Selection on growth rate and local adaptation drive genomic adaptation during experimental range expansions in the protist *Tetrahymena thermophila*

Felix Moerman[1,2,3,4,*], Emanuel A. Fronhofer[4],
Florian Altermatt[1,2] Andreas Wagner[1,3,5,*]

1. Department of Evolutionary Biology and Environmental Studies, University of Zürich, Winterthurerstrasse 190, Zürich CH-8057, Switzerland
2. Department of Aquatic Ecology, Eawag: Swiss Federal Institute of Aquatic Science and Technology, Überlandstrasse 133, Dübendorf CH-8600, Switzerland
3. Swiss Institute of Bioinformatics, Quartier Sorge—Bâtiment Génopode, Lausanne 1015, Switzerland
4. ISEM, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France
5. The Santa Fe Institute, Santa Fe, New Mexico 87501, USA

* Corresponding authors

Keywords:range expansions; whole genome resequencing; Tetrahymena; gene ontology; life-history; pH

Running head: Tetrahymena range expansion genomics

Abstract

Populations that expand their range can undergo rapid evolutionary adaptation, which can be aided or hindered by sexual reproduction and gene flow. Little is known about the genomic causes and consequences of such adaptation. We studied genomic adaptation during experimental range expansions of the protist *Tetrahymena thermophila* in landscapes with a uniform environment or a pH-gradient, both in the presence and absence of gene flow and sexual reproduction. We used pooled genome sequencing to identify genes subject to selection caused by the expanding range and by the pH-gradient. Adaptation to the range expansion affected genes involved in cell divisions and DNA repair, whereas adaptation to the pH gradient additionally affected genes involved in ion balance, and oxidoreductase reactions. These genetic changes may result from selection on growth and adaptation to low pH. Sexual reproduction affected both *de novo* mutation and standing genetic variation, whereas gene flow and the presence of a pH-gradient affected only standing variation. Sexual reproduction may have aided genetic adaptation during range expansion, but only in the absence of gene flow, which may have swamped expanding populations with maladapted alleles.

Introduction

Range expansions and biological invasions are increasingly common in many species. They often result from anthropogenic disturbances or species introductions (Parmesan et al., 1999; Kowarik, 2003; Chen et al., 2011). Because range expansions can have dramatic consequences on species distributions and ecosystems (see for example Grosholz, 2002; Phillips et al., 2009), it is ecologically and economically important to understand their ecological and evolutionary consequences. Theory predicts that expanding populations are under strong selection for both dispersal and population growth (Burton et al., 2010; Phillips et al., 2010; Shine et al., 2011). Recent studies have demonstrated that expanding populations indeed rapidly evolve their dispersal behaviour (Brown et al., 2007; Williams et al., 2016; Ochocki and Miller, 2017) and life-history strategy (Phillips, 2009; Fronhofer and Altermatt, 2015). Additionally, they can evolve specific adaptations to local abiotic conditions, such as temperature and novel food sources (Van Petegem et al., 2016; Szűcs et al., 2017). Such rapid evolution can in turn affect the range expansion rate (Weiss-Lehman et al., 2017; Ochocki and Miller, 2017; Williams et al., 2019).

Because selection pressures during range expansion can be strong, populations at a range edge may adapt quickly. Such adaptation can be facilitated by reshuffling existing genetic variation through sexual reproduction (Bell, 1982; Maynard-Smith, 1978; Otto and Lenormand, 2002). This benefit of sexual reproduction has been demonstrated repeatedly, (Colegrave, 2002; McDonald et al., 2016; Luijckx et al., 2017; Petkovic and Colegrave, 2019), and is especially pronounced in the presence of high genetic diversity (Lachapelle and Bell, 2012). Increased genetic variation or an influx of locally adapted genes can also increase the success of invasions and range expansions (Schmeller et al., 2005; Lavergne and Molofsky, 2007; Currat et al., 2008). However, sexual reproduction during range expansions can also hinder adaptation, for example in the process of ”gene swamping”. Gene swamping occurs when populations expand their range against an environmental gradient, and when genes ill-adapted to the range edge flow from a high density core population to a low density population at the range edge. According to the gene swamping hypothesis, such gene flow can decelerate adaptation of a population at the range edge (Haldane and Ford, 1956; García-Ramos and Kirkpatrick, 1997; Kirkpatrick and Barton, 1997; Polechová and Barton, 2015; Polechová, 2018). Such detrimen-

tal gene swamping has recently been demonstrated both in an experimental study (Moerman et al., 2020b) and a field study (Bachmann et al., 2020).

Little is known of the genetic mechanisms by which gene swamping may alter adaptation during range expansions. With few exceptions (Bosshard et al., 2020), studies on the evolutionary genetics of range expansions have focused on genetic population structure and diversity (see for example Swaegers et al., 2013; Bors et al., 2019). This population genetic work has demonstrated that, first, populations at the edge of a range typically harbour much less genetic variation than the core populations (Excoffier et al., 2009). Second, genetic drift can strongly affect populations at a range edge, which can lead to ”allele surfing”, the increase in frequency and potential fixation of neutral or maladative mutations at a range edge (Klopfstein et al., 2006; Excoffier et al., 2009). Despite such work, the genetic basis of phenotypic evolution during range expansions remains poorly understood. For example, we know little about the kinds of genes that are involved in phenotypic evolution during range expansion. We know even less about genes involved in gene swamping. Because phenotypic change could be caused by genetic change, by phenotypic plasticity, or by maternal effects (Bonduriansky and Day, 2018) it is also important to disentangle genetic and non-genetic causes of phenotypic change. For example, dispersal behaviour can have a strong genetic component (Saastamoinen et al., 2018), but it can also be a plastic response to population density or habitat quality (see chapter 1 in Clobert et al., 2012), and it can be modified by maternal effects (Donohue, 1999).

We here studied the population genomics of experimental range expansions in the protist *Tetrahymena thermophila*. We were especially interested in genomic changes caused by sexual reproduction and gene flow, when a population expands either into a uniform environment or into an abiotic (pH) gradient. To this end, we performed whole genome sequencing of pooled *T. thermophila* populations from a previously published range expansion experiment (Moerman et al., 2020b). In this experiment, we used a full factorial design in which populations reproduced either sexually or asexually, and in which gene flow from the range core to the range edge was either allowed or prohibited during range expansion. We identified the populations with the strongest phenotypic changes, pooled individuals from these populations, and sequenced the genomic DNA of the resulting pools. We then investigated how genetic changes

from standing variation and from *de novo* mutations for these pooled populations were affected by reproductive mode (asexual or sexual), gene flow, and a pH-gradient. Furthermore, we identified genes involved in adaptation to the range expansion itself and to the pH-gradient. We found that sexual reproduction increased the number of *de novo* mutations whose frequency rose to detectable values under all conditions we examined. In contrast, sexual reproduction only affected allele frequency changes in standing genetic variation when gene flow was absent. Genes involved in general adaptation were preferentially associated with DNA repair, mitosis, and gene expression. In contrast, genes involved in gradient-specific adaptation were associated with ion transport and oxidoreductase reactions.

Material and methods

Study organism and experimental populations

Tetrahymena thermophila is a freshwater ciliate commonly used in ecological and evolutionary microcosm experiments (Collins, 2012; Altermatt et al., 2015; Cairns et al., 2020; Moerman et al., 2020b). These ciliate species are also commonly used in the study of range expansions. (Giometto et al., 2014; Fronhofer and Altermatt, 2015; Giometto et al., 2017) We used ancestral and evolved populations from a range expansion experiment by Moerman et al. (2020b) to investigate the genomic basis of adaptation during range expansions. We explain the most pertinent details of this experiment here and refer the reader for other details to Moerman et al. (2020b).

Briefly, in this experiment, we mixed four phenotypically divergent (Moerman et al., 2020a) clonal strains of *T. thermophila* to create a genetically diverse ancestral population. For details on strain identities and the mixing of the ancestral population, see section S1.1 in the Supplementary Material. Subsequently, we allowed 40 replicates of this ancestral population to expand their range for ten weeks (250 generations). We kept all stock cultures and populations during experimental evolution in a climate room at 20 °C.

During this experimental range expansion, we used an established system of two-patch landscapes (Fronhofer and Altermatt, 2015) to emulate an expanding range edge (see also sec-

tion S1.2 in the Supplementary Material). These two-patch landscapes consist of two connected tubes, filled with modified Neff medium (Cassidy-Hanley, 2012). The connection between the tubes can be opened or closed using a plastic clamp. We initiated the experiment by inoculating one patch of the two-patch landscape (”home patch”) with 200 µL of the ancestral population. During experimental range expansion, we controlled 1) the abiotic conditions (”uniform”: pH always equals 6.5, ”pH-gradient”: pH gradually decreases from 6.5 of 4.0 during range expansion to emulate an environmental pH-gradient), 2) reproduction (”asexual”: pure asexual reproduction, ”sexual”: sexual reproduction induced at regular intervals) and 3) gene flow (”absent”: no gene flow; ”present”: regular gene flow). Altogether, our experiment thus used eight treatments. For each treatment, we evolved five replicate populations, resulting in a total of 40 populations.

Experimental evolution

We then subjected the populations to ten weeks of experimental evolution, during which we repeated the same cyclical procedure every 14 days. Specifically, we subjected each population to three dispersal events, which took place on days 1, 3 and 5 of the 14-day cycle. After the third dispersal event, we subjected populations of the appropriate treatment groups to a gene flow and/or sexual reproduction event on day 8. Lastly, we subjected the populations to an additional two dispersal events on days 10 and 12. For a detailed description and a visual representation of the experimental design, see the Supplementary Material sections S1.3 and S1.4.

Common garden and growth assessment

After a common garden phase of 18 generations to reduce epigenetic and maternal effects, we assessed the change in intrinsic population growth rate of the evolved populations compared to the ancestral population. To do so, we performed growth assays for the evolved and ancestral populations under multiple pH-values (pH 6.5, 6.0, 5.5, 5.0, 4.5, 4.0, 3.5 and 3.0). We then fit a Beverton-Holt population growth model (Beverton and Holt, 1993) to infer the intrinsic population growth rate r_0 (see section S1.5 of the Supplementary Material). We then calculated

the change in fitness that occurred in each evolved population as the change in the population's intrinsic growth rate r_0 relative to the ancestral population. More specifically, we calculated the log-ratio response as

$$ln(\frac{r_0, e}{r_0, a}), \tag{1}$$

where r_0, e is the intrinsic population growth rate of the evolved population and r_0, a is the intrinsic population growth rate of the ancestral population.

Sequencing and identification of variants

To identify candidate loci for genetic adaptation during range expansion, we performed population level sequencing ("PoolSeq"; Schlötterer et al., 2014).

Tetrahymena thermophila cells have a diploid non-expressed micronucleus and a highly polyploid (n=45) expressed macronucleus (Lynn and Doerder, 2012). We only wished to sequence the macronuclear DNA, because only the expressed macronuclear DNA affects the phenotype. We thus first separated macronuclei from micronuclei. Following nuclear separation, we extracted DNA from each population. Subsequently, we selected populations for genomic sequencing based on their intrinsic population growth rate r_0. Specifically, we sequenced DNA from the two populations with the highest intrinsic population growth rate r_0 in each of the eight treatment groups (uniform/gradient, asexual/sexual, gene flow/no gene flow; see Table S8 for r_0-values from all evolved populations). Next, we mapped the obtained sequences to the *T. thermophila* reference genome, and performed variant calling to identify genetic variants. In analyzing the resulting data, we studied standing genetic variation and its allele frequency change, as well as mutations that had occurred *de novo*. In the next sections, we describe each of these steps in detail.

Nuclear separation and DNA extraction

We separated macronuclei and micronuclei of the four ancestral clones and of all 34 surviving evolved populations using a differential centrifugation protocol adapted from Sweet and Allis

(2006, see Supplementary Material section S2 for details). As a result of this procedure, we obtained for every population three pellets containing macronuclei (pellet P1, P2 and P3). We extracted DNA from each of the pellets using a Qiagen® DNeasy Blood and Tissue kit (Cat No./ID: 69506), following the manufacturer's protocol (see `https://www.qiagen.com/us/resources/resourcedetail?id=63e22fd7-6eed-4bcb-8097-7ec77bcd4de6&lang=en`). After DNA extraction, we determined the quality and quantity of the DNA using a Nanodrop spectrophotometer (model ND-1000) and selected for each population the pellet with the highest quality for sequencing (see Supplementary Material section S3). We stored DNA samples in a $-80\,^{\circ}$C freezer until sequencing. Library preparation and sequencing was performed by the Functional Genomics Center Zürich, which prepared 2x150 bp paired-end libraries and sequenced them using the Illumina Novaseq 6000 platform. We aimed to obtain population level DNA-sequences at approximately 100x genome coverage (see Supplementary Material section S4.2 for coverage statistics).

Bioinformatic pipeline

We trimmed the reads of all fastq files using Trimmomatic version 0.39 (Bolger et al., 2014) to remove any Illumina adapter sequences (ILLUMINACLIP option) and low quality segments (SLIDINGWINDOW:4:15). Additionally, we removed any reads falling below a length of 36 base pairs (MINLEN:36). After trimming, we mapped the reads for each of the 20 sequenced populations to the *Tetrahymena thermophila* macronuclear reference genome (Eisen et al., 2006) using the Burrows-Wheeler aligner (Li and Durbin, 2009) version 0.7.17-r1188. On average, 97.7% of the reads aligned to the reference genome across all populations (see Supplementary Material section S4.1 for mapping rates of all populations).

We then called genetic variants using the multiallelic caller in BCFtools (Danecek et al., 2016). We first called multiallellic variants to identify sites where the four ancestral and 16 evolved populations differed from the reference genome, and stored the resulting allele counts for every position along the genome.

We mapped all variants to genes using bedtools (v2.27.1; Quinlan and Hall, 2010). To this end, we first obtained the genome annotation file from the *T. thermophila* reference genome

(Eisen et al., 2006), and filtered this file to only keep entries corresponding to protein-coding genes. We then used the intersect function of the bedtools package to map the variants to the gene entries.

Quantifying allele frequencies

To detect changes in allele frequencies, we first calculated the expected frequency of alleles, i.e., the frequency of each allele present in the ancestral population. Because we had sequenced the ancestor clones individually, we started by identifying all positions along the genome where at least one of the ancestral clones differed from the reference genome. We then calculated for all these positions the allele frequency in the ancestral population (see Supplementary Material section S7). Subsequently, we filtered these positions to exclude positions with low sequencing quality and potential artefacts (see Supplementary Material section S9.1 for details). We then calculated the allele frequencies for the 16 evolved populations at all positions along the genome that had not been eliminated by filtering. For each of the evolved populations, this resulted in a list of positions where genetic variation had initially been present in the ancestral population, and where allele frequency changes in this standing variation were therefore possible.

Identifying *de novo* mutations

To detect *de novo* mutations, we kept only those variants that differed from the reference genome in an evolved population, and that were not present in any of the four ancestral clones. We then filtered these positions to exclude positions with low sequencing quality and potential artefacts (see Supplementary Material section S9.2 for details).

Statistical analyses

We performed all statistical analyses, unless otherwise specified, using the R statistical language version 4.0.2 (R Core Team, 2020).

Analyzing fitness changes

We applied the fitness assay described in Moerman et al. (2020b) to the populations included in the genomic analysis, to verify that the results remained the same when only including populations used for the genomic analyses. Briefly, we determined the change in the intrinsic population growth rate r_0 of each evolved population relative to the ancestral population in the pH conditions experienced during range expansion (that is, pH 6.5 for populations expanding into a uniform environment, and pH 4.0 for populations expanding into a pH-gradient). For details and results on this analysis, see section S8 in the Supplementary Material.

Quantifying frequency changes in standing genetic variation.

We first determined for all polymorphic sites in the ancestral population if the alleles at the site had experienced a significant change in frequency for any of the 16 evolved populations. To do so, we performed a binomial test comparing the number of reference and non-reference reads in evolved populations with the expected allele frequency of the ancestral population. We used a p-value of 0.001 in this test, to which we applied a Bonferroni correction (rounded down to the nearest power of ten) to avoid falsely detecting allele frequency changes due to multiple testing. We therefore considered only variants with a p-value smaller than 10^{-10} as significantly different from the expectation.

In a preliminary analysis of standing genetic variation, we observed that the ancestor clone SB3539 showed a number of anomalies that made it desirable to exclude it from further analysis (see Supplementary Material section S5.1). To this end, we limited our analyses to those alleles whose frequency changed by an amount that is larger than could possibly be caused by selection against this clone. Specifically, we only analyzed alleles whose frequency changed by a value of 0.3 or higher (rounded up from 0.27, the starting frequency of clone SB3539).

Next, we assessed if the number of alleles whose frequency changed significantly was affected by the mode of reproduction, gene flow, or the pH-gradient. To do so, we counted for every evolved population the number of variants that significantly changed in allele frequency. In this analysis, we applied various cut-offs (0.3, 0.4, 0.5, 0.6, 0.7 and 0.8) for the magnitude of the allele frequency change, i.e., the minimum absolute difference between the ancestral allele

frequency and the observed allele frequency of the evolved population, required for variants to be included. We fit a generalized linear mixed model with a Poisson distribution to the resulting data, using the glmer-function (lme4 R-package, version 1.-23; Bates et al., 2015). We first created a model in which reproduction, gene flow, abiotic conditions and the cut-off value are all allowed to interact as fixed effects (full interaction model), and using the replicate population (i.e. the identity of each replicate evolved population) as a random effect. We then used the dredge function (MuMIn R-package, version 1.43.17; Bartoń, 2009) to compare all possible models using the Bayesian Information Criterion (BIC; Gelman et al., 2014) for model comparison (lower BIC score indicates better model fit). We report model weights following the BIC comparison, as well as summary statistics (type-III Anova table and model summary) of the best fitting model. Additionally, we plotted the population's fitness change against the number of alleles that changed significantly in frequency. We calculated the Pearson correlation between the two metrics, and did so for different cut-off values of the magnitude of allele frequency change.

Quantifying *de novo* mutations

To assess how the mode of reproduction, gene flow, and the presence/absence of a pH-gradient affect the evolutionary dynamics of *de novo* mutations in our populations, we counted in each of the 16 evolved populations the number of novel variants. To do so, we used eight cut-off values of increasing allele frequencies, ranging from 0.1 to 0.8. For each of these values we counted the number of *de novo* alleles that had reached an allele frequency larger than the cut-off in the post-evolution populations. To test whether this number of variants was affected by reproduction, gene flow, and the pH-gradient, we fit a linear mixed model using the lme-function (nlme R-package, version 3.1-147; Pinheiro et al., 2020) using reproduction, gene flow, presence of a pH-gradient, and cut-off as fixed effects, and replicate population as a random effect. Next, we used the dredge function (MuMIn R-package, version 1.43.17; Bartoń, 2009), to rank all possible models based on the Bayesian information criterion (BIC; Gelman et al., 2014). We report model weights as well as summary statistics (type-III anova table and model summary) of the best fitting model.

General adaptations during range expansion

To identify allele frequency changes likely to be associated with general adaptations —- they were subject to selection regardless of reproductive mode, gene flow, or the presence of a pH-gradient — we focused on alleles whose frequency increased or decreased consistently across all or most populations. Specifically, we started from allelic variants that had been present in the ancestral population, that changed significantly in frequency (p-value$<10^{-10}$), and that could not be attributed to selection against clone SB3539 (allele frequency change$>$0.3). Among these alleles, we then considered only those whose frequency changed significantly and in the same direction for 75% (12/16) of all evolved populations. We call the resulting dataset the ”general adaptation dataset”.

Gradient-specific adaptations: Cochran-Mantel-Haenszel test

To identify allelic changes specifically associated with the presence or absence of the pH-gradient, we performed a Cochran-Mantel-Haenszel test (PoPoolation2-package; Kofler et al., 2011). This test compares pairs of populations and it can identify genomic loci with standing genetic variation that show consistently different allele frequencies. To account for the effect of reproduction and gene flow, we always compared pairs of populations with the same reproduction or gene flow treatment, and that only differed in the presence of the gradient. For example, we compared an asexual population without gene flow expanding into a pH-gradient with another asexual population without gene flow but expanding into a uniform environment (see Supplementary Material section S14.1). With this approach, we aimed at detecting loci that display consistent differences between populations expanding into a pH-gradient and populations expanding into a uniform environment, independent of gene flow and reproductive mode. To account for multiple testing, we applied a Bonferroni correction (Bland and Altman, 1995). Because our genomes harboured approximately 10^5 genetically variable positions, and we wished to use a conservative p-value of 0.001, we only classified positions with a p-value smaller than 10^{-8} significant. We refer to this dataset, which contains all loci experiencing differential selection between populations expanding into a pH-gradient and populations expanding into a uniform environment as the ”gradient-specific adaptation” dataset. We then again

focused on alleles within protein-coding genes and examined the functions of these genes.

Gene ontology enrichment

To assess whether allele frequency changes occur in genes associated with specific functions, we performed a gene ontology (GO) term enrichment analysis using GOWINDA (version 1.12; Kofler and Schlötterer, 2012). To this end, we first queried the Uniprot Knowledgebase to obtain all entries associated with the *T. thermophila* reference genome (TaxID: 312017). This resulted in a list of genes encoded in the Tetrahymena genome together with GO-terms associated with each gene. We then performed a GOWINDA analysis using this dataset as a reference database. We performed this analysis for two sets of genes, the genes associated with general adaptations and the genes associated with gradient-specific adaptations

GOWINDA analysis corrects for multiple testing with the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995). For both GO-term analyses, we kept only those GO-terms with an FDR (false discovery rate) value$<$0.05. Next, we grouped the enriched GO-terms into 13 different categories of cellular mechanisms (see Supplementary Material table S16). We then compared the proportion of all GO-terms associated with the 13 categories between general adaptation genes and gradient-specific genes, using a χ^2-test. Finally, we performed a post-hoc test to assess which categories are differently enriched in the gradient-specific genes and the general adaptation genes using a Benjamini-Hochberg correction (Benjamini and Hochberg, 1995; Ebbert, 2019).

Results

Reproduction and gene flow alter allele frequency change.

Our first analysis assessed how reproductive mode, gene flow and the presence of a pH-gradient affected standing genetic variation in our evolving populations. After we had eliminated likely false positive allele frequency changes due to multiple testing and clonal selection on one of our ancestral clones (see Methods, ”Statistical analyses”), we observed significant allele frequency changes in 13,278 alleles. These changes occurred predominantly in non-coding regions of the

genome (see section S11.1 of the Supplementary Material for details).

We then constructed generalized linear models (Methods) that included the mode of reproduction, gene flow, and abiotic conditions as explanatory variables and the number of alleles which changed in frequency as response variables. We compared models that represented all possible interactions between the explanatory variables to find the best fitting model, according to the Bayesian Information Criterion (BIC; Gelman et al., 2014). We found only support for the full interaction model (weight=1, δBIC of second model = 136.09), showing that all explanatory variables, as well as their interaction affected the number of alleles that changed in frequency.

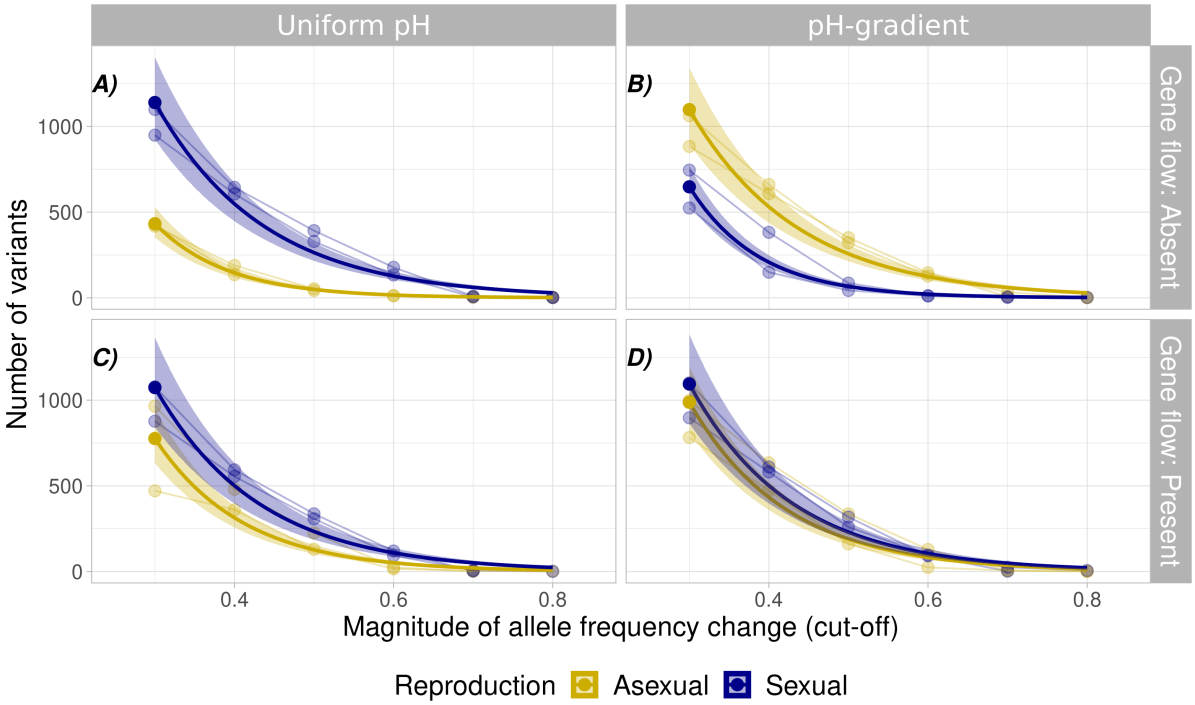

Figure 1: **Reproduction and gene flow alter genetic changes in standing genetic variation.** Number of alleles that changed significantly in frequency during range expansion into a uniform environment (panel A and C) or into a pH-gradient (panel B and D). The x-axis shows the cut-off value in the magnitude of allele frequency. The y-axis shows the number of variants analyzed at each cut-off value. Thin faint lines and circles represent allele frequency data, with lines connecting data from the same replicate population. Thick opaque lines show the model predictions for the best model. Shaded areas show 95%-confidence intervals from the best model. Colours represent mode of reproduction (yellow: asexual, blue:sexual).

The total number of alleles that changed significantly in frequency was on average similar for populations expanding into a uniform environment (Figure 1, panels A and C) and into a pH-gradient (Figure 1, panels B and D). However, the effect of reproduction and gene flow

differed between these populations. In the absence of gene flow, sexual reproduction increased the number of variants whose allele frequency changed when populations expanded into a uniform environment (Figure 1, panel A), but it decreased the number of such variants when populations expanded into a pH-gradient (Figure 1, panel B; Reproduction$\times$Abiotic conditions; $\chi^2{}_{1,92}$=8.3981; p=0.0038). When populations experienced gene flow during range expansion (Figure 1, panels C and D), this difference between populations with a sexual and asexual mode of reproduction disappeared. Complete model statistics can be found in Supplementary Material section S10.1.

Next, to investigate whether the number of genetic changes during range expansion can help explain the extent of fitness change, we compared the number of variants whose allele frequency changed significantly with a change in fitness of the evolved populations. To do so, we calculated the Pearson correlation coefficient between these two metrics. We calculated this coefficient for each cut-off value for the magnitude of allele frequency change. We did so separately for populations expanding into uniform conditions and populations expanding into a pH-gradient, as well as for sexual and asexual populations (Figure 2). For asexual populations (yellow circles and triangles), the allele frequency change and fitness are not correlated. A possible reason is that the lack of sexual recombination in asexual populations limits the separation of maladaptive and beneficial alleles, and may therefore limit the efficacy of selection on beneficial alleles.

In contrast, for sexual populations (blue circles and triangles), the number of alleles changing their frequency and the extent of fitness change are correlated, depending on the abiotic conditions. Specifically, for populations expanding into a uniform environment, populations in which more variants changed their frequency also increased their fitness to a greater extent, and significantly so for three out of five cut-off values (2, left panels). In contrast for populations expanding into a pH-gradient, the number of variants changing in allele frequency and fitness change are not correlated (2, right panels). One candidate explanation is that in a new and stressful environments, such as our pH-gradient, *de novo* mutations may be more important than standing variation for adaptation. However, this is not the case (see section S12 in the Supplementary Material).

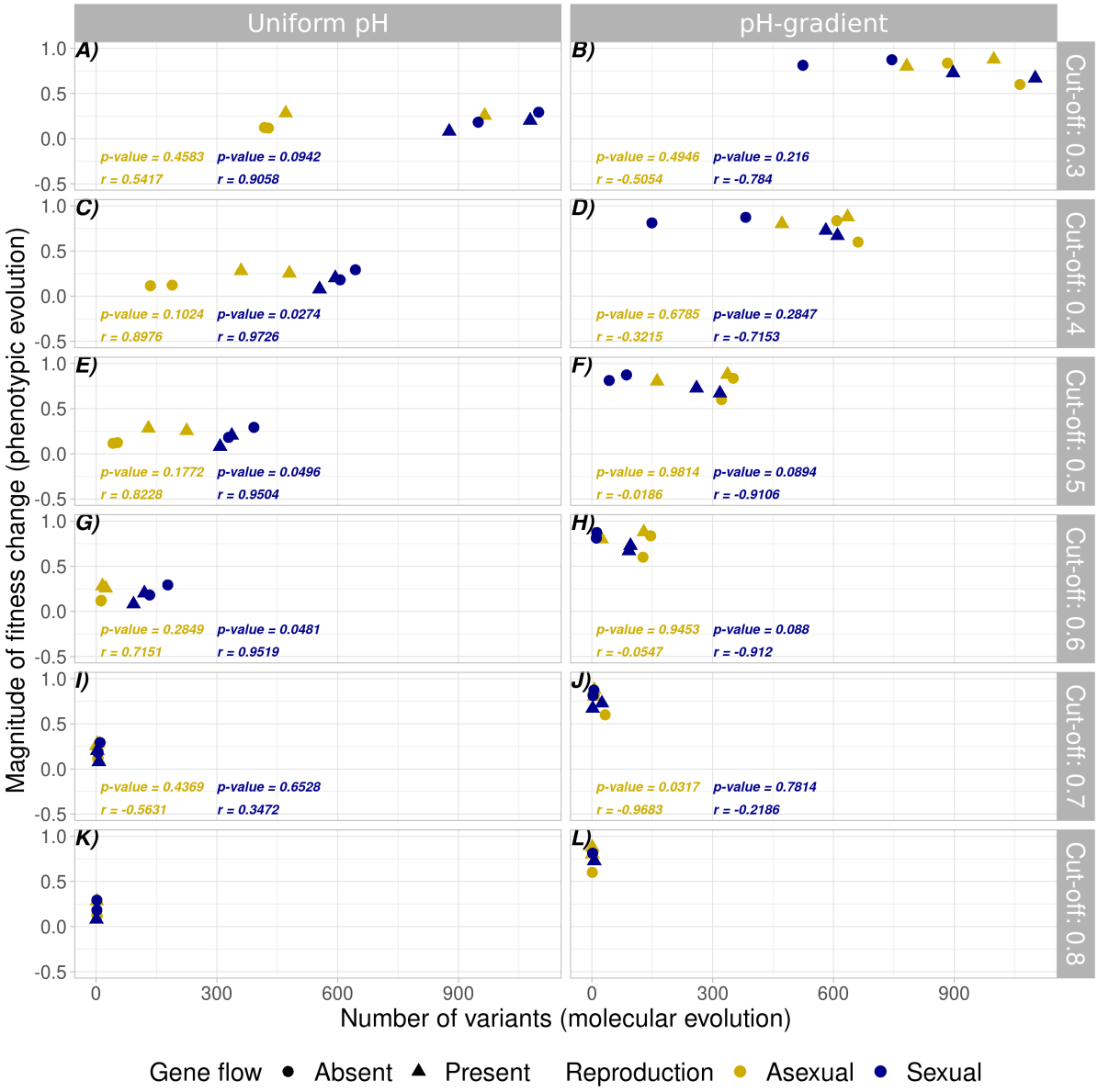
Uniform pH
pH-gradient
Cut-off: 0.3
Cut-off: 0.4
Cut-off: 0.5
Cut-off: 0.6
Cut-off: 0.7
Cut-off: 0.8
A)
p-value = 0.4583
r = 0.5417
p-value = 0.0942
r = 0.9058
B)
p-value = 0.4946
r = -0.5054
p-value = 0.216
r = -0.784
C)
p-value = 0.1024
r = 0.8976
p-value = 0.0274
r = 0.9726
D)
p-value = 0.6785
r = -0.3215
p-value = 0.2847
r = -0.7153
E)
p-value = 0.1772
r = 0.8228
p-value = 0.0496
r = 0.9504
F)
p-value = 0.9814
r = -0.0186
p-value = 0.0894
r = -0.9106
G)
p-value = 0.2849
r = 0.7151
p-value = 0.0481
r = 0.9519
H)
p-value = 0.9453
r = -0.0547
p-value = 0.088
r = -0.912
I)
p-value = 0.4369
r = -0.5631
p-value = 0.6528
r = 0.3472
J)
p-value = 0.0317
r = -0.9683
p-value = 0.7814
r = -0.2186
K)
L)
1.0
0.5
0.0
-0.5
0
300
600
900
Magnitude of fitness change (phenotypic evolution)
Number of variants (molecular evolution)
Gene flow
Absent
Present
Reproduction
Asexual
Sexual

Figure 2: **A positive relationship between fitness change and genetic change when sexual populations expand into a uniform environment.** Statistical association between change in fitness and the number of allele frequency changes in standing genetic variation, for populations expanding into a uniform environment (panels A, C, E, G, I, and K) and populations expanding into a pH-gradient (panels B, D, F, H, J, and L). The x-axis shows the number of standing genetic variants that changed significantly in allele frequency. The y-axis shows the change in growth rate (fitness) of evolved populations compared to the ancestral population ($log_2(r_0$ of evolved population / r_0 of ancestral population)). Horizontal subplots show the data for different cut-off values used for the minimum change in allele frequency at the end of evolution for inclusion in the analysis. Each geometric symbol within a subplot represents a single population, with colour representing reproductive mode (yellow=asexual, blue=sexual) and shape representing gene flow (circle=gene flow absent, triangle=gene flow present). Text insets shows the Pearson correlation coefficient *r* and significance p, based on a correlation test for sexual populations (blue text) and asexual populations (yellow text). For the most stringent cut-off value (allele frequency change >0.8), we did not calculate the correlation, because several populations did not harbour any variants that changed so dramatically in allele frequency.

Frequencies of *de novo* mutations change most strongly in sexually reproducing populations.

In a second analysis, we investigated the presence of *de novo* mutations that rose to detectable allele frequencies. We found a total of 27,694 *de novo* mutations across all 16 evolved populations. They reached frequencies between 1.9% and 100%. Only 0.02% of these *de novo* variants went to fixation. These *de novo* variants occurred predominantly in the non-coding region of the genome (see Supplementary Material section S11.2 for details).

We then studied the role of reproductive mode, gene flow and the pH-gradient (explanatory variables) on the number of *de novo* mutations (response variable) in generalized linear models (see Methods section). We again compared all possible models based on the Bayesian Information Criterion (BIC; Gelman et al., 2014), and found clear support for only one model, which had the cut-off for allele frequency that mutations reached, mode of reproduction and the interaction between these two variables as fixed effects (weight=0.753, δBIC of second model = 3.66). Specifically, we detected similar numbers of *de novo* mutations in populations expanding into a uniform environment (Figure 3, panel A) and into a pH-gradient (Figure 3, panel B). We observed more *de novo* mutations in sexually reproducing populations (Figure 3, blue lines and circles) than in asexual populations (Figure 3, yellow lines and circles; ${\chi^2}_{1,14}$=13.567). At increasingly higher allele frequency cut-offs, the number of novel variants became more similar

for populations with asexual and sexual reproduction (Reproduction×Cut-off; $\chi^2_{1,126}$=18.696).

The increased number of *de novo* mutations that rise to detectable frequency in sexually reproducing populations is consistent with the notion that recombination facilitates adaptative evolution by separating maladaptive mutations from beneficial mutations (Peck, 1994; Otto and Barton, 1997; Gerrish and Lenski, 1998; Burt, 2000; Johnson and Barton, 2002; Gray and Goddard, 2012). See Supplementary Material section S10.2 for statistical details on the best model.

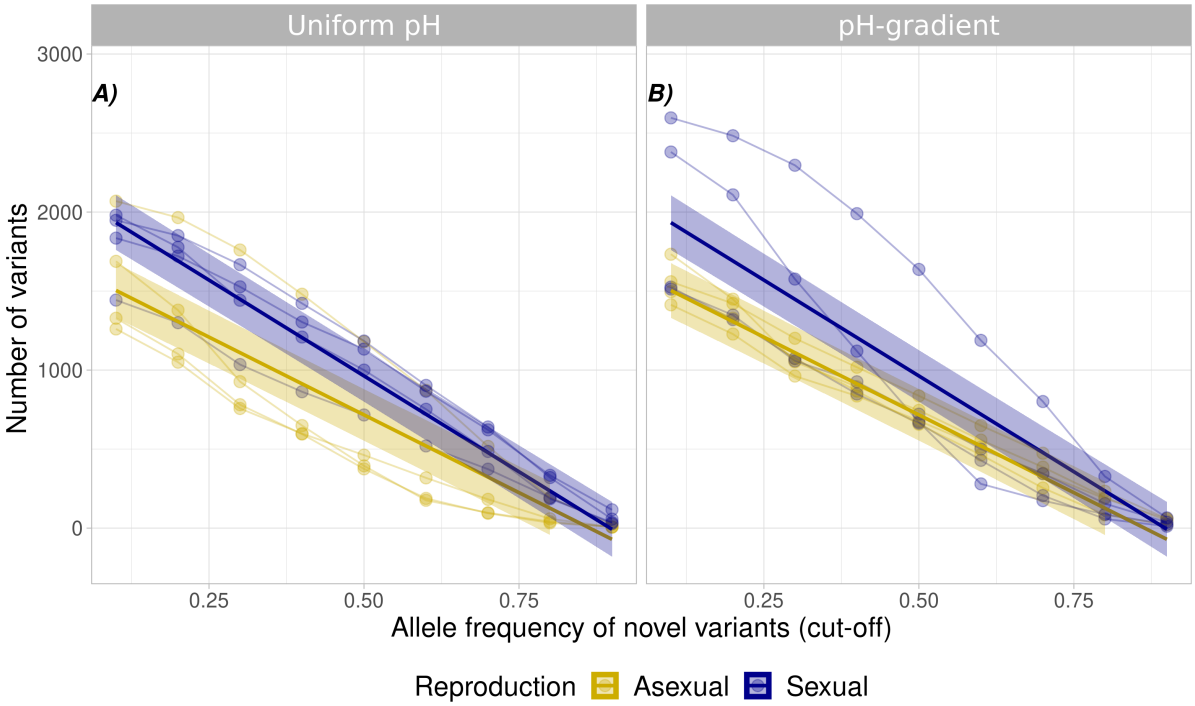

Figure 3: ***De novo* variants increase to higher frequency in sexual populations.** The figure shows the number of novel variants for the evolved populations expanding into a uniform environment (panel A) or into a pH-gradient (panel B). The x-axis shows the cut-off value in minimum allele frequency of the variants. The y-axis shows the number of *de novo* variants (mutations) whose frequency is above a given cut-off. Thin faint lines and circles represent the data, with every line representing a single replicate population. The thicker opaque lines and shaded areas show the predicted means and 95% confidence intervals for the best model. Colours represent the mode of reproduction (yellow=asexually reproducing populations, blue=sexually reproducing populations).

Genes involved in adaptive evolution through allele frequency changes.

In our next analysis, we wanted to identify specific classes of genes that may be involved in adaptive evolution. In such an analysis, it is useful to distinguish between genes involved in

general adaptation to conditions experienced by all populations, such as the growth medium and the expanding range, and genes that are specifically involved in adaptation to our pH-gradient. For both classes of genes, we here focus on alleles that changed significantly in frequency during evolution, and report an analogous analysis of *de novo* mutations in Supplementary Material section S13.

Changes in standing variation associated with general adaptation involve transmembrane proteins and kinase domains.

Of all polymorphic genomic loci at which significant allele frequency changes occurred, 242 loci fit our general adaptation criterion. Of these 242 loci, 53 occurred in protein-coding regions, which fell into 43 different genes (Table S17). Of these 43 genes, 12 encode transmembrane proteins, four encode proteins with a kinase domain, three encode proteins with a cyclic nucleotide-binding domain, three encode zinc finger proteins, and two encode cation channel proteins. All other genes either encode uncharacterized proteins (10 out of 43 genes) or fell into unique functional categories (see also Table S17).

Gradient-specific changes in allele frequency are associated with ion balance.

We found 97,690 genomic positions for which we could quantify differences in allele frequencies through a Cochran-Mantel-Haenszel test (Methods). After Bonferroni correction, we retained 4,388 positions with significantly different allele frequencies between populations expanding into a pH-gradient and populations expanding into a uniform environment. 1,758 of these positions occurred in protein-coding regions, and fell into 636 different genes (Table S18 in the Supplementary Material). The largest subset of these genes with a functional annotation (136 genes) encode transmembrane proteins. Other highly represented groups include genes that encode kinase domain proteins (99 genes), cyclic-nucleotide binding domain proteins (39 genes), cation channel proteins (18 genes), and zinc-fingers (14 genes). The remaining genes either encode uncharacterized proteins (132 genes) or they fell into functional categories rep-

resented by only few genes. Notably, two genes, a gene encoding a cation channel family protein and a gene encoding an oxalate/formate antiporter were involved in both gradient-specific changes in allele frequency and gradient-specific *de novo* mutations (Table S18 and Table S20 in the Supplementary Material). Because both cation channels and oxalate/formate antiporters can play a role in ion transport (Wada et al., 1987; Ruan et al., 1992; Cyert and Philpott, 2013), we speculate that these genes help maintain the ion balance in the cell.

General adaptation preferentially affects genes involved in DNA repair and gene expression.

To find out whether general adaptation and gradient-specific adaptation involved genes with different functions, we studied the differential enrichment of gene ontology terms in these sets of genes.

Among 701 GO-terms associated with the 43 general adaptation genes, 72 terms in 13 major categories were significantly enriched at a false discovery rate (FDR) of <0.05 after multiple testing correction (Benjamini and Hochberg, 1995; Kofler et al., 2011). Likewise, among 1928 GO-terms associated with 626 gradient-specific genes, 693 terms in 13 major categories were significantly enriched.

In Figure 4, the percentage of the 13 major categories is shown for the GO-terms associated with general adaptation (Figure 4 panel A) and gradient-specific adaptation (Figure 4 panel B). We found only two significantly differentially enriched GO-term categories. The first comprise transcription and translation functions, which were more likely to be associated with genes involved in general adaptation (8.97% of all GO-terms) than with genes involved in gradient-specific adaptation (1.3% of all GO-terms; FDR<0.001). The second comprises mitosis, DNA repair and chromosome division, which were also more likely to be associated with genes involved in general adaptation (16.67% of all GO-terms) than in genes associated with genes gradient-specific adaptation (3.75% of all GO-terms; FDR<0.001). We speculate that these cellular functions may be important when selection for faster population growth results in faster cell division and therefore potentially more DNA damage.

Genes involved in gradient-specific adaptation were associated more often with ion binding

and transport, as well as with mitochondrial functioning and oxidoreductase reactions. However, these differences were no longer significant after correcting for multiple testing. Nonetheless, both functions may play a role in the maintenance of ionic balance and the oxidative stress response associated with low pH (Lambert and Brand, 2004; Lemarie et al., 2011; Cyert and Philpott, 2013; Lindberg and Collins, 2020).

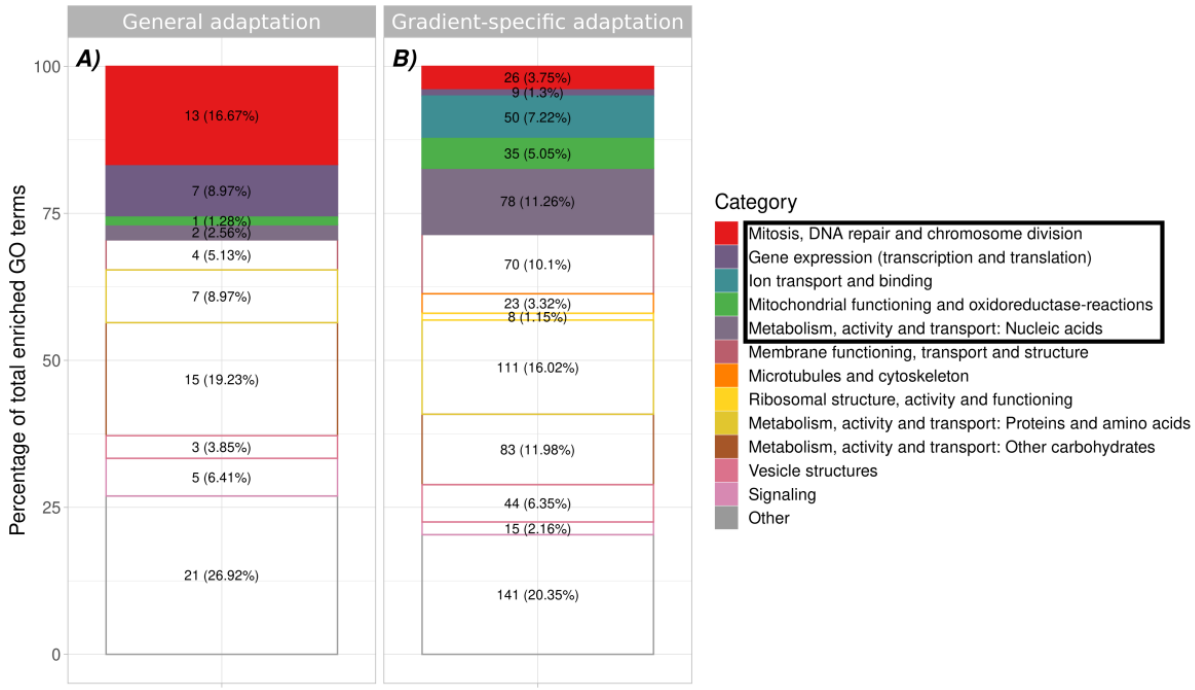

Figure 4: **General and gradient-specific adaptations during range expansion.** Partitioning of all enriched gene ontology terms among 13 major functional categories for genes involved in (panel A) general adaptations and (panel B) gradient specific adaptations. Colours correspond to the 13 major GO categories we identified. Numbers in each rectangle represent the number and percentage of genes whose annotation fell in each category. Differentially enriched categories are displayed as filled rectangles and are boxed in the color legend. The remaining categories are displayed as open rectangles with a coloured border.

Discussion

We investigated the genetic basis of adaptations that occur when populations of *T. thermophila* expand their range. Specifically, we investigated how a pH-gradient, sexual reproduction, and gene flow affect the number of allele frequency changes in standing genetic variation and the number of *de novo* mutations that rise to detectable frequency. Additionally, we investigated which kinds of genes may have been involved in general adaptations to the growth environment

and the expanding range, as opposed to adaptations specific to a pH-gradient.

We found that sexual reproduction increased the incidence of *de novo* mutations that rise to detectable frequencies. The effect of sexual reproduction on standing genetic variation was more complex. In the absence of gene flow, sexual reproduction increased the number of allele frequency changes for populations expanding into uniform environments, but decreased the number of such changes for populations expanding into a pH-gradient. In the presence of gene flow, this difference between sexual and asexual populations disappeared. We found that general adaptation mostly involved genes associated with mitosis, DNA-repair and gene expression. Gradient-specific adaptation mostly involved genes associated with ion transport and oxidoreductase reactions.

The role of sexual recombination on genetic changes during range expansions

By reshuffling genetic variation, sexual reproduction can have important consequences on adaptive evolution (Maynard-Smith, 1978; Bell, 1982; Otto, 2009). This reshuffling of geentic variation may affect both existing variation and new mutations that occur during evolution.

Most mutations that arise are neutral or deleterious (Loewe and Hill, 2010). The accumulation of mutations therefore often leads to a gradual decrease in fitness. Sexual reproduction can help avert this decrease by purging deleterious mutations and separating them from rare beneficial mutations (Muller, 1932; Maynard-Smith, 1978; Peck, 1994; Fisher, 2000; Johnson and Barton, 2002; McDonald et al., 2016). This role of sexual reproduction can help explain that more mutations rose to detectable allele frequency in populations with a sexual mode of reproduction. Because chromosome division in the polyploid macronucleus of asexual lineages of *T. thermophila* happens randomly, offspring of the same individual may vary in the copy numbers of specific mutations (Ruehle et al., 2016). Therefore, in asexual *T. thermophila* populations, selection likely acts against individuals carrying deleterious mutations at a high copy number. In contrast, sexual recombination may reduce this mutational load by separating maladaptive mutations from neutral or beneficial mutations (Maynard-Smith, 1978), which may alleviate selection against *de novo* mutations.

To explain why, in the absence of gene flow, sexual reproduction resulted in more allele frequency changes for populations expanding into a uniform environment, and fewer such changes for populations expanding into a pH-gradient, we considered the role of *de novo* mutations. Specifically, such mutations may have been more important for populations expanding into a pH-gradient, whereas standing genetic variation may have been more important for populations expanding into a uniform environment. Because the conditions a population experiences when expanding into a uniform environment are similar to those experienced by the ancestral population, selection on standing genetic variation may suffice for adaptation in this case. This possibility is supported by the positive correlation between the number of genetic changes and the fitness increase for populations expanding into a uniform environment (Figure 2). Sexual recombination could be adaptive for such populations by bringing together adaptive alleles from the different ancestor clones (Maynard-Smith, 1978). In asexual populations, however, the lack of sexual recombination implies that any mutation that arises cannot be separated from its genetic background. Thus, if a beneficial mutation is under strong positive selection, neutral or deleterious alleles that co-occur with it may hitch-hike to high allele frequencies. The importance of hitchhiking is supported by the observation that in asexual populations, the number of allelic changes does not correlate with fitness (see also Figure 2), suggesting that most allelic changes don't affect fitness strongly. In both cases, sexual reproduction may have aided adaptation, either by bringing together adaptive variants in the uniform environment, or by separating adaptive mutations from their genetic background in the pH-gradient.

When gene flow was present, we observed no clear difference in the number of variants whose frequency changed between asexual and sexual populations. The reason may be the detrimental effect of gene flow on adaptation (Lenormand, 2002). Because gene flow and sexual reproduction were linked to each other in our experiment, sexual reproduction may have led to the swamping of expanding populations with maladapted alleles (Haldane and Ford, 1956; García-Ramos and Kirkpatrick, 1997; Kirkpatrick and Barton, 1997), thus counteracting the adaptive effect of selection prior to sexual reproduction. This is why sexual reproduction may not have been beneficial in the presence of gene flow. Both asexual and sexual populations alike may have undergone adaptation primarily via selection on asexual clones.

Genes under selection during range expansions

Genes related to mitosis, DNA repair and chromosome division are involved in general adaptations that occur during range expansions in our experiment. This observation is consistent with range expansion theory, which predicts strong selection on population growth rate (Burton et al., 2010; Phillips et al., 2010; Shine et al., 2011). Faster growth leads to a decreased nucleotide generation time (i.e., the average time it takes before a nucleotide is copied), which can be associated with greater mutation rate per unit time (Martin and Palumbi, 1993; Friedberg et al., 2002; Kunkel, 2004; Friedberg et al., 2005). General selection in our experiment also acted on genes related to DNA repair, which may be an adaptation to reduce the accumulation of mutations due to faster cell division. Because maladaptive mutations accumulate at the edge of a population's range (Excoffier et al., 2009; Peischl and Excoffier, 2015; Peischl et al., 2015) an increase in mutation rate may be especially harmful at this edge.

Range expansion theory also predicts strong selection for increased dispersal (Burton et al., 2010; Shine et al., 2011). This prediction has been validated true in several empirical and experimental studies (Brown et al., 2007; Williams et al., 2016; Ochocki and Miller, 2017; Weiss-Lehman et al., 2017). Although dispersal behaviour often has a genetic basis (Saastamoinen et al., 2018), we did not observe selection on genes obviously associated with dispersal. This observation could have at least two explanations. Firstly, dispersal is often energetically costly (Bonte et al., 2012), and may require investment of resources. Selection on dispersal could therefore be exerted through genes involved in metabolism or feeding behaviour. Secondly, changes in dispersal may be caused by phenotypic plasticity rather than by genetic change (Clobert et al., 2012, see also Petegem et al. 2018). Because *Tetrahymena* shows strong plasticity in dispersal behaviour (Jacob et al., 2016; Fronhofer et al., 2017), and because we observed no stable phenotypic changes in dispersal (Moerman et al., 2020b), the latter explanation may be more likely.

We found that gradient-specific adaptations were mostly associated with ion transport and binding, as well as with oxidoreductase reactions. Selection on genes related to ion transport and binding are likely a response to osmotic stress under low pH conditions (Bremer and Krämer, 2019). Oxidoreductases play an important role in the production of reactive oxygen

species (ROS), which are harmful to cells (Esterházy et al., 2008). Low pH can also contribute to elevated production of ROS (Lambert and Brand, 2004; Lemarie et al., 2011). A previous study found that selection for the capacity to metabolize ROS is an important aspect of evolution under pH-stress in phytoplankton (Lindberg and Collins, 2020). Consequently, genes associated with mitochondrial functioning and oxidoreductase reactions may help cope with the harmful consequences of ROS.

Limitations

While our experiments identified multiple genomic changes associated with range expansions, our ability to generalize from them is limited by the strong selection we observed for one ancestral clone (see section S5.1 in the Supplementary Material). This phenomenon required us to apply stringent criteria to identify genes under selection, and may thus have masked some signatures of general adaptation. This possibility is consistent with the larger number of genes we observed to be associated with gradient-specific adaptation than with general adaptation. Conversely, this increased stringency also means that the signatures of adaptation we identified are likely true positives. A second limitation is that we cannot quantify the fitness effect of individual genetic changes. The reasons lie in the large number of genetic changes we identified, and the difficulty in genetic manipulation of *T. thermophila*. Thirdly, because *T. thermophila* has a large genome, we could only sequence pooled samples from a limited number of populations. One consequence is a loss of linkage information for specific variants, which limits our understanding of the role of sexual reproduction. Another consequence is limited replication of sequencing across populations, which reduces our ability to identify signatures of selection based on parallel evolution across populations. The sequencing of more populations or individuals rather than pooled populations remains an important task for future work.

Conclusions

We showed that selection during range expansions can help populations adapt rapidly to their local environment. It affected traits known to be associated with growth rate and ion balance, but not with dispersal behavior. We do not know how well the latter observation might gen-

eralize to other organisms, where evolution of dispersal-related traits may be more important than we identified (e.g. Brown et al., 2007; Williams et al., 2016; Ochocki and Miller, 2017; Weiss-Lehman et al., 2017). Additionally, we showed that sexual reproduction and gene flow can alter genetic adaptation during range expansion. Little is known about the importance of either process for populations in the wild. More generally, the identification of genomic changes associated with range expansions in wild populations, for example through genome-wide association analysis (Bush and Moore, 2012), remain an important task for future work.

Acknowledgments

We thank Samuel Hürlemann, Silvana Käser and Sarah Bratschi for help with laboratory work, Vanessa Weber De Melo for aid on the nuclear separation protocol, and Carla Bello and Hélène Boulain for help with the bioinformatics. Funding is from the University of Zurich URPP Evolution in Action and the Swiss National Science Foundation, Grant No PP00P3_179089. This is publication ISEM-YYYY-XXX of the Institut des Sciences de l'Evolution – Montpellier. We would also like to acknowledge support by Swiss National Science Foundation grant 31003A_172887 and European Research Council Advanced Grant No. 739874.

Data availability statement

Phenotypic data and output files from variant calling, Cochran-Mantel-Haenszel test and GO-enrichment analyses will be uploaded to Dryad upon acceptance. Raw sequence files will be uploaded to the European Sequence Archive upon acceptance. All analysis scripts and video analysis parameters will be deposited to Github on acceptance.

Author contributions

FM, EAF, AW and FA designed the experiment. FM performed experimental work and statistical analyses. FM, FA, AW and EAF interpreted the results. FM and AW wrote the first version of the manuscript and all authors commented on and approved of the final version.

References

Altermatt, F.; Fronhofer, E. A.; Garnier, A.; Giometto, A.; Hammes, F.; Klecka, J.; Legrand, D.; Mächler, E.; Massie, T. M.; Pennekamp, F.; Plebani, M.; Pontarp, M.; Schtickzelle, N.; Thuillier, V.; and Petchey, O. L. 2015: Big answers from small worlds: a user's guide for protist microcosms as a model system in ecology and evolution. Methods Ecol. Evol. 6(2):218–231.

Bachmann, J. C.; Jansen van Rensburg, A.; Cortazar-Chinarro, M.; Laurila, A.; and Van Buskirk, J. 2020: Gene flow limits adaptation along steep environmental gradients. Am. Nat. 195(3):E67–E86.

Bartoń, K. 2009: MuMIn: Multi-model inference.

Bates, D.; Mächler, M.; Bolker, B.; and Walker, S. 2015: Fitting linear mixed-effects models using lme4. J. Stat. Softw 67(1):1–48.

Bell, G. 1982: The masterpiece of nature : The evolution and genetics of sexuality. CUP Archive.

Benjamini, Y. and Hochberg, Y. 1995: Controlling the false discovery rate: A practical and powerful approach to multiple testing. J. Royal Stat. Soc. B 57(1):289–300.

Beverton, R. and Holt, S. 1993: On the dynamics of exploited fish populations. Springer.

Bland, J. M. and Altman, D. G. 1995: Multiple significance tests: the Bonferroni method. BMJ 310(6973):170.

Bolger, A. M.; Lohse, M.; and Usadel, B. 2014: Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30(15):2114–2120.

Bonduriansky, R. and Day, T. 2018: Extended heredity: A new understanding of inheritance and evolution. Princeton.

Bonte, D.; Van Dyck, H.; Bullock, J. M.; Coulon, A.; Delgado, M.; Gibbs, M.; Lehouck, V.; Matthysen, E.; Mustin, K.; Saastamoinen, M.; Schtickzelle, N.; Stevens, V. M.; Van-

dewoestijne, S.; Baguette, M.; Barton, K.; Benton, T. G.; Chaput-Bardy, A.; Clobert, J.; Dytham, C.; Hovestadt, T.; Meier, C. M.; Palmer, S. C. F.; Turlure, C.; and Travis, J. M. J. 2012: Costs of dispersal. Biol. Rev. 87(2):290–312.

Bors, E. K.; Herrera, S.; Morris, J. A.; and Shank, T. M. 2019: Population genomics of rapidly invading lionfish in the Caribbean reveals signals of range expansion in the absence of spatial population structure. Ecol. Evol. 9(6):3306–3320.

Bosshard, L.; Peischl, S.; Ackermann, M.; and Excoffier, L. 2020: Dissection of the mutation accumulation process during bacterial range expansions. BMC Genomics 21(1):253.

Bremer, E. and Krämer, R. 2019: Responses of microorganisms to osmotic stress. Ann. Rev. Microb. 73(1):313–334.

Brown, G. P.; Shilton, C.; Phillips, B. L.; and Shine, R. 2007: Invasion, stress, and spinal arthritis in cane toads. Proc. Natl. Acad. Sci. U.S.A. 104(45):17698–17700.

Burt, A. 2000: Perspective: Sex, recombination, and the efficacy of selection—Was Weismann right? Evolution 54(2):337–351.

Burton, O. J.; Phillips, B. L.; and Travis, J. M. J. 2010: Trade-offs and the evolution of life-histories during range expansion. Ecol. Lett. 13(10):1210–1220.

Bush, W. S. and Moore, J. H. 2012: Chapter 11: Genome-wide association studies. PLOS Comput. Biol. 8(12):e1002822.

Cairns, J.; Moerman, F.; Fronhofer, E. A.; Altermatt, F.; and Hiltunen, T. 2020: Evolution in interacting species alters predator life-history traits, behaviour and morphology in experimental microbial communities. Proc. R. Soc. B-Biol. Sci. 287(1928):20200652.

Cassidy-Hanley, D. M. 2012: *Tetrahymena* in the laboratory: Strain resources, methods for culture, maintenance, and storage. Methods in Cell Biology, volume 109, pages 237–276. Elsevier.

Chen, I.-C.; Hill, J. K.; Ohlemüller, R.; Roy, D. B.; and Thomas, C. D. 2011: Rapid range shifts of species associated with high levels of climate warming. Science 333(6045):1024–1026.

Clobert, J.; Baguette, M.; Benton, T. G.; and Bullock, J. M. 2012: Dispersal ecology and evolution. Oxford University Press, USA, Oxford, first edition.

Colegrave, N. 2002: Sex releases the speed limit on evolution. Nature 420(6916):664–666.

Collins, K. (editor). 2012: *Tetrahymena thermophila*. Academic Press, Amsterdam, 1 edition edition.

Currat, M.; Ruedi, M.; Petit, R. J.; and Excoffier, L. 2008: The hidden side of invasions: massive introgression by local genes. Evolution 62(8):1908–1920.

Cyert, M. S. and Philpott, C. C. 2013: Regulation of cation balance in *Saccharomyces cerevisiae*. Genetics 193(3):677–713.

Danecek, P.; Schiffels, S.; and Durbin, R. 2016: Multiallelic calling model in bcftools (-m).

Donohue, K. 1999: Seed dispersal as a maternally influenced character: Mechanistic basis of maternal effects and selection on maternal characters in an annual plant. Am. Nat. 154(6):674–689.

Ebbert, D. 2019: chisq.posthoc.test: Perform post hoc analysis based on residuals of Pearson's Chi-Squared Test for Count Data.

Eisen, J. A.; Coyne, R. S.; Wu, M.; Wu, D.; Thiagarajan, M.; Wortman, J. R.; Badger, J. H.; Ren, Q.; Amedeo, P.; Jones, K. M.; Tallon, L. J.; Delcher, A. L.; Salzberg, S. L.; Silva, J. C.; Haas, B. J.; Majoros, W. H.; Farzad, M.; Carlton, J. M.; Smith, R. K.; Garg, J.; Pearlman, R. E.; Karrer, K. M.; Sun, L.; Manning, G.; Elde, N. C.; Turkewitz, A. P.; Asai, D. J.; Wilkes, D. E.; Wang, Y.; Cai, H.; Collins, K.; Stewart, B. A.; Lee, S. R.; Wilamowska, K.; Weinberg, Z.; Ruzzo, W. L.; Wloga, D.; Gaertig, J.; Frankel, J.; Tsao, C.-C.; Gorovsky, M. A.; Keeling, P. J.; Waller, R. F.; Patron, N. J.; Cherry, J. M.; Stover, N. A.; Krieger, C. J.; del Toro, C.; Ryder, H. F.; Williamson, S. C.; Barbeau, R. A.; Hamilton, E. P.; and Orias, E. 2006: Macronuclear genome sequence of the ciliate *Tetrahymena thermophila*, a model eukaryote. PLoS Biol. 4(9):e286.

Esterházy, D.; King, M. S.; Yakovlev, G.; and Hirst, J. 2008: Production of reactive oxygen species by Complex I (NADH:Ubiquinone Oxidoreductase) from *Escherichia coli* and comparison to the enzyme from mitochondria. Biochemistry 47(12):3964–3971.

Excoffier, L.; Foll, M.; and Petit, R. J. 2009: Genetic consequences of range expansions. Annu. Rev. Ecol. Evol. Syst. 40(1):481–501.

Fisher, R. A. 2000: The genetical theory of natural selection. Oxford University Press, Oxford, annotated edition.

Friedberg, E. C.; Wagner, R.; and Radman, M. 2002: Specialized DNA polymerases, cellular survival, and the genesis of mutations. Science 296(5573):1627–1630.

Friedberg, E. C.; Walker, G. C.; Siede, W.; Wood, R. D.; Schultz, R. A.; and Ellenberger, T. 2005: DNA repair and mutagenesis. ASM Press, Washington, D.C, 2nd edition edition.

Fronhofer, E. A. and Altermatt, F. 2015: Eco-evolutionary feedbacks during experimental range expansions. Nat. Commun. 6:6844.

Fronhofer, E. A.; Govaert, L.; O'Connor, M. I.; Schreiber, S. J.; and Altermatt, F. 2018: The shape of density dependence and the relationship between population growth, intraspecific competition and equilibrium population density. bioRxiv page 485946.

Fronhofer, E. A.; Nitsche, N.; and Altermatt, F. 2017: Information use shapes the dynamics of range expansions into environmental gradients. Glob. Ecol. Biogeogr. 26(4):400–411.

García-Ramos, G. and Kirkpatrick, M. 1997: Genetic models of adaptation and gene flow in peripheral populations. Evolution 51(1):21–28.

Gelman, A.; Hwang, J.; and Vehtari, A. 2014: Understanding predictive information criteria for Bayesian models. Stat. Comput. 24(6):997–1016.

Gerrish, P. J. and Lenski, R. E. 1998: The fate of competing beneficial mutations in an asexual population. Genetica 102(0):127.

Giometto, A.; Altermatt, F.; and Rinaldo, A. 2017: Demographic stochasticity and resource autocorrelation control biological invasions in heterogeneous landscapes. Oikos 126(11):1554–1563.

Giometto, A.; Rinaldo, A.; Carrara, F.; and Altermatt, F. 2014: Emerging predictable features of replicated biological invasion fronts. Proc. Natl. Acad. Sci. U.S.A. 111(1):297–301.

Gray, J. C. and Goddard, M. R. 2012: Sex enhances adaptation by unlinking beneficial from detrimental mutations in experimental yeast populations. BMC Evol. Biol. 12(1):43.

Grosholz, E. 2002: Ecological and evolutionary consequences of coastal invasions. Trends Ecol. Evol. 17(1):22–27.

Haldane, J. B. S. and Ford, E. B. 1956: The relation between density regulation and natural selection. Proc. R. Soc. B-Biol. Sci. 145(920):306–308.

Hallatschek, O.; Hersen, P.; Ramanathan, S.; and Nelson, D. R. 2007: Genetic drift at expanding frontiers promotes gene segregation. Proc. Natl. Acad. Sci. U.S.A. 104(50):19926–19930.

Hartl, D. L. and Clark, A. G. 2006: Principles of population genetics. Sinauer Associates is an imprint of Oxford University Press, Sunderland, Mass, 4th edition edition.

Hurvich, C. M. and Tsai, C.-L. 1989: Regression and time series model selection in small samples. Biometrika 76(2):297–307.

Jacob, S.; Wehi, P.; Clobert, J.; Legrand, D.; Schtickzelle, N.; Huet, M.; and Chaine, A. 2016: Cooperation-mediated plasticity in dispersal and colonization. Evolution 70(10):2336–2345.

Johnson, T. and Barton, N. H. 2002: The effect of deleterious alleles on adaptation in asexual populations. Genetics 162(1):395–411.

Kirkpatrick, M. and Barton, N. H. 1997: Evolution of a species' range. Am. Nat. 150(1):1–23.

Klopfstein, S.; Currat, M.; and Excoffier, L. 2006: The fate of mutations surfing on the wave of a range expansion. Mol. Biol. Evol. 23(3):482–490.

Kofler, R.; Pandey, R. V.; and Schlötterer, C. 2011: PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics 27(24):3435–3436.

Kofler, R. and Schlötterer, C. 2012: Gowinda: unbiased analysis of gene set enrichment for genome-wide association studies. Bioinformatics 28(15):2084–2085.

Kowarik, I. 2003: Human agency in biological invasions: Secondary releases foster naturalisation and population expansion of alien plant species. Biol. Invasions 5(4):293–312.

Kunkel, T. A. 2004: DNA replication fidelity. J. Biol. Chem. 279(17):16895–16898.

Lachapelle, J. and Bell, G. 2012: Evolutionary rescue of sexual and asexual populations in a deteriorating environment. Evolution 66(11):3508–3518.

Lambert, A. J. and Brand, M. D. 2004: Superoxide production by NADH:ubiquinone oxidoreductase (complex I) depends on the pH gradient across the mitochondrial inner membrane. Biochem. J. 382(2):511–517.

Lavergne, S. and Molofsky, J. 2007: Increased genetic variation and evolutionary potential drive the success of an invasive grass. Proc. Natl. Acad. Sci. U.S.A. 104(10):3883–3888.

Lemarie, A.; Huc, L.; Pazarentzos, E.; Mahul-Mellier, A.-L.; and Grimm, S. 2011: Specific disintegration of complex II succinate:ubiquinone oxidoreductase links pH changes to oxidative stress for apoptosis induction. Cell Death Differ. 18(2):338–349.

Lenormand, T. 2002: Gene flow and the limits to natural selection. Trends Ecol. Evol. 17(4):183–189.

Li, H. and Durbin, R. 2009: Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25(14):1754–1760.

Lindberg, R. T. and Collins, S. 2020: Quality–quantity trade-offs drive functional trait evolution in a model microalgal 'climate change winner'. Ecol. Lett. 23(5):780–790.

Loewe, L. and Hill, W. G. 2010: The population genetics of mutations: good, bad and indifferent. Philos. Trans. Royal Soc. B 365(1544):1153–1167.

Luijckx, P.; Ho, E. K. H.; Gasim, M.; Chen, S.; Stanic, A.; Yanchus, C.; Kim, Y. S.; and Agrawal, A. F. 2017: Higher rates of sex evolve during adaptation to more complex environments. Proc. Natl. Acad. Sci. U.S.A. 114(3):534–539.

Lynn, D. H. and Doerder, F. P. 2012: The life and times of *Tetrahymena*. Methods Cell Biol. 109:9–27.

Martin, A. P. and Palumbi, S. R. 1993: Body size, metabolic rate, generation time, and the molecular clock. Proc. Natl. Acad. Sci. U.S.A. 90(9):4087–4091.

Maynard-Smith, J. 1978: The evolution of sex. Cambridge University Press.

McDonald, M. J.; Rice, D. P.; and Desai, M. M. 2016: Sex speeds adaptation by altering the dynamics of molecular evolution. Nature 531(7593):233–236.

Moerman, F.; Arquint, A.; Merkli, S.; Wagner, A.; Altermatt, F.; and Fronhofer, E. A. 2020a: Evolution under pH stress and high population densities leads to increased density-dependent fitness in the protist *Tetrahymena thermophila*. Evolution 74(3):573–586.

Moerman, F.; Fronhofer, E. A.; Wagner, A.; and Altermatt, F. 2020b: Gene swamping alters evolution during range expansions in the protist *Tetrahymena thermophila*. Biol. Lett. 16(6):20200244.

Muller, H. J. 1932: Some genetic aspects of sex. Am. Nat. 66(703):118–138.

Ochocki, B. M. and Miller, T. E. X. 2017: Rapid evolution of dispersal ability makes biological invasions faster and more variable. Nat. Commun. 8:14315.

Otto, S. 2009: The evolutionary enigma of sex. Am. Nat. 174(S1):S1–S14.

Otto, S. P. and Barton, N. H. 1997: The evolution of recombination: removing the limits to natural selection. Genetics 147(2):879–906.

Otto, S. P. and Lenormand, T. 2002: Resolving the paradox of sex and recombination. Nat. Rev. Genet. 3(4):252–261.

Parmesan, C.; Ryrholm, N.; Stefanescu, C.; Hill, J. K.; Thomas, C. D.; Descimon, H.; Huntley, B.; Kaila, L.; Kullberg, J.; Tammaru, T.; Tennent, W. J.; Thomas, J. A.; and Warren, M. 1999: Poleward shifts in geographical ranges of butterfly species associated with regional warming. Nature 399(6736):579–583.

Peck, J. R. 1994: A ruby in the rubbish: beneficial mutations, deleterious mutations and the evolution of sex. Genetics 137(2):597–606.

Peischl, S. and Excoffier, L. 2015: Expansion load: recessive mutations and the role of standing genetic variation. Mol. Ecol. 24(9):2084–2094.

Peischl, S.; Kirkpatrick, M.; and Excoffier, L. 2015: Expansion load and the evolutionary dynamics of a species range. Am. Nat. 185(4):E81–E93.

Pennekamp, F.; Schtickzelle, N.; and Petchey, O. L. 2015: BEMOVI, software for extracting behavior and morphology from videos, illustrated with analyses of microbes. Ecol. Evol. 5(13):2584–2595.

Petegem, K. V.; Moerman, F.; Dahirel, M.; Fronhofer, E. A.; Vandegehuchte, M. L.; Leeuwen, T. V.; Wybouw, N.; Stoks, R.; and Bonte, D. 2018: Kin competition accelerates experimental range expansion in an arthropod herbivore. Ecol. Lett. 21(2):225–234.

Petkovic, N. and Colegrave, N. 2019: Sex increases the probability of evolutionary rescue in the presence of a competitor. J. Evol. Biol. 32(11):1252–1261.

Phillips, B. 2009: The evolution of growth rates on an expanding range edge. Biol. Lett. 5(6):802–804.

Phillips, B. L.; Brown, G. P.; and Shine, R. 2010: Life-history evolution in range-shifting populations. Ecology 91(6):1617–1627.

Phillips, I. D.; Vinebrooke, R. D.; and Turner, M. A. 2009: Ecosystem consequences of potential range expansions of *Orconectes virilis* and *Orconectes rusticus* crayfish in Canada — a review. Environ. Rev. 17:235–248.

Pinheiro, J.; Bates, D.; DebRoy, S.; Sarkar, D.; and R Core Team. 2020: nlme: Linear and nonlinear mixed effects models.

Polechová, J. 2018: Is the sky the limit? On the expansion threshold of a species' range. PLoS Biol. 16(6):e2005372.

Polechová, J. and Barton, N. H. 2015: Limits to adaptation along environmental gradients. Proc. Natl. Acad. Sci. U.S.A. 112(20):6401–6406.

Quinlan, A. R. and Hall, I. M. 2010: BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26(6):841–842.

R Core Team. 2020: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rosenbaum, B.; Raatz, M.; Weithoff, G.; Fussmann, G. F.; and Gaedke, U. 2019: Estimating parameters from multiple time series of population dynamics using Bayesian inference. Front. Ecol. Evol. 6:234.

Ruan, Z. S.; Anantharam, V.; Crawford, I. T.; Ambudkar, S. V.; Rhee, S. Y.; Allison, M. J.; and Maloney, P. C. 1992: Identification, purification, and reconstitution of OxlT, the oxalate: formate antiport protein of *Oxalobacter formigenes*. J. Biol. Chem. 267(15):10537–10543.

Ruehle, M. D.; Orias, E.; and Pearson, C. G. 2016: *Tetrahymena* as a unicellular model Eukaryote: Genetic and genomic tools. Genetics 203(2):649–665.

Saastamoinen, M.; Bocedi, G.; Cote, J.; Legrand, D.; Guillaume, F.; Wheat, C. W.; Fronhofer, E. A.; Garcia, C.; Henry, R.; Husby, A.; Baguette, M.; Bonte, D.; Coulon, A.; Kokko, H.; Matthysen, E.; Niitepõld, K.; Nonaka, E.; Stevens, V. M.; Travis, J. M. J.; Donohue, K.; Bullock, J. M.; and Delgado, M. d. M. 2018: Genetics of dispersal. Biol. Rev. 93(1):574–599.

Schlötterer, C.; Tobler, R.; Kofler, R.; and Nolte, V. 2014: Sequencing pools of individuals — mining genome-wide polymorphism data without big funding. Nat. Rev. Genet. 15(11):749–763.

Schmeller, D. S.; Seitz, A.; Crivelli, A.; and Veith, M. 2005: Crossing species' range borders: interspecies gene exchange mediated by hybridogenesis. Proc. R. Soc. B-Biol. Sci. 272(1572):1625–1631.

Shine, R.; Brown, G. P.; and Phillips, B. L. 2011: An evolutionary process that assembles phenotypes through space rather than through time. Proc. Natl. Acad. Sci. U.S.A. 108(14):5708–5711.

Swaegers, J.; Mergeay, J.; Therry, L.; Larmuseau, M. H. D.; Bonte, D.; and Stoks, R. 2013: Rapid range expansion increases genetic differentiation while causing limited reduction in genetic diversity in a damselfly. Heredity 111(5):422–429.

Sweet, M. T. and Allis, C. D. 2006: Isolation and purification of *Tetrahymena* nuclei. Cold Spring Harb Protoc 2006(4):pdb.prot4500.

Szűcs, M.; Vahsen, M. L.; Melbourne, B. A.; Hoover, C.; Weiss-Lehman, C.; and Hufbauer, R. A. 2017: Rapid adaptive evolution in novel environments acts as an architect of population range expansion. Proc. Natl. Acad. Sci. U.S.A. 114(51):13501–13506.

Thieme, H. R. 2003: Mathematics in population biology. Princeton University Press.

Van Petegem, K. H. P.; Boeye, J.; Stoks, R.; and Bonte, D. 2016: Spatial selection and local adaptation jointly shape life-history evolution during range expansion. Am. Nat. 188(5):485–498.

Wada, Y.; Ohsumi, Y.; Tanifuji, M.; Kasai, M.; and Anraku, Y. 1987: Vacuolar ion channel of the yeast, *Saccharomyces cerevisiae*. J. Biol. Chem. 262(36):17260–17263.

Weiss-Lehman, C.; Hufbauer, R. A.; and Melbourne, B. A. 2017: Rapid trait evolution drives increased speed and variance in experimental range expansions. Nat. Commun. 8:14303.

Williams, J. L.; Hufbauer, R. A.; and Miller, T. E. X. 2019: How evolution modifies the variability of range expansion. Trends Ecol. Evol. 34(10):903–913.

Williams, J. L.; Kendall, B. E.; and Levine, J. M. 2016: Rapid evolution accelerates plant population spread in fragmented experimental landscapes. Science 353(6298):482–485.