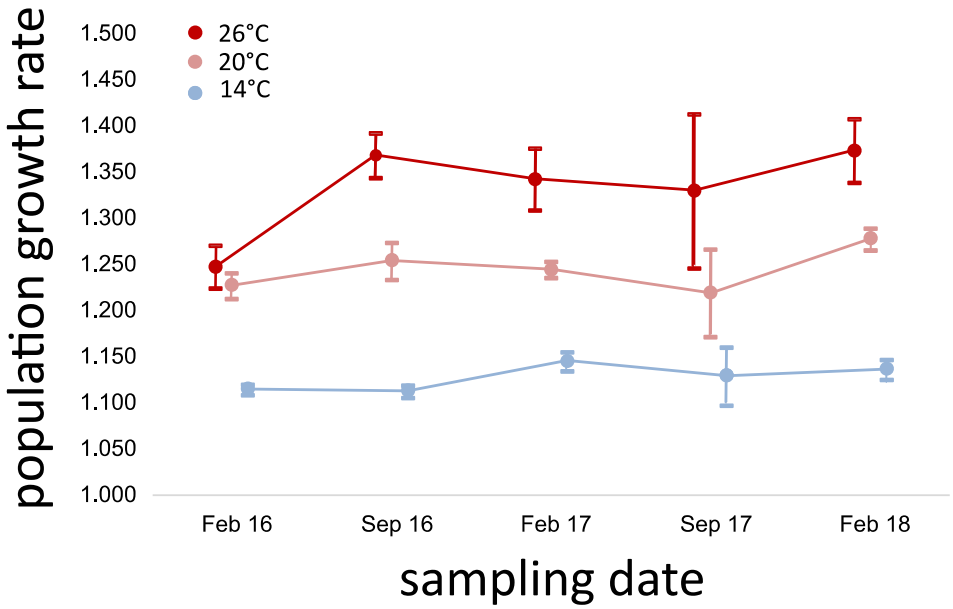


Figure 2. Temporal dynamics and complexity of allele frequencies trajectories with signs of selection. The left side of the panel shows a UPGMA-dendrogram on the pairwise distance (1 - correlation coefficients) of all 35,574 haplotype AFTs with at least one significant AFC. The hatched vertical line presents the inferred clustering threshold of 0.25. The right side of the panel shows the AFTs for each cluster. For better visibility and comparability, the individual AFTs in each cluster were polarised into the same direction and normalised by substracting the respective mean.

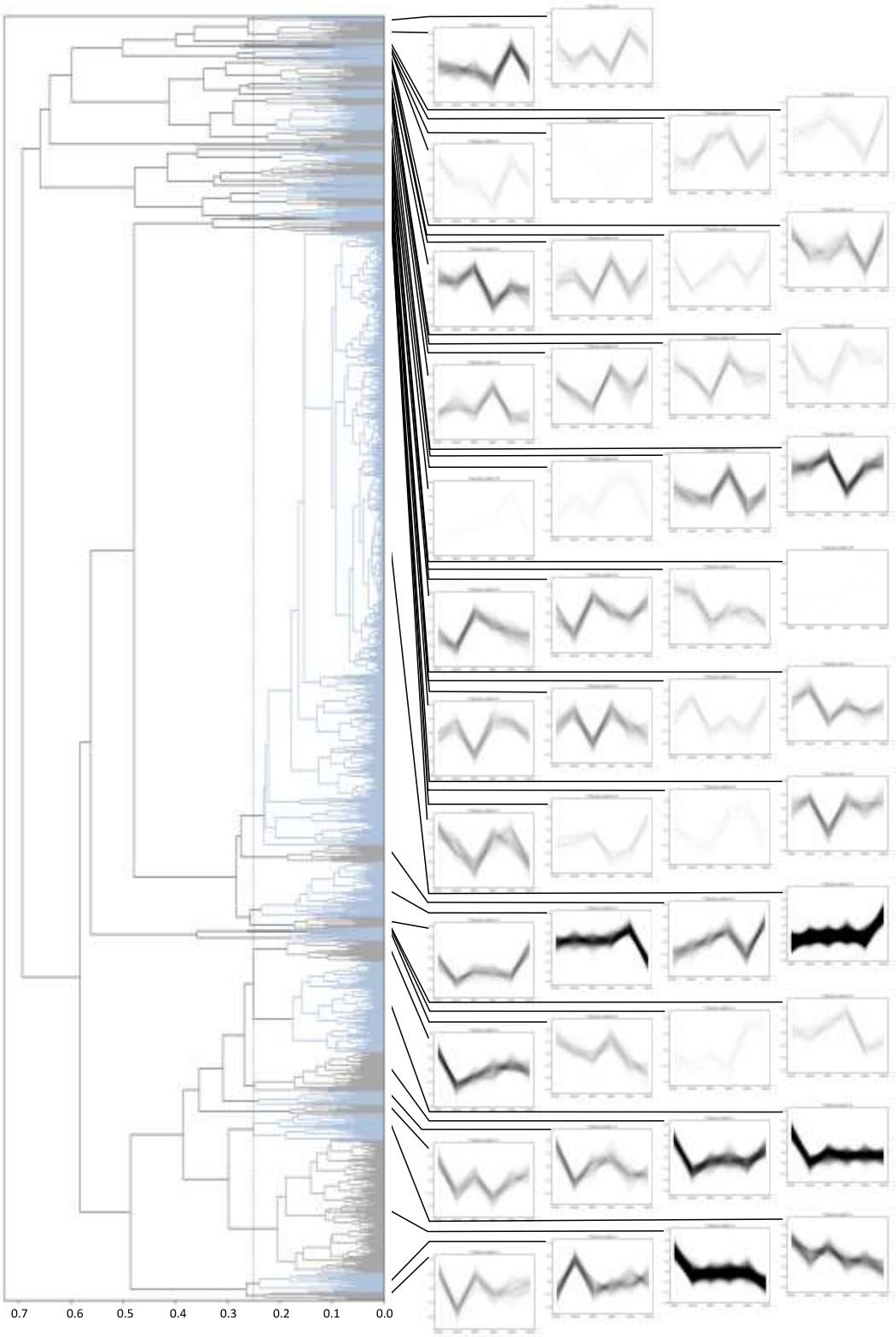
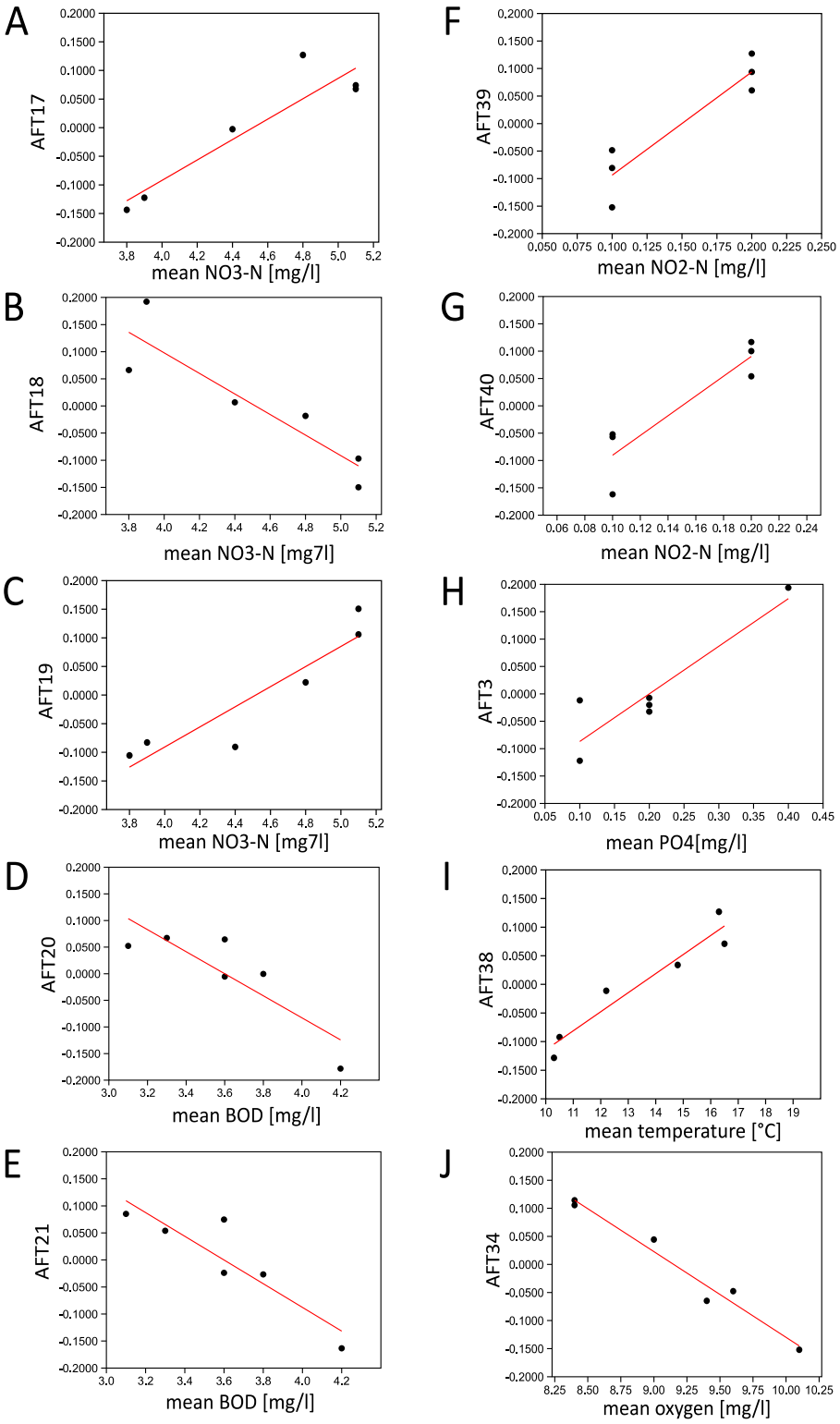


Fig 3. Correlations between measured environmental parameters and AFTs. A: AFT19 and NO_3, B: AFT17 and NO_3, AFT39 and NO_2, AFT38 and temperature, AFT34 and oxygen. Please note that the direction of the correlation depends on the arbitrary polarisation of the alleles in the allele frequency trajectory.



SUPPLEMENT

SUPPL. RESULTS

Table 1. Results of two way ANOVA on fitness data of the field population to three experimental temperatures (14°C, 20°C, 26°C) at 5 different sampling dates (Feb. 16, Sept. 16, Feb. 17, Sep. 17, Feb. 18).

Factor	Sum of sqrs	d.f.	Mean square	F	p (same)
Sampling date	0.037	4	0.009	9.509	5.07E-06
Exp. Temp	0.527	2	0.263	269.6	1.05E-30
Interaction	0.028	8	0.004	3.631	0.001668
Within	0.059	60	0.001		
Total	0.651	74			

Table 2. Frequency distribution of AFT cluster patterns, environmental parameter with correlation coefficient > 0.9, number of genes affected and first four most significantly enriched GFO terms for the 46 observed AFT clusters.

Cluster ID	**N AFT**	**Env. Parameter r r > 0.9**	**N genes**	**Significant GO Terms**			
1	244	-	72	0045454, cell redox homeostasis, 41, 2, 0.28, 0.031	0006606, protein import into nucleus, 6, 1, 0.04, 0.040	0006355, regulation of transcription, DNA-templat..., 322, 6, 2.2, 0.040	0007165, signal transduction, 397, 8, 2.71, 0.046
2	521	-	152	0007205, protein kinase C-activating G protein-co..., 5, 2, 0.06, 0.0016	0007017, microtubule-based process, 64, 4, 0.82, 0.0163	0007264, small GTPase mediated signal transductio..., 55, 3, 0.7, 0.0233	
3	3943	PO_4	1126	0006811, ion transport, 300, 33, 22.94, 2.7e-05	0007156, homophilic cell adhesion via plasma memb..., 25, 7, 1.91, 0.0021	0006836, neurotransmitter transport, 14, 5, 1.07, 0.0028	0042737, drug catabolic process, 11, 4, 0.84, 0.0072
4	742	-	262	0051704, multi-organism process, 7, 2, 0.14, 0.0081	0006952, defense response, 9, 2, 0.18, 0.0136	0055085, transmembrane transport, 439, 16, 8.99, 0.0226	0006816, calcium ion transport, 17, 2, 0.22, 0.0372
5	289	-	94	0006355, regulation of transcription, DNA-templat..., 322, 7, 2.12, 0.012	0007268, chemical synaptic transmission, 6, 1, 0.04, 0.039		
6	352	-	111	0055114, oxidation-reduction process, 433, 10, 4.14, 0.0099	0009116, nucleoside metabolic process, 23, 2, 0.22, 0.0199	0072522, purine-containing compound biosynthetic ..., 58, 2, 0.55, 0.0373	0006120, mitochondrial electron transport, NADH t..., 5, 1, 0.05, 0.0469
7	1183	-	354	0035556, intracellular signal transduction, 132, 11, 3.54, 0.0030	0007275, multicellular organism development,	0006836, neurotransmitter transport, 14, 3, 0.38, 0.0055	0042246, tissue regeneration, 5, 2, 0.13,

					25, 4, 0.67, 0.0040		0.0068
8	2611	-	816	0015698, inorganic anion transport, 15, 5, 0.93, 0.00043	0006811, ion transport, 300, 32, 18.57, 0.00378	0055085, transmembrane transport, 439, 41, 27.17, 0.01175	0006629, lipid metabolic process, 189, 13, 11.7, 0.02091
9	660	-	215	0007156, homophilic cell adhesion via plasma memb..., 25, 4, 0.44, 0.00087	0034220, ion transmembran e transport, 88, 4, 1.56, 0.01161	3, 0009072, aromatic amino acid family metabolic pro..., 13, 2, 0.23, 0.02137	0055085, transmembran e transport, 439, 16, 7.79, 0.02999
10	208	-	61	0007264, small GTPase mediated signal transductio..., 55, 2, 0.25, 0.025	0006207, 'de novo' pyrimidine nucleobase biosynth..., 6, 1, 0.03, 0.027		
11	29	-	13	-			
12	102	-	24	0032269, negative regulation of cellular protein ..., 6, 1, 0.01, 0.012	0006417, regulation of translation, 9, 1, 0.02, 0.018	0006413, translational initiation, 13, 1, 0.03, 0.026	2000113, negative regulation of cellular macromol..., 14, 1, 0.03, 0.028
13	355	-	136	0007165, signal transduction, 397, 14, 4.61, 0.00043			
14	1580	-	569	0007156, homophilic cell adhesion via plasma memb..., 25, 6, 1.15, 0.00076	0007155, cell adhesion, 49, 11, 2.26, 0.00373	0035556, intracellular signal transduction, 132, 10, 6.1, 0.02015	0006012, galactose metabolic process, 6, 2, 0.28, 0.02816
15	514	-	175	0007156, homophilic cell adhesion via plasma memb..., 25, 3, 0.4, 0.0069	0006979, response to oxidative stress, 22, 2, 0.35, 0.0471		
16	17914	-	4831	0055085, transmembrane transport, 439, 202, 159.22, 1.5e-05	0007166, cell surface receptor signaling pathway, 53, 32, 19.22, 9.9e-05	0007156, homophilic cell adhesion via plasma memb..., 25, 18, 9.07, 0.00029	0006811, ion transport, 300, 139, 108.81, 0.00059
17	333	NO_3	106	0006813, potassium ion transport, 46, 5, 0.33, 0.0016	0071805, potassium ion transmembran e transport, 11, 2, 0.08, 0.0027	0051260, protein homoolimerizatio n, 14, 2, 0.1, 0.0044	0010646, regulation of cell communicatio n, 53, 3, 0.39, 0.0273
18	66	NO_3	19	0006364, rRNA processing, 20, 1, 0.02, 0.023			
19	25	NO_3	6	0009072, aromatic amino acid family metabolic pro..., 13, 1, 0.01, 0.0088			

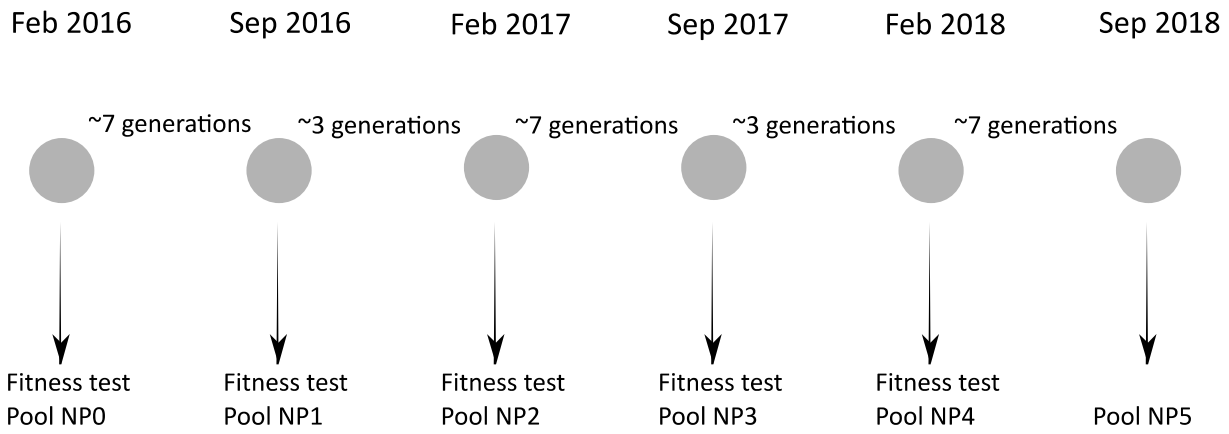
20	340	-	100	0007156, homophilic cell adhesion via plasma memb..., 25, 3, 0.19, 0.00079	0006486, protein glycosylation, 36, 2, 0.27, 0.02938	0006072, glycerol-3-phosphate metabolic process, 5, 1, 0.04, 0.03700	
21	242	BOD	83	0005975, carbohydrate metabolic process, 172, 4, 1.21, 0.015			
22	324	BOD	110	0007608, sensory perception of smell, 12, 2, 0.11, 0.0055	0006979, response to oxidative stress, 22, 2, 0.21, 0.0182	0006368, transcription elongation from RNA polyme..., 5, 1, 0.05, 0.0469	
23	67	-	24	0006955, immune response, 6, 1, 0.01, 0.012	0051604, protein maturation, 20, 1, 0.04, 0.040	0018205, peptidyl-lysine modification, 22, 1, 0.05, 0.044	
24	253	-	91	0006979, response to oxidative stress, 22, 2, 0.16, 0.011	0051259, protein complex olimerization, 16, 2, 0.12, 0.014	0055114, oxidation-reduction process, 433, 7, 3.15, 0.033	0005975, carbohydrate metabolic process, 172, 4, 1.25, 0.035
25	392	-	128	0042325, regulation of phosphorylation, 7, 2, 0.07, 0.0093	0006811, ion transport, 300, 8, 2.87, 0.0113	0043086, negative regulation of catalytic activit..., 5, 1, 0.05, 0.0469	0007205, protein kinase C-activating G protein-co..., 5, 1, 0.05, 0.0469
26	325	-	111	0044262, cellular carbohydrate metabolic process, 20, 2, 0.21, 0.018	0007264, small GTPase mediated signal transductio..., 55, 3, 0.58, 0.019	0034220, ion transmembrane transport, 88, 3, 0.92, 0.031	
27	197	-	57	0009395, phospholipid catabolic process, 6, 1, 0.02, 0.023	0007179, transforming growth factor beta receptor..., 7, 1, 0.03, 0.027		
28	3	-	1	-			
29	8	-	4	0007264, small GTPase mediated signal transductio..., 55, 1, 0.03, 0.025			
30	29	-	8	0032012, regulation of ARF protein signal transdu..., 7, 1, 0, 0.0032			
31	406	-	125	0042133, neurotransmitter metabolic process, 8, 2, 0.08, 0.0024	0007205, protein kinase C-activating G protein-co..., 5, 1, 0.05, 0.0469		

32	697	-	206	0006596, polyamine biosynthetic process, 9, 2, 0.12, 0.0064	0006508, proteolysis, 618, 16, 8.58, 0.0076	0007165, signal transduction, 397, 11, 5.51, 0.0088	0018393, internal peptidyl-lysine acetylation, 9, 2, 0.12, 0.0137
33	230	-	67	0006470, protein dephosphorylatio n, 28, 2, 0.16, 0.011	0007018, microtubule-based movement, 39, 2, 0.22, 0.020	0055085, transmembrane transport, 439, 7, 2.5, 0.035	0034030, ribonucleoside bisphosphate biosynthetic..., 7, 1, 0.04, 0.039
34	222	O_2	81	0007275, multicellular organism development, 25, 2, 0.18, 0.013	1901642, nucleoside transmembran e transport, 5, 1, 0.04, 0.035	0007169, transmembrane receptor protein tyrosine ..., 6, 1, 0.04, 0.042	0046173, polyol biosynthetic process, 6, 1, 0.04, 0.042
35	151	-	44	0007205, protein kinase C-activating G protein-co..., 5, 1, 0.02, 0.020	0005978, glycogen biosynthetic process, 5, 1, 0.02, 0.020	0006207, 'de novo' pyrimidine nucleobase biosynth..., 6, 1, 0.02, 0.024	0006013, mannose metabolic process, 6, 1, 0.02, 0.024
36	71	-	20	0006779, porphyrin-containing compound biosynthet..., 5, 1, 0.01, 0.0091	0046148, pigment biosynthetic process, 7, 1, 0.01, 0.0127	0006629, lipid metabolic process, 189, 2, 0.34, 0.0434	
37	504	-	157	0006979, response to oxidative stress, 22, 4, 0.32, 0.00025	0055114, oxidation-reduction process, 433, 13, 6.31, 0.00833	0006928, movement of cell or subcellular componen..., 41, 3, 0.6, 0.02827	0006813, potassium ion transport, 46, 3, 0.67, 0.02866
38	187	Temp.	60	0007179, transforming growth factor beta receptor..., 7, 1, 0.04, 0.035	0000226, microtubule cytoskeleton organization, 9, 1, 0.05, 0.044		
39	52	NO_2	19	0070588, calcium ion transmembrane transport, 14, 1, 0.01, 0.0095			
40	269	NO_2	90	0007186, G protein-coupled receptor signaling pat..., 137, 4, 0.9, 0.012	0006355, regulation of transcription, DNA-templat..., 322, 6, 2.12, 0.017	1901362, organic cyclic compound biosynthetic pro..., 533, 8, 3.52, 0.027	4, 0008202, steroid metabolic process, 6, 1, 0.04, 0.039
41	52	-	17	0006368, transcription elongation from RNA polyme..., 5, 1, 0, 0.0045	0005975, carbohydrate metabolic process, 172, 2, 0.16, 0.0087	0016570, histone modification, 21, 1, 0.02, 0.0190	0007186, G protein-coupled receptor signaling pat..., 137, 1, 0.12, 0.1190
42	5	-	-	0006368, transcription elongation from RNA polyme..., 5,	0005975, carbohydrate metabolic process, 172,	0016570, histone modification, 21, 1, 0.02, 0.0190	0007186, G protein-coupled receptor

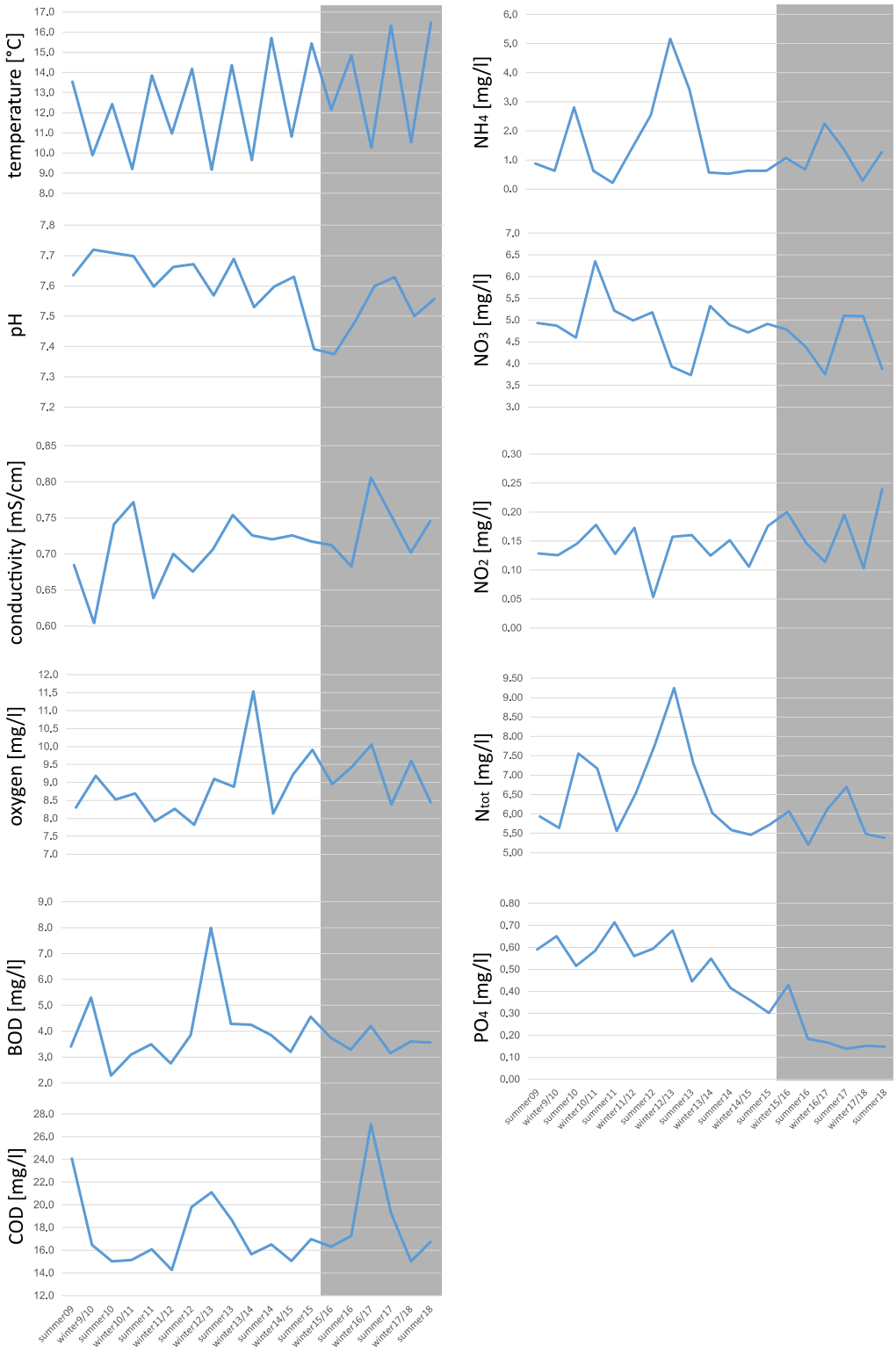
				1, 0, 0.0045	2, 0.16, 0.0087		signaling pat..., 137, 1, 0.12, 0.1190
43	120	-	45	0006511, ubiquitin-dependent protein catabolic pr..., 52, 2, 0.19, 0.015	0007268, chemical synaptic transmission, 6, 1, 0.02, 0.022	0006506, GPI anchor biosynthetic process, 8, 1, 0.03, 0.029	
44	63	-	19	0008272, sulfate transport, 8, 1, 0.01, 0.011	0016055, Wnt signaling pathway, 10, 1, 0.01, 0.014		
45	556	-	198	0006468, protein phosphorylation, 215, 7, 3.23, 0.041			
46	138	-	46	0006012, galactose metabolic process, 6, 1, 0.03, 0.026	0018342, protein prenylation, 6, 1, 0.03, 0.026	0045892, negative regulation of transcription, DN..., 9, 1, 0.04, 0.038	

629

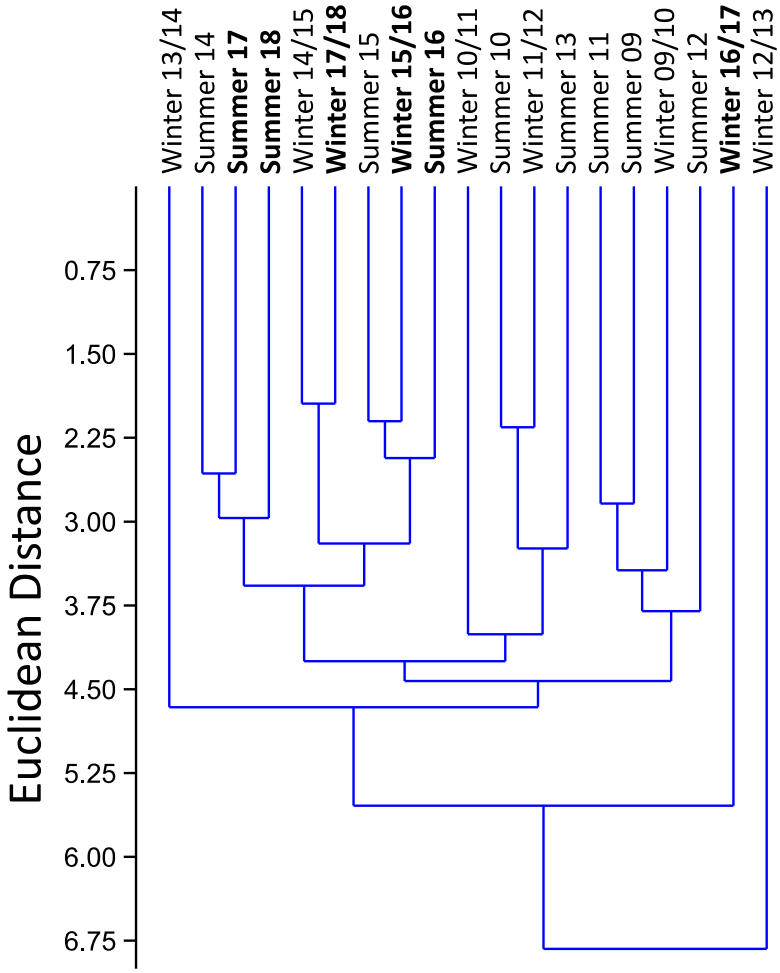
Suppl. Figure 1. Experimental design. The grey circles indicate the sampling periods.



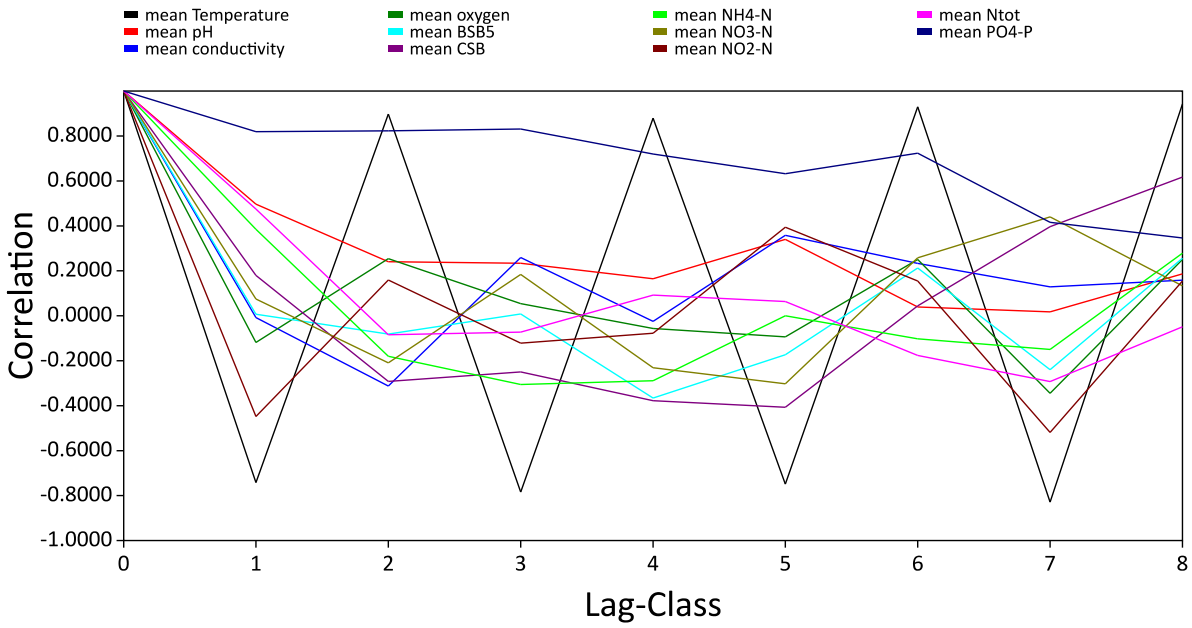
Suppl. Fig. 2. Temporal changes in mean seasonal environmental parameters in the period from 2009-2018 in the Hasselbach. The period relevant for the population genomic analyses is highlighted in grey.



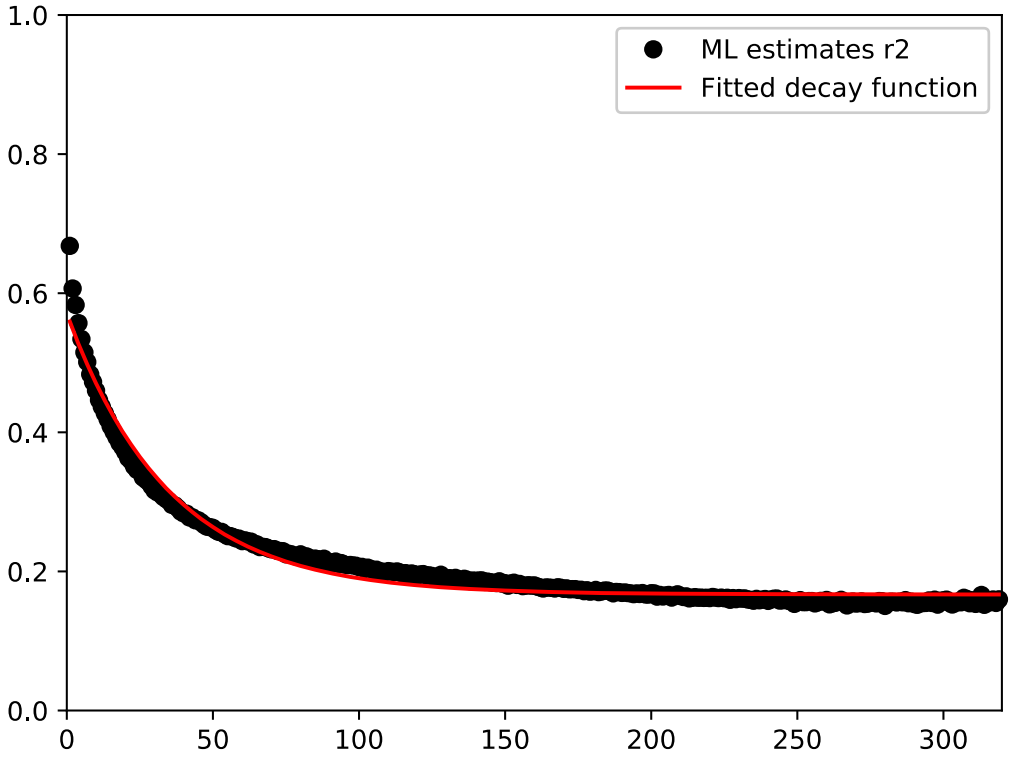
Suppl. Figure 3. UPGMA dendrogram of Euclidean distance matrix from standardised seasonal mean water parameters in the Hasselbach at the sampling site for the period 2009-2018. The seasons relevant for the population genomic sampling are indicated in bold.



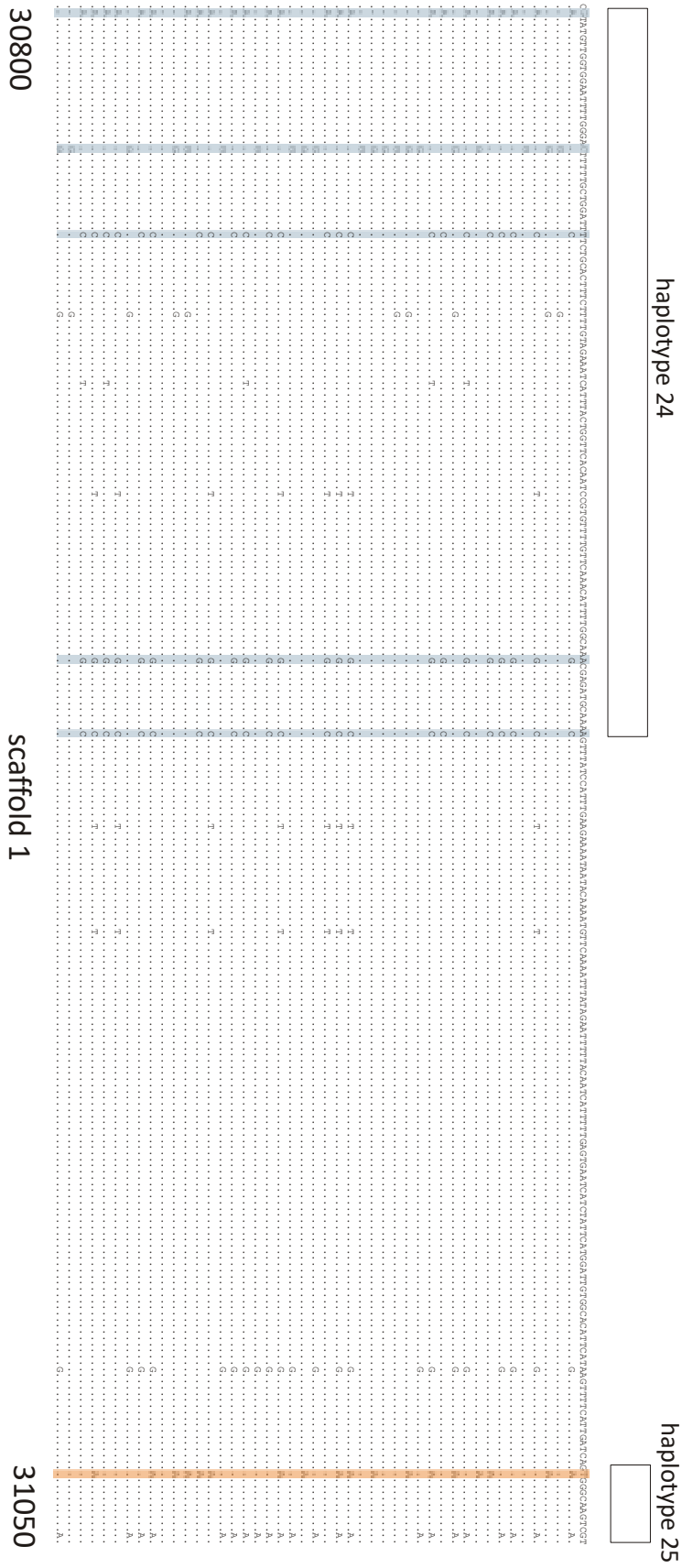
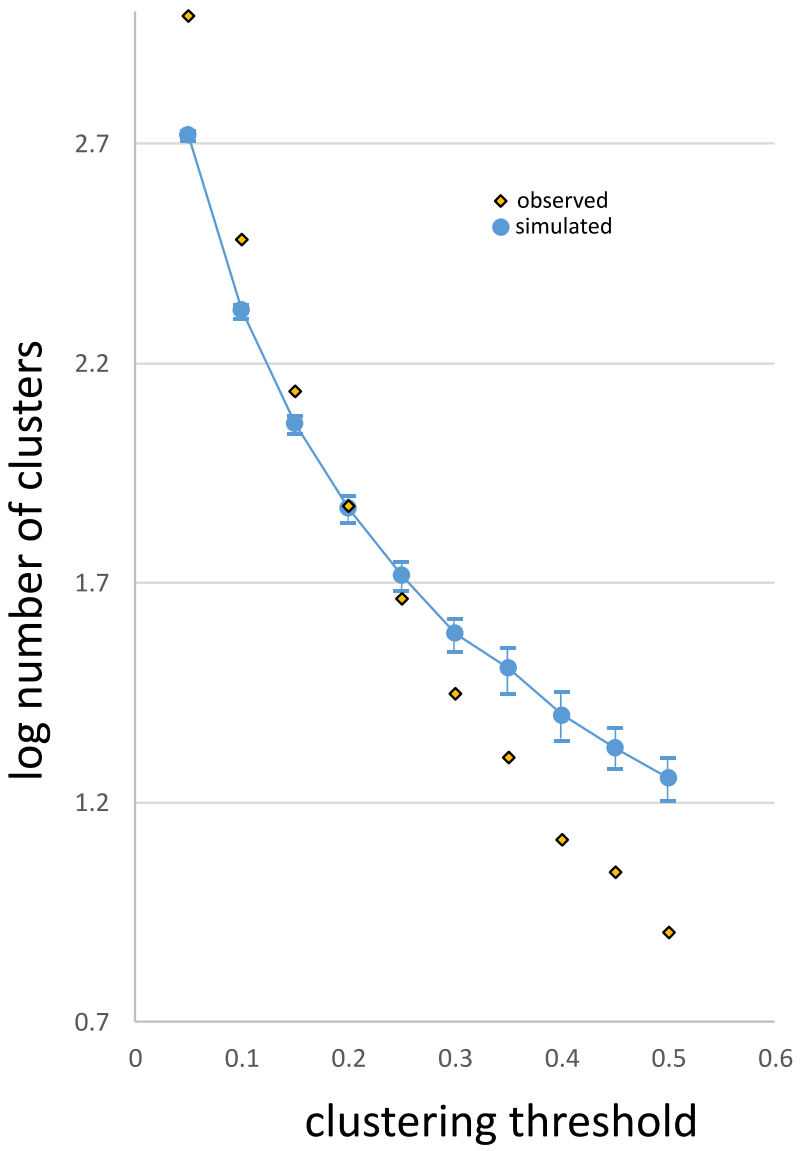
Suppl. Fig 4. Temporal autocorrelations of environmental parameters.



Suppl. Fig 5. Genome wide mean decay of linkage disequilibrium (r^2) in dependence of distance in basepairs.



Suppl. Fig. 6. Example for inferred haplotypes of 30 re-sequenced individuals for haplotype 24 and the first SNP of haplotype 25 on scaffold 1 (positions 30800 – 31050). The marker SNPs for haplotype 24 are indicated in blue, for haplotype 25 in orange. The SNPs of haplotype 24 are in full linkage and unlinked to the SNP of haplotype 25.



Suppl. Fig. 7. Determination of clustering threshold. Plot of clustering threshold against log number of clusters. In blue, the simulated values with 95% confidence intervals, in yellow with black outline the observed value.