Molecular signatures of resource competition: clonal interference drives the emergence of ecotypes

Massimo Amicone[+, *] and Isabel Gordo[+, *]

[+]Instituto Gulbenkian de Ciência (IGC)

[*]corresponding authors: mamicone@igc.gulbenkian.pt, igordo@igc.gulbenkian.pt

Short running title: Clonal interactions drive ecotypes formation.

Keywords: Eco-evolutionary dynamics, rapid adaptation, clonal interference, resource competition, competitive exclusion, diversification, community assembly.

Type of article: Letters.

Content (number of words/items): Abstract (141), Main text (4990), References (70), Figure legends (6).

Author contributions

MA and IG designed the study, all authors wrote the manuscript and provided final approval for publication.

Conflict of interest

The authors declare no conflict of interest.

Data accessibility

Should the manuscript be accepted, the data supporting the results will be archived in an appropriate public repository (Figshare) and the data DOI will be included at the end of the article.

Abstract

Microbial ecosystems harbor an astonishing diversity that can persist for long times. To understand how such diversity is generated and maintained, ecological and evolutionary processes need to be integrated at similar timescales, but this remains a difficult challenge. Here, we extend an ecological model of resource competition to allow for evolution via *de novo* mutation, focusing on large and rapidly adapting asexual populations. Through numerical and analytical approaches, we characterize adaptation and diversity at different levels and show how clonal interference – the interaction between simultaneously emerging lineages – shapes the eco-evolutionary dynamics. We find that large mutational inputs can foster diversification under sympatry, increasing the probability that phenotypically and genetically distinct clusters arise and stably coexist, constituting an initial form of community. Our findings have implications beyond microbial populations, providing novel insights about the interplay between ecology and evolution in clonal populations.

Introduction

Understanding the mechanisms behind the evolution of biodiversity and the formation of communities remains a difficult challenge. One must integrate ecology and evolution over similar timescales, as taken together, they can give rise to phenomena that could not be explained by either alone (Schoener, 2011). The competitive exclusion principle (Hardin, 1960) theoretically bounds the number of species by the number of limiting resources. This principle generated an apparent contradiction between theoretical expectations and observations, as ecosystems can be replete of diversity even in limiting environments, both in nature (Hutchinson, 1961; Tilman, 1982; Huston, 1994) and in more controlled laboratory conditions (Maharjan *et al.*, 2006; Gresham *et al.*, 2008; Kinnersley, Holben and Rosenzweig, 2009; Herron and Doebeli, 2013). Much theoretical work has been done to resolve such controversy, often referred to as the "paradox of the plankton" (Hutchinson, 1961). Different ecological mechanisms were proposed to maintain diversity, including trade-offs on the species' traits (Posfai, Taillefumier and Wingreen, 2017), heterogeneity in space (Abrams, 1988) and time (Litchman and Klausmeier, 2001) or gene regulation (Pacciani-Mori *et al.*, 2020). However, solutions often rely on specific conditions, where a small change in the parameters could cause a collapse of the species richness. What amount of diversity can be maintained if the species' traits are not constant but rather evolve? is a critical question that remains to be answered.

Ecological theories have been challenged by evolutionary processes (Geritz *et al.*, 1998). In eco-evolutionary frameworks, mutations generate new genetic variants whose fate depends on the state of the ecosystem and, in turn, their increase in frequency can alter the populations. A common outcome of such eco-evolutionary feedbacks is that

evolution limits diversity by reducing the effectiveness of coexistence mechanisms (Edwards *et al.*, 2018). The diversity that could be possible by ecological principles alone, is reduced by selection of the fittest and competitive exclusion. Although several studies have produced novel understanding on the evolution of diversity (Dieckmann and Doebeli, 1999; Shoresh, Hegreness and Kishony, 2008; Doebeli, 2011; Kremer and Klausmeier, 2017), the majority of them rely on a timescale separation between ecological and evolutionary processes, i.e. on the strong-selection-weak-mutation assumption. The emergence of beneficial mutations is assumed to be much slower than the ecological dynamics, thus, before a new lineage arises, the population has already reached ecological equilibrium. While allowing for analytical tractability this assumption comes at a cost: it neglects the overlap between multiple evolving lineages – clonal interference – a phenomena that has been extensively observed in microbial communities *in vitro* and *in vivo* (Desai, Fisher and Murray, 2007; Barroso-Batista *et al.*, 2014). Population genetics models incorporating clonal interference have generated predictions for the adaptation rate, fixation probabilities and genetic diversity of a population (Gerrish and Lenski, 1998; Park and Krug, 2007; Good *et al.*, 2012; de Sousa *et al.*, 2016), yet excluding ecological interactions. When the population size (N) and/or the mutation rate (U) are not small ($NU >> 1$), multiple lineages can emerge simultaneously, ecologically interact with each other and evolve in non-trivial ways. Although these processes are inevitably intertwined (Lawrence *et al.*, 2012; Barroso-Batista *et al.*, 2014, 2020; Garud *et al.*, 2019), little was done to investigate how they act in chorus.

Here, we study how clonal interference affects the interplay between ecological and evolutionary processes and structures communities. Following a model of competition

for resources (Posfai, Taillefumier and Wingreen, 2017), we build an eco-evolutionary framework, where no restrictive assumption on the timescales is made, as common in other models (Geritz *et al.*, 1998; Shoresh, Hegreness and Kishony, 2008; Good, Martis and Hallatschek, 2018). We follow an initially isogenic population throughout time and show how different patterns of adaptation and diversity emerge. Every generation, multiple mutations arise simultaneously in the population and interactions between clones, which occur via competition, affect both the sign and the strength of selection. To our knowledge this is the first study to investigate adaptation under competition for resources across phenotypic and genetic diversity and to characterize the probability of diversification under different molecular evolution parameters. We describe how a large amount of variation can be maintained via a balance between selection and mutation, and how ecotypes - functionally distinct clusters of genotypes - can emerge, coexist and establish a first form of community.

Model and Methods

Eco-Evolutionary model

We model the dynamics of a single clonal lineage evolving to consume a set of different substitutable resources, constantly replenished in a well-mixed environment. Individuals mutate at a rate U and the fate of the emerging mutations depends on their phenotypic effects, on the resource concentration, on the other individuals present in the environment and on drift.

The underlying dynamics are determined by the MacArthur's consumer resource model (Mac Arthur, 1969), recently formalized by *Posfai et al.* to explain high levels of diversity

in a purely ecological context, *i.e.* in the absence of mutations (Posfai, Taillefumier and Wingreen, 2017). Briefly, let M be the number of types present at time t with densities (*#cells/V*) n_i, *(i=1...M)* and R the number of substitutable resources with input concentrations r_j, *(j=1...R)*. The expected density dynamics of each type are:

$$\frac{dn_i}{dt} = n_i(t)\left(\sum_{j=1}^{R} \frac{\alpha_j^{(i)} r_j}{\sum_{k=1}^{M(t)} n_k(t)\cdot\alpha_j^{(k)}} - \delta\right) \quad (1)$$

where $\alpha_j^{(i)}$ represents the consumption rate of resource j by type i and δ is the death rate. The resource amounts are constant in this model since, as *Posfai et al.*, we assume that metabolic reactions occur much faster than cell division (Posfai, Taillefumier and Wingreen, 2017).

We also assume a finite amount of energy available for each cell and limit their ability of consuming resources by an energetic constraint (E):

$$0 \leq \sum_{j=1}^{R} \alpha_j^{(i)} \leq E\ ,\ \forall i = 1, \dots, M \quad (2)$$

A trade-off has been considered in previous studies as a fixed energy budget (Posfai, Taillefumier and Wingreen, 2017; de Oliveira, Amado and Campos, 2018; Amado and Campos, 2019), but in our model it acts as an upper limit. Assuming equally supplied resources ($r_j = r\ \forall j$) and unitary energy, volume and death rate ($E, V, \delta=1$), the population size is $N=Rr$.

We model an initial isogenic population ($M(t_0)=1$) with given traits $\vec{\alpha}^{(1)}$ and allow for mutations that change the heritable traits and give rise to new genotypes. Every generation, each genotype *i (i=1...M(t))* will generate a Poisson number of mutants with expected value

$$E[\#mut_i] = n_i(t) \cdot U$$

with U being the per-genome, per-generation rate of non-lethal mutations. Assuming an infinite site model, a mutation on genotype i will result into a new individual with unique genotype (i') whose phenotypes differ from the parental traits as:

$$\vec{\alpha}^{(i')} = \vec{\alpha}^{(i)} + \vec{\Delta}, \qquad \vec{\Delta} \in \mathbb{R}^R$$

For simplicity, we assume the Δ_j to be normally distributed. The mutation effects are defined as follows:

$$\vec{\Delta}: \begin{cases} \Delta_j \sim N(0,\sigma), j \text{ sampled from } \{1,\dots,R\} \\ \Delta_z \sim N(0,\rho\cdot\sigma), \text{for all } z \neq j \end{cases}$$

If $\rho=1$, a mutation changes all the traits with equal probability. If $0<\rho<1$, mutations target one trait (randomly sampled with probability $1/R$), but also alter partially the others. If $\rho=0$, a mutation only changes one trait. The parameter ρ modulates different degrees of trait interdependence or equivalently the pleiotropic effect of mutations, while the parameter σ modulates the magnitude of the mutation effects (see Fig. 1B).

In order to respect the boundary condition (2), we assumed that: i) mutations leading to negative values of α_j are loss of function and thus assigned $\alpha_j=0$; ii) mutations that do not respect the energy constraint cannot exist, therefore Δ_j are drawn until the upper limit of (2) is satisfied.

In the limit of discrete time steps, we define the selection acting on genotype i at time t, $s_i(t)$, as the expected increase in abundance in the absence of drift, such that:

$$E[n_i(t+1)] = n_i(t)(1+s_i(t))$$

and from (1):

$$s_i(t) = \sum_{j=1}^{R} \frac{\alpha_j^{(i)} r_j}{\sum_{k=1}^{M(t)} n_k(t)\cdot\alpha_j^{(k)}} - \delta\,. \qquad (3)$$

Thus, the fate of each genotype depends on its ability to consume each of the resources and on the ecosystem's ecology.

To simulate drift, we draw the final abundances via multinomial sampling with probabilities $\frac{E[n_i(t+1)]}{n_{tot}}$ $\forall i=1...M(t)$. Every generation, the number of genotypes $M(t+1)$ is updated together with their abundances and traits. Along the simulations, we record the genotypic and phenotypic composition of the population, from which classical statistics used in population genetics are calculated as discussed below.

Phenotypic and genotypic diversity

In the two-resources case, within a population, each type i is characterized by a vector of its consumption traits $(\alpha_1^{(i)}, \alpha_2^{(i)})$ and a vector of the mutations that gave rise to it, each with a unique identifier (e.g. the vector [1,2,7,10] represents genotype 10 whose ancestors are, in order, genotypes 7,2 and 1, and 1 is the ancestor common to every type). From this implementation we can reconstruct the entire phylogenetic tree of a population at any time point and map it on the phenotypic space. To measure the genetic diversity of a population we quantify the average pairwise genetic distance π_G in a sample of m individuals: $\pi_G(m) = \frac{\sum_{(i,j)} d_G(i,j)}{\binom{m}{2}}$, where $m=100$ and $d_G(i,j)$ is the number of mutations that separate genotype i from j. From π_G and the number of segregating sites in the sample, we further compute another population genetics statistic: Tajima's D, whose expectation under the simplest population genetics neutral model is zero (Tajima, 1989). At the functional level, we compute the average pairwise phenotypic distance π_P, defined as: $\pi_P(m) = \frac{\sum_{(i,j)} d_E(i,j)}{\binom{m}{2}} \cdot \frac{1}{\sigma}$, where d_E is the classical Euclidean distance.

Neutral mutation model

To understand the dynamics in the absence of selection, we run simulations with only drift. Genotypes acquire mutations with the same trait effects and probability as described before, but their growth probabilities are equal and do not depend on the phenotypes. Modelling the explicit α_j under neutrality, instead of assuming that the mutations have no effect (i.e. $\Delta_j=0$), allows for a better comparison with the model of selection. These neutral simulations were run for the same time as the selection case. Even though this time is much smaller than that needed to accumulate neutral changes in a classical Wright-Fisher neutral model (Wright, 1930; Fisher, 1958), the outcome informs on how drift contributes to the patterns of molecular and phenotypic evolution measured here.

Parameters of the numerical simulations

In each simulation, an initially maladapted ($\alpha_1^{(1)} = \alpha_2^{(1)} = 0.05$) monomorphic population undergoes the eco-evolutionary process described above for 10000 generations. Its genotypic and phenotypic compositions are analyzed along time, for the following set of parameters: $N=10^7$, $U: \{10^{-8}, 5\cdot10^{-8}, 10^{-7}, 5\cdot10^{-7}, 10^{-6}, 5\cdot10^{-6}, 10^{-5}, 5\cdot10^{-5}\}$, $\sigma: \{0.0125, 0.025, 0.05\}$, $\rho: \{0, 0.5, 1\}$. The parameter combinations are studied under selection or neutrality, with the exception of fixing $\sigma=0.05$ and $\rho=0$ in the latter, as by changing these values, the outcome would be equivalent. Of these, each set was simulated in 100 independent replicates to obtain the statistics of diversity. In order to disentangle the role of the energetic constraint assumption, we simulated adaptation under selection but without any boundary condition (i.e. $E=+\infty$ in (2)), with parameters: $N=10^7$, $U: \{10^{-8}, 10^{-7}, 10^{-6}, 10^{-5}\}$, $\sigma: \{0.0125, 0.025, 0.05\}$, $\rho=0$. When we tested additional parameters, this is specified in the text. The algorithm was written in *R* (*version 3.6.1*) and the results analyzed in *RStudio*.

Results

Competition-driven diminishing return and the rate of adaptation

The initially monomorphic population, which is poorly adapted, is expected to acquire mutations that improve the ability to consume the available resources and to advance in the phenotypic space towards better adapted states. Our aim is to identify what influences the speed of this adaptive process.

We first obtain some analytical approximations on how the selective pressure changes in time and with the genetic composition of the population, under the competition for resources set by (1). Selection acting on an emerging genotype *i* is given by (3) and depends on the population investment on each resource *j*: $e_j^{(M(t))} := \sum_{i=1}^{M(t)} n_i(t) \cdot \alpha_j^{(i)}$. Let us first simplify the problem by considering a monomorphic (*M=1*) population whose phenotype ($\vec{\alpha}$) mirrors the resource input proportions: $\frac{\alpha_j}{\sum_k^R \alpha_k} = \frac{r_j}{\sum_k^R r_k}, \forall j = 1, \dots, R$ which consists of the local optimal strategy for a given energetic investment ($\sum_k^R \alpha_k$). In the absence of mutations, such population at equilibrium satisfies:

$$e_j^{(1)} = r_j \cdot \sum_{k=1}^R \alpha_k \qquad \forall j = 1, \dots, R. \tag{4}$$

Now consider a mutant that emerges from this population with phenotypes $\vec{\alpha}_{mut} = \vec{\alpha} + \vec{\Delta}$. From (3) and (4) it follows that the selection acting on such mutant is:

$$s_{mut}(\vec{\alpha}, \vec{\Delta}) = \frac{\sum_j^R \Delta_j}{\sum_j^R \alpha_j}. \tag{5}$$

While bigger steps result in stronger selection, equation (5) also implies that the same mutations are subject to weaker selection when emerging on a better adapted background. Thus, this system exhibits diminishing returns epistasis that emerges from the competition dynamics. Because we sample the phenotypic changes from a normal

distribution (see *Methods*), their additive effect will also follow a normal distribution $\sum_j^R \Delta_j \sim N(0, \sigma^*)$ where $\sigma^* = \sigma\sqrt{(R-1)\cdot\rho^2+1}$; therefore, stronger pleiotropic effects lead to stronger selection. From the continuous univariate distribution theory (Johnson *et al.*, 1994) we can retrieve the expected beneficial mutation effect ($E[\sum_j^R \Delta_j]$) and, from (5), compute the corresponding expected selection coefficients ($E[s^+]$) for varying values of $\sum_j^R \alpha_j$ (see Appendix S1 in Supporting Information). Figure 2A shows how the selection strength decreases for better adapted genetic backgrounds across different σ and ρ conditions. It is important to note that the diminishing return epistasis in this model is not due to the energy constraint; nonetheless, such boundary condition truncates the distribution of mutation effects (see Fig. S1A and Appendix S1) and further slows down the rate of adaptation of well adapted populations (inset in Fig. 2A).

Following (5), the proportion of beneficial mutations, computed as

$$P(s^+) = \frac{\int_0^{(1-\sum_j^R \alpha_j)} N(0,\sigma*)}{\int_{-\infty}^{(1-\sum_j^R \alpha_j)} N(0,\sigma*)}$$

drops rapidly in well adapted populations as the energetic trade-off makes the deleterious mutations more common (Fig. S1). Next, we tested how well the analytical approximations – obtained by assuming monomorphic populations at consecutive equilibria – predict the regimes of extensive clonal interference, where the adapting populations are polymorphic and out of equilibrium. To do so, after running simulations with *NU=100* and two resources, we calculate the average population trait sum $\hat{\alpha}(t) := \frac{\sum_i^{M(t)}\left(\alpha_1^{(i)}+\alpha_2^{(i)}\right)\cdot n_i(t)}{N}$, compute the expected beneficial selection coefficient as $E[s^+|\hat{\alpha}] = \frac{E[\Delta_1+\Delta_2|\hat{\alpha}]}{\hat{\alpha}}$ and compare it with the mean beneficial selection observed in the simulations, during the first 300 generations. The strength of selection acting on

239 polymorphic populations follows the predicted diminishing return pattern, but is often
240 underestimated (Fig. 2B). In fact, it decreases with the mean population phenotype $\hat{\alpha}$,
241 but it increases with the population phenotype variance (Fig. S2).

242 Numerical simulations further allow us to link the strength of selection with the speed
243 of phenotypic adaptation: larger phenotypic changes imply stronger selection, resulting
244 in faster adaptation (Fig. 2C). We find that, when time is scaled by $\int_{\hat{\alpha}(t0)}^{1} E[s^{+}|\hat{\alpha}]d\hat{\alpha} \cdot \sigma^{*}$,
245 the populations' mean phenotype moves with similar velocity, demonstrating that both
246 the complex form of selection and the mutation type mediate the speed of adaptation
247 (see Fig. 2D and Appendix S1). Finally, the simulations show how larger mutation rates
248 further accelerate adaptation. Specifically, populations under intense clonal
249 interference ($NU \gg 1$) can approach the phenotypic optimum rapidly (Fig. 2E).

250 In summary, these results show how the availability and the effect of beneficial
251 mutations, together with the complex form of selection, dictate the rate of adaptation
252 that slows down over time and drives the populations from an initially strong to a
253 finally weak selection regime.

254 **Number of coexisting genotypes**

255 During adaptation of an initially monomorphic population, *de novo* mutations can
256 generate polymorphism but at the same time purifying selection tends to reduce such
257 diversity. How many genotypes are generated and maintained under competition for
258 resources? The simulations show that, after a first burst of diversity, the mean number
259 of genotypes first declines and later plateaus (Fig. 3A). This reduction in the mean
260 number of genotypes is due to the energetic constraint; In fact, populations evolving
261 under neutrality or without such boundary do not suffer any decline (Fig. S3). In

contrast, when the populations' phenotypes approach the energetic constraint, beneficial mutations become rarer and selection reduces the number of coexisting genotypes (Fig. S3). Despite the more abundant deleterious mutations, the populations can maintain a dynamic balance between the lineages that are purged and the newly emerging ones (inset of Fig. 3A). The long-term number of genotypes, M^*, averaged across the last 1000 generations, can be fitted by a linear function of the population mutation rate whose slope decreases for stronger selection: $M^* \approx aNU+1$, with an inferred $a= \{8.25\pm0.06, 6.83\pm0.02, 5.48\pm0.02\}$ for $\sigma= \{0.0125,0.025,0.05\}$, respectively (Fig. 3B). Additional simulations, run over longer times (50000 generations), confirmed the long-term plateau of the number of genotypes (M), which can be fitted to a power law decay $M(t)=ct^{-\beta}+d$, with asymptote $d \approx 1$ (see Fig. S4, for parameter values $N=10^7$, $U=10^{-5}$, $\sigma=0.05$, $\rho=0$).

The number of genotypes, M^*, deviates from NU/σ - the expected mean number of deleterious mutations under mutation-selection balance in a simple model of constant negative selection (Haigh, 1978) (see Fig. S5).

Population diversification into ecotypes

In this model, adapting populations consist of a cloud of many genotypes (Fig. 3) and we now characterize the phenotypic structures of these clouds. In trait space, the population optimum Ω (represented by the star in Fig. 1B) has $\sum_j^R \alpha_j^{(\Omega)} = 1$ and all individuals' phenotypes mirroring the resource supply proportions $\left(\alpha_j^{(\Omega)} = \frac{r_j}{r_{tot}}, \forall j = 1, \dots, R\right)$. Such state cannot be invaded by any mutant, thus excluding diversity. However, *Posfai et al.* have shown that, in the absence of *de novo* mutations, large collections of phenotypes can stably coexist when they are distributed around the optimum, if polymorphism and metabolic trade-off are the initial conditions (Posfai,

286 Taillefumier and Wingreen, 2017). Thus, we now ask: if mutation is the only source of
287 variation, will an initially maladapted isogenic population evolve towards a single
288 optimal state or towards distinct states?

289 The simulations show that, due to the stochastic nature of mutation, populations
290 adapting under exactly the same conditions can evolve either one or multiple states
291 (Fig. 4A). Remarkably, the same ancestral genotype can give rise to a better adapted
292 population with many genotypes, all functionally similar (Fig. 4A, left panel), or
293 diversify into different ecotypes: clusters of genotypes with distinct metabolic
294 preferences, capable of coexisting indefinitely (Fig. 4A, right panel).

295 We now investigate the conditions favoring such ecotype diversification, in the regime
296 of strong selection ($Ns \gg 1$). Using the mean shift clustering algorithm (Cheng, 1995),
297 which tests for multimodality, one can group each adapted population into functional
298 clusters (see details in Appendix S2). We find that the proportion of populations that
299 evolved into 1, 2 or more clusters, changes dramatically with NU. Under regimes of
300 more intense clonal interference, distinct ecotypes more probably emerge and coexist.
301 Importantly, when NU is very large, the number of ecotypes may exceed the number of
302 limiting resources (Fig. 4B). The probability of diversification (P) – computed as the
303 proportion of populations that evolved more than one cluster – is close to zero when
304 $NU<1$ but significantly increases when $NU \geq 1$. Such increase can be fitted by a logistic
305 function: $P = \frac{1}{1+e^{a(\ln(NU)-b)}}$ (see Fig. 4C and Tables S1-2). In contrast, the populations do
306 not diversify into clusters under neutrality (Fig. 4C).

307 Both the rate and the type of mutations influence the diversification process. Under
308 intense clonal interference, larger mutation effects (σ) and/or smaller pleiotropic
309 effects (ρ) promote the formation of multiple ecotypes (Fig. 4C, Fig S6 and S7).

310 In summary, during the process of adaptation to improve consumption of the available
311 resources, different outcomes can evolve: i) if the input of new beneficial mutations is
312 low, the recurrently fittest takes over as a cloud of genotypes until the population forms
313 a single generalist ecotype (e.g. Video S1); ii) if *NU* is large, the availability of many
314 beneficial mutations provides the potential for several genotypes to coevolve and
315 distinct ecotypes to stably coexist (e.g. Video S2).

316 **Phenotypic and genotypic diversity within populations**

317 Can we predict ecotypes from the genotypes in the population? Populations composed
318 by a single functional cluster can have considerable genetic diversity if multiple lineages
319 have converged to similar phenotypes (see examples in Fig. S8). We characterized the
320 adapting populations by calculating their average pairwise genetic (π_G) and phenotypic
321 (π_P) distances. Both measures of diversity increase with *NU* and always exceed the
322 neutral case (Fig. 5A). The increase of π_P, but not of π_G, is negatively affected by larger
323 pleiotropic effect (ρ) (ANOVA p-value = 3.2×10^{-10} and 0.09, respectively). Correlating π_P
324 and π_G in the evolved populations, we confirmed that more intense clonal interference
325 fosters functional diversification but stronger pleiotropy constrains it (Fig. S9A-B).
326 We next attempted to predict the existence of multiple ecotypes using the genotypic
327 diversity. The populations that evolved more than one cluster have on average larger
328 π_G. However, it is difficult to establish a threshold above which multiple ecotypes can be
329 predicted (Fig. 5B). The Receiver Operating Characteristic (ROC) curves – obtained via
330 varying thresholds – show the predictive power of π_G and confirm that this improves for
331 those populations that evolved under weaker pleiotropy (Fig. 5C). For example, the
332 thresholds needed to predict ecotypes with 95% specificity are: $\pi_G > \{2.7, 3.6, 4.5\}$ for ρ=
333 $\{0, 0.5, 1\}$, respectively (as shown by the dotted lines in Fig. 5B and the dots in Fig. 5C).

334 A common statistic used to infer population genetic structure is the Tajima's D (Tajima,
335 1989), which compares the number of segregating mutations with the average pairwise
336 genetic distance. The expected value of D is zero under an equilibrium model of
337 mutation and drift, without selection. We note that the time simulated (10^4 generations)
338 is much shorter than that required to achieve a neutral equilibrium at these population
339 sizes (N=10^7), so as expected our results in the absence of selection show a negative
340 Tajima's D, whose mean decreases with *NU* due to larger numbers of newly emerging
341 segregating sites. Under selection, populations with the same *NU* may present different
342 genetic structures (see three examples in Fig. S8) and Tajima's D ranges from very
343 negative to very positive values (Fig. 6). Interestingly, the mean Tajima's D is maximal
344 for intermediate *NU* and it can be close to zero under both low or very large *NU*
345 (compare the black and the gold distributions in Fig. 6). These patterns suggest that
346 other statistics beyond Tajima's D are needed to understand the structure of non-
347 equilibrium populations undergoing strong selection.

348

349 **Discussion**

350 Microbial communities are vital for humans and many other host species (Nicholson *et*
351 *al.*, 2012; Sunagawa *et al.*, 2015). Emerging observations of evolution in such
352 ecosystems (Barroso-Batista *et al.*, 2014; Garud *et al.*, 2019; Zhao *et al.*, 2019) motivate
353 new theories where the mechanisms that generate diversity involve complex forms of
354 selection and clonal interference (Gordo, 2019). We propose that an eco-evolutionary
355 model of resource competition, describing the mechanisms behind ecological
356 divergence, can help understand diversification within ecosystems. Our framework can
357 be generalized to incorporate other evolutionary mechanisms, including transmission

358 and horizontal gene transfer, and can serve as a bridge between ecology and population
359 genetics.

360 Population genetics models of clonal interference have greatly advanced our
361 understanding on adaptation under this condition (Gerrish and Lenski, 1998; Park and
362 Krug, 2007; Good *et al.*, 2012; de Sousa *et al.*, 2016). However, clonal interference is
363 rarely considered in theoretical studies of ecosystems (Farahpour *et al.*, 2018), even
364 though it greatly impacts the evolution of microbes within communities (Barroso-
365 Batista *et al.*, 2014) and may be relevant in key ecosystems such as the human
366 microbiota (Zhao *et al.*, 2019). Commensal species in the gut have large population sizes
367 $\sim 10^8$ cells/g; If each bacterium mutates in the gut as it does in the laboratory (Drake,
368 1991), then each gram of material will host around 10^5 new mutant cells every
369 generation. Even if only 0.1% brings up a benefit (Perfeito *et al.*, 2007), clonal
370 interference still extensively affects the gut microbiota dynamics.

371 Here we have studied an ecological model where clones do not compete for fixation but
372 for resources. Modeling competition explicitly allows to make testable predictions
373 about diversity as both the traits and genomes can now easily be studied. We show that
374 clonal interactions can drive an initial monomorphic population to polymorphism with
375 distinct ecotypes, deviating from the simple expectation of adapting to a single optimum
376 phenotype (Fig. 4).

377 Taking the MacArthur model, *Posfai et al.* have demonstrated that metabolic trade-offs
378 promote coexistence of more species than resource types, thus overcoming the
379 competitive exclusion principle in the absence of evolution (Posfai, Taillefumier and
380 Wingreen, 2017). Extensions of this framework already demonstrated its power in
381 recapitulating experimental results from studies of soil, plant (Goldford *et al.*, 2018) or

382 mammalian gut microbiotas (Leónidas Cardoso *et al.*, 2020). In addition, a similar
383 framework provided analytical descriptions of how populations adapt under
384 competition for resources and demonstrated that directional selection can limit
385 ecological diversification (Good, Martis and Hallatschek, 2018). This pattern is also
386 observed in our simulations; In fact, stronger pleiotropy, which causes stronger
387 directional selection, limits the emergence of clusters (Fig. 4C). Our work differs from
388 the latter by focusing on clonal interference and by introducing an energetic trade-off.
389 We show that under clonal interference, the outcome of phenotypic adaptation is
390 probabilistic, whereby the populations can evolve paths that would be impossible to
391 observe through equilibrium assumptions.

392 Trade-offs are commonly assumed and expected to affect evolutionary trajectories
393 (Farahpour *et al.*, 2018; Amado and Campos, 2019), but this is not always observed.
394 While many empirical results have confirmed the role of trade-offs during adaptation
395 (Bell and Reboud, 1997; Bull, Badgett and Wichman, 2000; Turner and Elena, 2000;
396 Dykhuizen and Dean, 2004; Greene *et al.*, 2005; Duffy, Turner and Burch, 2006; Coffey *et*
397 *al.*, 2008; Ward, Perron and MacLean, 2009; Bailey and Kassen, 2012; Li, Petrov and
398 Sherlock, 2019), others did not find evidence for any (Reboud and Bell, 1997; Kassen
399 and Bell, 1998; Turner and Elena, 2000; Trindade *et al.*, 2009; Bedhomme, Lafforgue
400 and Elena, 2012). Here we assumed a trade-off in the form of an energetic constraint,
401 which only affects well adapted genotypes. Thus, in our model the observation of a
402 trade-off depends on the time at which it is measured. It would be interesting to test for
403 trade-offs at different times during adaptation, as this could explain some of the
404 contrasting findings outlined above. Compatible with this hypothesis, a trade-off in
405 Escherichia coli ability to grow in the presence of both glucose and lactose was found,

406 but it only emerged after a period of constraint-free adaptation (Satterwhite and
407 Cooper, 2015). We find that the metabolic trade-off in the resource consumption is not
408 required for the formation of distinct ecotypes but it promotes their stable coexistence
409 (Fig. 4, S8 and Video S2).

410 We introduced mutation in a model of competition for resources to couple ecology and
411 evolution. This framework allowed some analytical approximations of the selection that
412 underlies adaptation. Consistent with empirical data, competition-driven selection is
413 characterized by diminishing return epistasis - the benefit decline in populations with
414 higher mean fitness (Chou *et al.*, 2011; Kryazhimskiy *et al.*, 2014; Schoustra *et al.*, 2016;
415 Wünsche *et al.*, 2017).

416 In many ecosystems coexisting types seem to outnumber the limiting resources, and,
417 solving this contention has motivated numerous studies. Previous eco-evolutionary
418 analysis suggest that adding evolutionary changes confirms (Edwards *et al.*, 2018) or
419 even exacerbates (Shoresh, Hegreness and Kishony, 2008) this paradox. Perhaps
420 surprisingly, our simulations show that large mutational inputs maintain a dynamically
421 stable number of types that overcome the competitive exclusion (Fig. 3). And at the
422 functional level, we show that diversity generally respects the exclusion principle
423 (number of types ≤ number of resources) but with exceptions: in a regime of strong
424 clonal interference, the number of extant ecotypes can even exceed the number of
425 limiting resources (Fig. 4B). The process leading to such ecological diversification is
426 strongly influenced by the underlying molecular parameters: regimes of low clonal
427 interference (NU <1) lead to the evolution of a single generalist population, but large
428 mutational inputs ($NU \geq 1$) can lead to the formation of two or more differentially
429 specialized ecotypes (Fig. 4 and Video S1-S2).

430 Our analysis demonstrates that both large mutation effects and weak pleiotropy foster
431 ecological diversification (Fig. 4C). Different pleiotropic effects are meant to represent
432 different interactions between the traits under selection. If the available resources are
433 similar (e.g. chemical composition) and/or the metabolic processes involved in their
434 consumption share many genes, this could increase the chances that a mutation affects
435 the two traits simultaneously leading to large pleiotropy. Contrarily, less related
436 resources could involve more independent effects leading to smaller pleiotropy,
437 promoting diversification (Fig. 4C). This interpretation could explain why adaptive
438 diversification occurred in some experimental evolution setups (Friesen *et al.*, 2004;
439 Sandberg *et al.* 2017) but was not observed in others (Satterwhite and Cooper, 2015;
440 Sandberg *et al.* 2017). In agreement with this hypothesis, Sandberg and colleagues
441 showed that evolving on less metabolically related resources promoted ecological
442 diversification (Sandberg *et al.* 2017).

443 Sympatric diversification can be observed experimentally and predicted by theoretical
444 models (Friesen *et al.*, 2004). The framework of adaptive dynamics has been extensively
445 used in this context as it describes evolution on fitness landscapes that change
446 dynamically due to frequency-dependent interactions (Geritz *et al.*, 1998; Doebeli,
447 2011). While very useful to understand diversification, adaptive dynamics are based on
448 equilibrium assumptions: the populations first evolve to an equilibrium state before
449 diversification occurs, as explained by the concept of evolutionary branching points.
450 Our approach in contrast follows adaptation through an individual-based model that
451 undergoes strong non-equilibrium dynamics. Previous studies (Rosindell, Harmon and
452 Etienne, 2015; Ispolatov, Madhok and Doebeli, 2016) highlighted how considering
453 evolution at the individual level is necessary to fully understand the adaptation process.

454 We find that the genetic diversity of a population can be used to predict the underlying
455 phenotypic structure, but with limitations. The accuracy of such prediction decreases if
456 mutations are not constrained and can affect equally all the traits under selection (Fig.
457 5). Other statistics based on genetic data (such as the Tajima's D) can strongly deviate
458 from the expected values if the populations are not yet at (evolutionary) equilibrium
459 and can be hard to interpret if the underlying populations' parameters are unknown.

460 The model studied here can be considered a first step of integration. It predicts that the
461 typical high mutational input of bacterial species and cancer cells, coupled with an
462 energetic constraint, is a mechanism capable of generating functionally diverse clonal
463 communities. Future frameworks addressing microbial ecology and evolution will need
464 to address how space, migration and/or fluctuating conditions affect the patterns of
465 diversity observed here. In addition, cooperation between genotypes (such as cross-
466 feeding or use of costly public goods) have the potential to shape diversity levels and
467 should be investigated in future studies.

468

469 **Acknowledgements**

470 The authors acknowledge C Bank and the members of the Evolutionary Dynamics and
471 Evolutionary Biology labs of the IGC for their assistance throughout the development of
472 this work; PR Campos, T Paixão, S Miller, G Sgarlata and R Ramiro for their comments
473 on the manuscript.
474 This work was supported by Portuguese Science Foundation (FCT) Grant
475 PTDC/BIA-EVL/31528/2017; Deutsche Forschungs-gemeinschaft (DFG) Grant
476 SFB 1310; and a cooperation agreement between University of

Cologne and Gulbenkian Institute. MA was supported by FCT Grant PD/BD/138735/2018.

References

Abrams, P. A. (1988) 'How should resources be counted?', *Theoretical Population Biology*. doi: 10.1016/0040-5809(88)90014-7.

Amado, A. and Campos, P. R. A. (2019) 'Ecological specialization under multidimensional trade-offs', *Evolutionary Ecology*. Springer International Publishing, 33(6), pp. 769–789. doi: 10.1007/s10682-019-10013-4.

Mac Arthur, R. (1969) 'Species Packing, and What Competition Minimizes', *Proceedings of the National Academy of Sciences*, 64(4), pp. 1369–1371. doi: 10.1073/pnas.64.4.1369.

Bailey, S. F. and Kassen, R. (2012) 'Spatial structure of ecological opportunity drives adaptation in a bacterium', *American Naturalist*. doi: 10.1086/666609.

Barroso-Batista, J. *et al.* (2014) 'The First Steps of Adaptation of Escherichia coli to the Gut Are Dominated by Soft Sweeps', *PLoS Genetics*. doi: 10.1371/journal.pgen.1004182.

Barroso-Batista, J. *et al.* (2020) 'Specific Eco-evolutionary Contexts in the Mouse Gut Reveal Escherichia coli Metabolic Versatility', *Current Biology*. doi: 10.1016/j.cub.2020.01.050.

Bedhomme, S., Lafforgue, G. and Elena, S. F. (2012) 'Multihost experimental evolution of a plant RNA virus reveals local adaptation and host-specific mutations', *Molecular Biology and Evolution*. doi: 10.1093/molbev/msr314.

Bell, G. and Reboud, X. (1997) 'Experimental evolution in Chlamydomonas II. Genetic variation in strongly contrasted environments', *Heredity*. doi: 10.1038/hdy.1997.78.

Bull, J. J., Badgett, M. R. and Wichman, H. A. (2000) 'Big-benefit mutations in a bacteriophage inhibited with heat', *Molecular Biology and Evolution*. doi: 10.1093/oxfordjournals.molbev.a026375.

Cheng, Y. (1995) 'Mean Shift, Mode Seeking, and Clustering', *IEEE Transactions on Pattern Analysis and Machine Intelligence*. doi: 10.1109/34.400568.

Chou, H. H. *et al.* (2011) 'Diminishing returns epistasis among beneficial mutations decelerates adaptation', *Science*, 332(6034), pp. 1190–1192. doi: 10.1126/science.1203799.

Coffey, L. L. *et al.* (2008) 'Arbovirus evolution in vivo is constrained by host alternation', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0712130105.

Desai, M. M., Fisher, D. S. and Murray, A. W. (2007) 'The Speed of Evolution and Maintenance of Variation in Asexual Populations', *Current Biology*. doi: 10.1016/j.cub.2007.01.072.

Dieckmann, U. and Doebeli, M. (1999) 'On the origin of speciesby sympatric speciation', *Science*, 400(July), pp. 0–3. doi: 10.1038/22521.

Doebeli, M. (2011) *Adaptive diversification (MPB-48)*, *Adaptive Diversification (MPB-48)*. doi: 10.23943/princeton/9780691128931.001.0001.

Drake, J. W. (1991) 'A constant rate of spontaneous mutation in DNA-based microbes', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.88.16.7160.

Duffy, S., Turner, P. E. and Burch, C. L. (2006) 'Pleiotropic costs of niche expansion in the RNA bacteriophage Φ6', *Genetics*. doi: 10.1534/genetics.105.051136.

Dykhuizen, D. E. and Dean, A. M. (2004) 'Evolution of specialists in an experimental microcosm', *Genetics*. doi: 10.1534/genetics.103.025205.

Edwards, K. F. *et al.* (2018) 'Evolutionarily stable communities: a framework for understanding the role of trait evolution in the maintenance of diversity', *Ecology Letters*, 21(12), pp. 1853–1868. doi: 10.1111/ele.13142.

Farahpour, F. *et al.* (2018) 'Trade-off shapes diversity in ecoevolutionary dynamics', *eLife*, 7. doi: 10.7554/eLife.36273.

Fisher, R. A. (1958) 'The genetical theory of natural selection, 2nd edn.', *Dover Publication, New York*. doi: 10.1111/jeb.12566.

Friesen, M. L. *et al.* (2004) 'Experimental evidence for sympatric ecological diversification due to frequency-dependent competition in Escherichia coli', *Evolution*, 58(2), pp. 245–260. doi: 10.1111/ele.12430.

Garud, N. R. *et al.* (2019) 'Evolutionary dynamics of bacteria in the gut microbiome within and across hosts', *PLoS biology*, 17(1), p. e3000102. doi: 10.1371/journal.pbio.3000102.

Geritz, S. A. H. *et al.* (1998) 'Evolutionary singular strategies and the adaptive growth and branching of the evolutionary tree', *Evolutionary Ecology*, 12, pp. 35–57.

Goldford, J. E. *et al.* (2018) 'Emergent simplicity in microbial community assembly', *Science*. doi: 10.1126/science.aat1168.

Good, B. H. *et al.* (2012) 'Distribution of fixed beneficial mutations and the rate of adaptation in asexual populations', *Proceedings of the National Academy of Sciences of the United States of America*, 109(13), pp. 4950–4955. doi: 10.1073/pnas.1119910109.

Good, B. H., Martis, S. and Hallatschek, O. (2018) 'Adaptation limits ecological diversification and promotes ecological tinkering during the competition for substitutable resources', *Proceedings of the National Academy of Sciences of the United States of America*, 115(44), pp. E10407–E10416. doi: 10.1073/pnas.1807530115.

Gordo, I. (2019) 'Evolutionary change in the human gut microbiome: From a static to a dynamic view', *PLoS Biology*, 17(2), pp. 1–5. doi: 10.1371/journal.pbio.3000126.

Greene, I. P. *et al.* (2005) 'Effect of Alternating Passage on Adaptation of Sindbis Virus to Vertebrate and Invertebrate Cells', *Journal of Virology*. doi: 10.1128/jvi.79.22.14253-14260.2005.

Gresham, D. *et al.* (2008) 'The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast', *PLoS Genetics*. doi: 10.1371/journal.pgen.1000303.

Haigh, J. (1978) 'The accumulation of deleterious genes in a population-Muller's Ratchet', *Theoretical Population Biology*, 14(2), pp. 251–267. doi: 10.1016/0040-5809(78)90027-8.

Hardin, G. (1960) 'The Competitive Exclusion Principle', *Science*, 131(November).

Herron, M. D. and Doebeli, M. (2013) 'Parallel Evolutionary Dynamics of Adaptive Diversification in Escherichia coli', *PLoS Biology*. doi: 10.1371/journal.pbio.1001490.

Huston, M. A. (1994) 'Biological diversity: the coexistence of species on changing landscapes', *Biological diversity: the coexistence of species on changing landscapes*. doi: 10.2307/2404662.

Hutchinson, G. E. (1961) 'THE PARADOX OF THE PLANKTON * New Haven , Connecticut Osborn Zoological Laboratory , The problem that I wish to discuss in the present contribution is raised by the very paradoxical situation of the plankton , particularly the phyto- plankton , of relativ', XCV(882), pp. 137–145.

Ispolatov, I., Madhok, V. and Doebeli, M. (2016) 'Individual-based models for adaptive diversification in high-dimensional phenotype spaces', *Journal of Theoretical Biology*. Elsevier, 390, pp. 97–105. doi: 10.1016/j.jtbi.2015.10.009.

Johnson, Norman L *et al.* (1994) 'Continuous Univariate Distributions', *Wiley*, 1, p. Section 10.1.

Kassen, R. and Bell, G. (1998) 'Experimental evolution in Chlamydomonas. IV. Selection in environments that vary through time at different scales', *Heredity*. doi: 10.1046/j.1365-2540.1998.00329.x.

Kinnersley, M. A., Holben, W. E. and Rosenzweig, F. (2009) 'E unibus plurum: Genomic analysis of an experimentally evolved polymorphism in Escherichia coli', *PLoS Genetics*. doi: 10.1371/journal.pgen.1000713.

Kremer, C. T. and Klausmeier, C. A. (2017) 'Species packing in eco-evolutionary models of seasonally fluctuating environments', *Ecology Letters*. doi: 10.1111/ele.12813.

Kryazhimskiy, S. *et al.* (2014) 'Global epistasis makes adaptation predictable despite sequence-level stochasticity', *Science*, 344(6191), pp. 1519–1522. doi: 10.1126/science.1250939.

Lawrence, D. *et al.* (2012) 'Species interactions alter evolutionary responses to a novel environment', *PLoS Biology*. doi: 10.1371/journal.pbio.1001330.

Leónidas Cardoso, L, Durão, P, Amicone, M and Gordo, I (2020) 'Dysbiosis individualizes the fitness effect of antibiotic resistance in the mammalian gut', *Nature Ecology and Evolution*. Springer US, 4(9), pp. 1268–1278. doi: 10.1038/s41559-020-1235-1.

Li, Y., Petrov, D. A. and Sherlock, G. (2019) 'Single nucleotide mapping of trait space reveals Pareto fronts that constrain adaptation', *Nature Ecology and Evolution*. Springer US, 3(11), pp. 1539–1551. doi: 10.1038/s41559-019-0993-0.

Litchman, E. and Klausmeier, C. A. (2001) 'Competition of phytoplankton under fluctuating light', *American Naturalist*. doi: 10.1086/318628.

Maharjan, R. *et al.* (2006) 'Clonal adaptive radiation in a constant environment', *Science*. doi: 10.1126/science.1129865.

Nicholson, J. K. *et al.* (2012) 'Host-gut microbiota metabolic interactions', *Science*. doi: 10.1126/science.1223813.

de Oliveira, V. M., Amado, A. and Campos, P. R. A. (2018) 'The interplay of tradeoffs within the framework of a resource-based modelling', *Ecological Modelling*. Elsevier, 384(June), pp. 249–260. doi: 10.1016/j.ecolmodel.2018.06.026.

Pacciani-Mori, L. *et al.* (2020) 'Dynamic metabolic adaptation can promote species coexistence in competitive microbial communities', *PLoS Computational Biology*, 16(5), pp. 1–18. doi: 10.1371/journal.pcbi.1007896.

Park, S. C. and Krug, J. (2007) 'Clonal interference in large populations', *Proceedings of the National Academy of Sciences of the United States of America*, 104(46), pp. 18135–18140. doi: 10.1073/pnas.0705778104.

Perfeito, L. *et al.* (2007) 'Adaptive mutations in bacteria: High rate and small effects', *Science*, 317(5839), pp. 813–815. doi: 10.1126/science.1142284.

Philip J. Gerrish & Richard E. Lenski (1998) 'The fate of competing beneficial mutations in an asexual population', *Genetica*, 102/103, pp. 127–144,.

Posfai, A., Taillefumier, T. and Wingreen, N. S. (2017) 'Metabolic Trade-Offs Promote Diversity in a Model Ecosystem', *Physical Review Letters*, 118(2), pp. 1–5. doi: 10.1103/PhysRevLett.118.028103.

Reboud, X. and Bell, G. (1997) 'Experimental evolution in Chlamydomonas III. Evolution of specialist and generalist types in environments that vary in space and time', *Heredity*. doi: 10.1038/hdy.1997.79.

Rosindell, J., Harmon, L. J. and Etienne, R. S. (2015) 'Unifying ecology and macroevolution with individual-based theory', *Ecology Letters*, 18(5), pp. 472–482. doi: 10.1111/ele.12430.

Satterwhite, R. S. and Cooper, T. F. (2015) 'Constraints on adaptation of Escherichia coli to mixed-

resource environments increase over time', *Evolution*, 69(8), pp. 2067–2078. doi: 10.1111/evo.12710.

Schoener, T. W. (2011) 'The Newest Synthesis: Understanding Ecological Dynamics', *Science*, 331(January), pp. 426–429. doi: 10.1126/science.1193954.

Schoustra, S. *et al.* (2016) 'Diminishing-returns epistasis among random beneficial mutations in a multicellular fungus', *Proceedings of the Royal Society B: Biological Sciences*, 283(1837). doi: 10.1098/rspb.2016.1376.

Shoresh, N., Hegreness, M. and Kishony, R. (2008) 'Evolution exacerbates the paradox of the plankton', *Proceedings of the National Academy of Sciences of the United States of America*, 105(34), pp. 12365–12369. doi: 10.1073/pnas.0803032105.

de Sousa, J. A. M. *et al.* (2016) 'Competition and fixation of cohorts of adaptive mutations under Fisher geometrical model', *PeerJ*, 2016(2), pp. 1–17. doi: 10.7717/PEERJ.2256.

Sunagawa, S. *et al.* (2015) 'Structure and function of the global ocean microbiome', *Science*. doi: 10.1126/science.1261359.

T. E. Sandberg, C. J. Lloyd, B..O. Palsson, A. M. F. (2017) 'Laboratory Evolution to Alternating Substrate Environments Yields Distinct Phenotypic and Genetic Adaptive Strategies', 83(13), pp. 1–15.

Tajima, F. (1989) 'Statistical method for testing the neutral mutation hypothesis by DNA polymorphism', *Genetics*.

Tilman, D. (1982) 'Resource competition and community structure.', *Monographs in population biology*. doi: 10.2307/4549.

Trindade, S. *et al.* (2009) 'Positive epistasis drives the acquisition of multidrug resistance', *PLoS Genetics*. doi: 10.1371/journal.pgen.1000578.

Turner, P. E. and Elena, S. F. (2000) 'Cost of host radiation in an RNA virus', *Genetics*.

Ward, H., Perron, G. G. and MacLean, R. C. (2009) 'The cost of multiple drug resistance in Pseudomonas aeruginosa', *Journal of Evolutionary Biology*. doi: 10.1111/j.1420-9101.2009.01712.x.

Wright, S. (1930) 'The genetical theory of natural selection: A review', *Journal of Heredity*. doi: 10.1093/oxfordjournals.jhered.a103361.

Wünsche, A. *et al.* (2017) 'Diminishing-returns epistasis decreases adaptability along an evolutionary trajectory', *Nature Ecology and Evolution*, 1(4), pp. 1–6. doi: 10.1038/s41559-016-0061.

Zhao, S. *et al.* (2019) 'Adaptive Evolution within Gut Microbiomes of Healthy People', *Cell Host and Microbe*. Elsevier Inc., 25(5), pp. 656-667.e8. doi: 10.1016/j.chom.2019.03.007.

Figure legends

Figure 1. **Ecological dynamics and individual-based evolutionary processes**. **A)** Illustration of the eco-evolutionary dynamics. Bacterial genotypes, represented by circles of different colors, grow according to the resources' concentration (squares and semicircles) and their own phenotypic traits, represented by the enzyme-like structures on the cell wall. Mutation events (red arrow) can generate new types whose fate will depend on drift and selection. **B)** Phenotypic space and mutation process. An initially maladapted monomorphic population (green circle) can acquire de-novo mutations according to

the given assumptions, as explained in the inset. **C)** Examples of different clonal interference regimes, NU << 0.1 above, NU >>1 below. Each color represents a different mutant type.

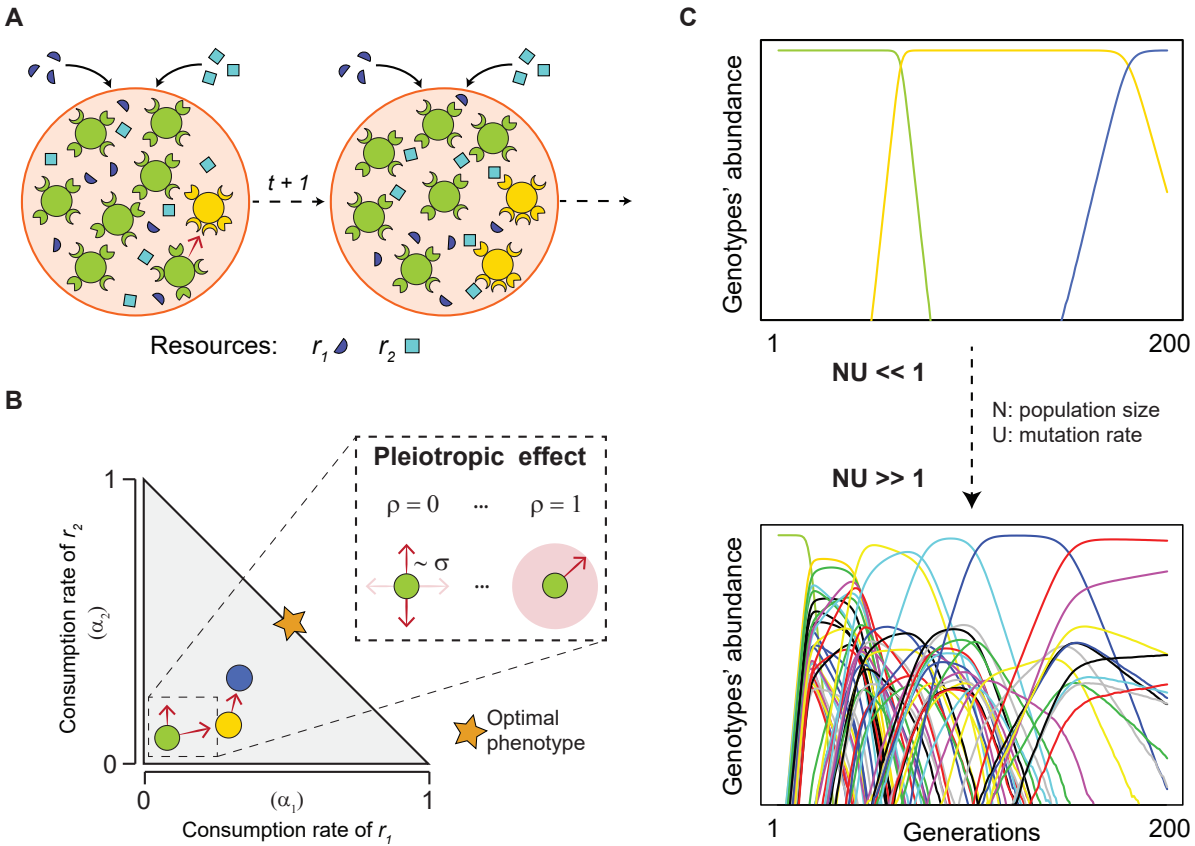
Figure 2. Diminishing return epistasis and the adaptation rate. A) Analytical predictions of the diminishing return epistasis in monomorphic populations. The inset shows the effect of the energy constraint. **B)** Dots represent the mean of the observed positive selection during the first 300 generations of adaptation. The dotted line represents the analytical expectation at each generation t, given the average population trait sum $\left(\hat{\alpha}(t) := \frac{\sum_i^{M(t)} n_i(t)(\alpha_1^{(i)}+\alpha_2^{(i)})}{N}\right)$. Each color represents an independent simulated population, which adapted with $\sigma = 0.05$, $\rho = 0.5$, $N = 10^7$ and U= 10^{-5}. **C)** The population average trait sum $\hat{\alpha}(t)$ is shown as proxy of adaptation under different σ and ρ conditions. Other parameters: $R = 2$, $N = 10^7$, $U = 10^{-8}$. Lines are the averages over 100 simulations and the shaded areas represent the confidence interval. **D)** Same dynamics as in C, but on a different scale: Generations · I · σ*, where I and σ* are defined in the figure. **E)** Adaptation across different mutational inputs (color coded) and the expected beneficial effect of mutations at the end of the adaptation process (generation 10000). The inset panel highlights the dynamics over the last 100 generations. Points in the right panel are represented by the mean +/- one standard deviation in *log-log* scale. Other parameters for panel E: $N = 10^7$, $\sigma = 0.05$, $\rho = 0.5$.

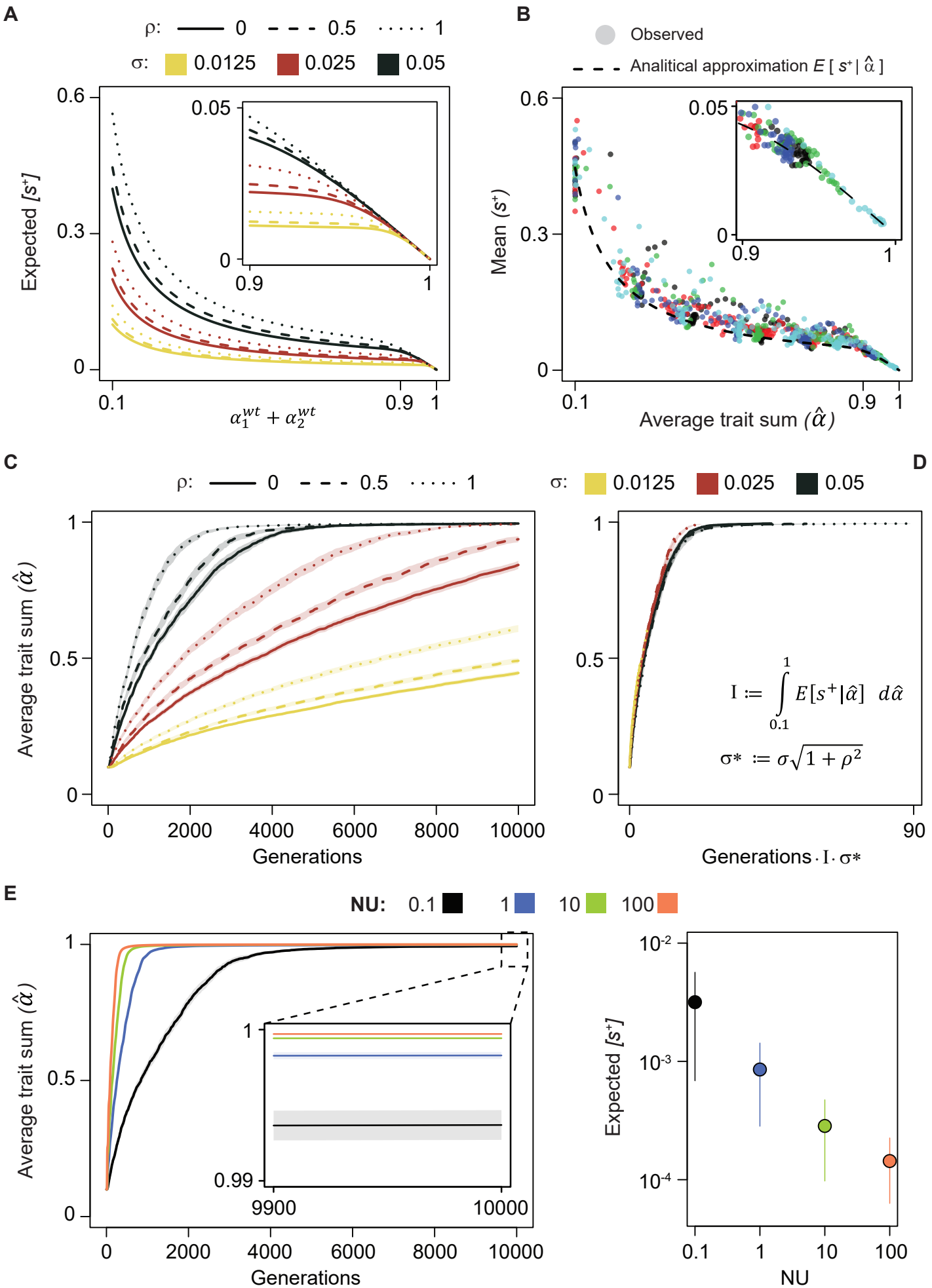
Figure 3. Genotypes' dynamics and the mutation-selection balance under competition for two resources. A) Number of genotypes present in the environment over time, under neutrality (diamonds) or under selection (lines). Lines represent the average across 100 populations and the shaded area their confidence interval. The populations eventually plateau and dynamically maintain large number of genotypes (M^*). Other parameters: $\sigma = 0.05$, $\rho = 0$ and $N = 10^7$. **B)** Long-lasting number of genotypes, computed as the average over the last 1000 generations (e.g. dotted line in the right panel of A). The lines represent the linear regressions with model: $M^* = aNU + 1$ and NU: $\{0.1, 0.5, 1, 5, 10, 50, 100, 500\}$ where $a = \{8.25 \pm 0.06, 6.83 \pm 0.02, 5.48 \pm 0.02\}$ for $\sigma = \{0.0125, 0.025, 0.05\}$, respectively. Both axes are represented in *log* scale with ticks every $\{1, \ldots, 9\} \cdot 10^x$. Other parameters: $\rho = 0$ and $N = 10^7$.

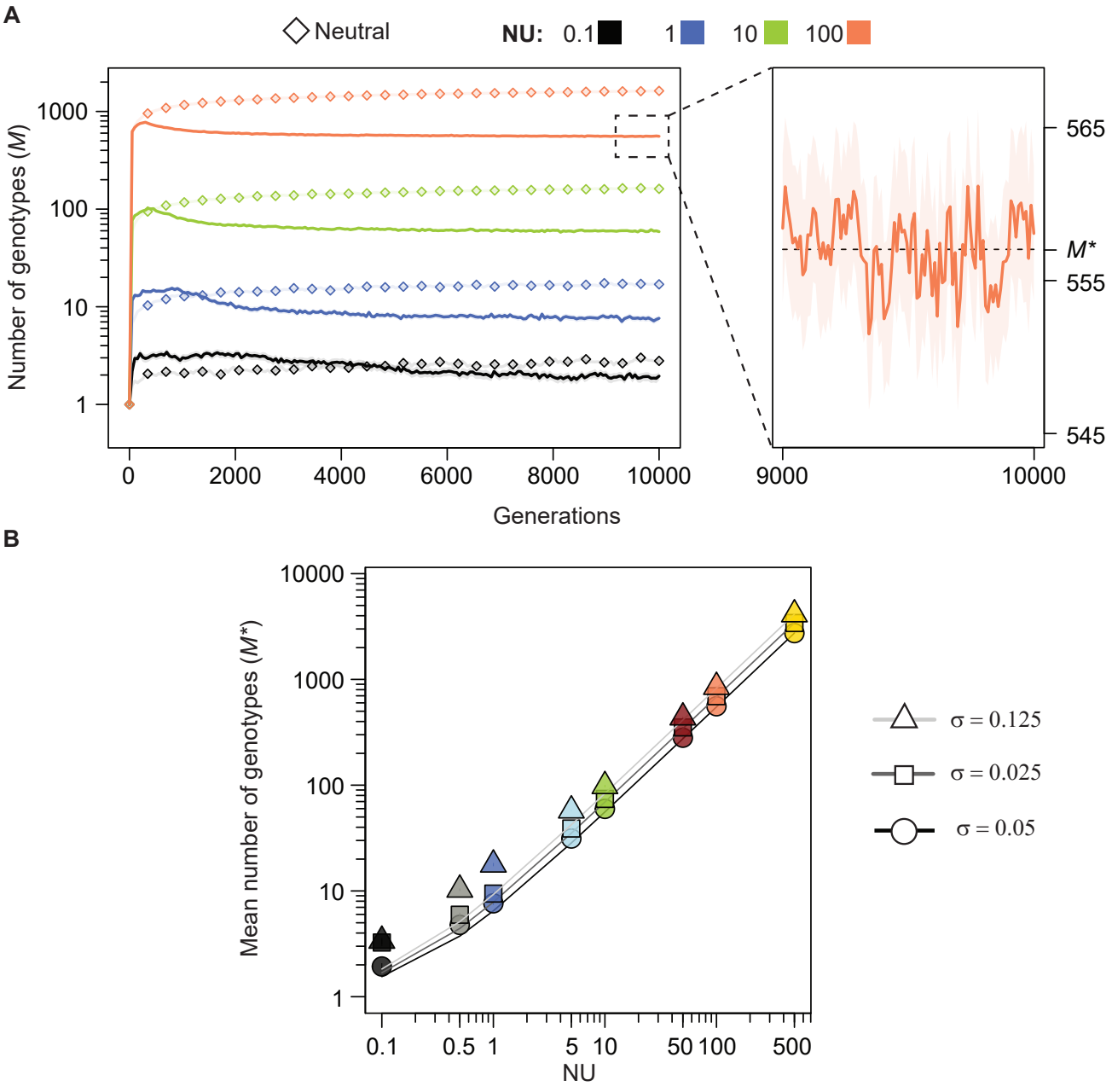
Figure 4. Ecological diversification under competition for two resources. A) Two example populations evolving under the same conditions ($N = 10^7, U = 10^{-5}, \rho = 0.5, \sigma = 0.05, \mathrm{R} = 2$). The phenotypes and the preference distributions show one population that has evolved into a single optimal cluster (squares) and another population that gave rise to a stable diverse community composed by two clusters (circles). Lines connecting the shapes represent mutations. **B)** Counts of populations that evolved into 1,2 or more phenotypic clusters. Here, $N = 10^7, \sigma = 0.05, \mathrm{R} = 2, \rho = 0$. **C)** Populations diversify with a probability that increases with $\ln(NU)$ and σ but decreases with ρ. The lines represent the fit of the data to the logistic function: $P = \frac{1}{1+e^{a(\ln(NU)-b)}}$, NU: $\{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 1, 5, 10, 50, 100, 500\}$. The inferred parameters a and b are reported in Tables S1-2 and the full set of data is shown in Fig. S6. In the left plot: $\rho = 0$, while in the right one: $\sigma = 0.05$. The probabilities were computed as proportions out of 100 independent populations and their 95% confidence interval by normal approximation: $P \pm z\sqrt{\frac{P(1-P)}{100}}$, $z = 1.96$.

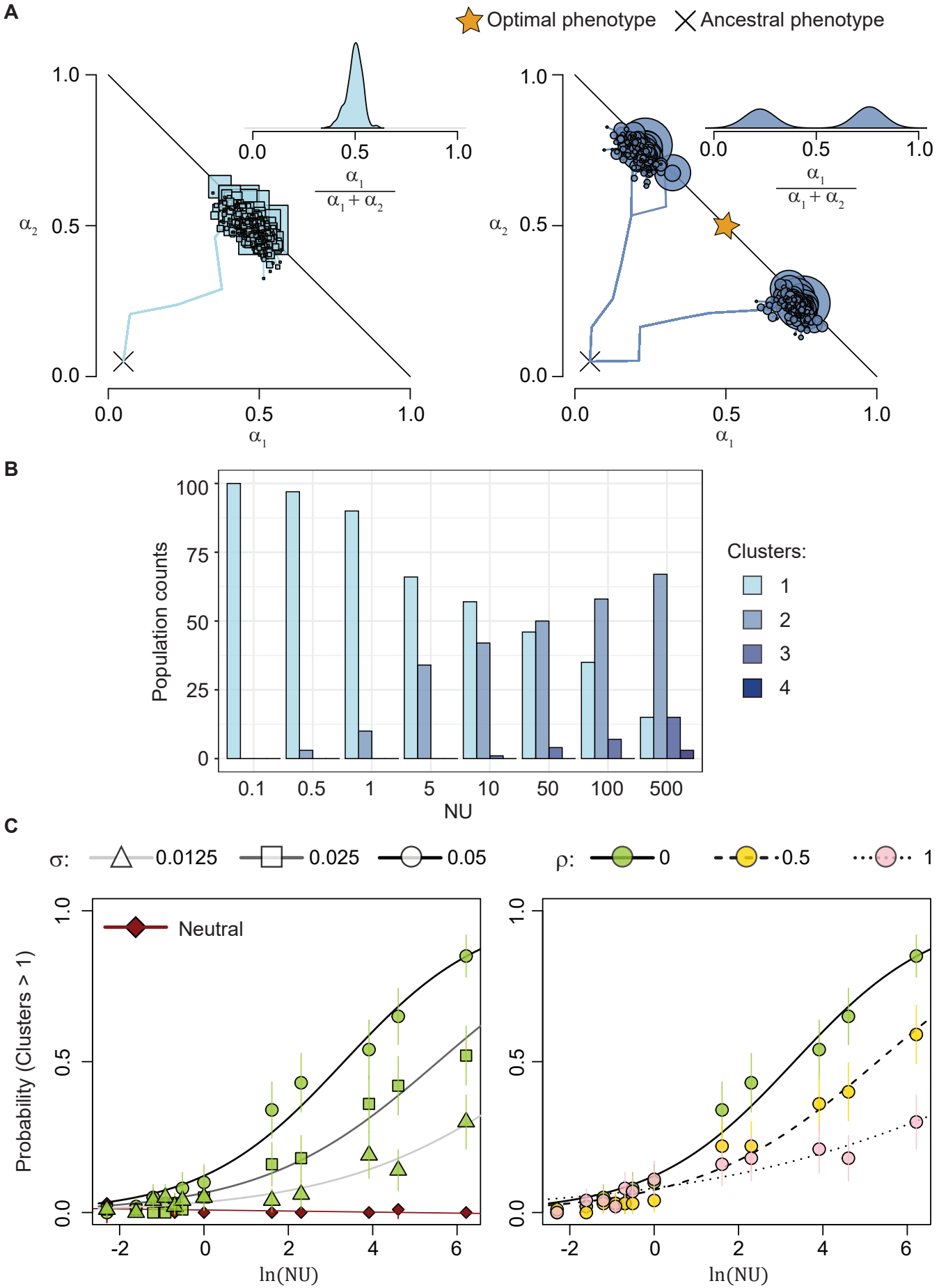
Figure 5. Diversity within and between populations and ecotypes' prediction. A) Average pairwise genotypic (π_G) and phenotypic (π_P) diversity were measured within each population, as defined in the *Methods*. 100 independent populations were simulated under the conditions specified on the x axis and by the colors. Other parameters: σ = 0.05, R=2 and $N = 10^7$. **B)** π_G distributions of the populations that evolved in more than one cluster (dark blue) or in a single cluster (light blue). Data form panel A with different NU were pulled together for a total of 800 populations per condition (ρ). The violin plots in A and B show the distribution of the data and their median. The dotted lines represent the π_G threshold that would ensure 95% specificity in ecotype prediction (i.e. 5% false positive) and correspond to the circles in panel C. **C)** π_G was used to predict whether each population is composed by one or multiple ecotypes. The ROC curves represent the sensitivity over one minus specificity of the prediction outcomes, for varying thresholds. The dotted line and the circles represent the thresholds that ensure 5% false positive rate and these are: $\pi_G > 2.7, 3.6$ or 4.5 for ρ=0, 0.5 or 1, respectively. The area under the curve (AUC) is reported on the figure. * or *** indicate p-value <0.01 or <0.0001, respectively.

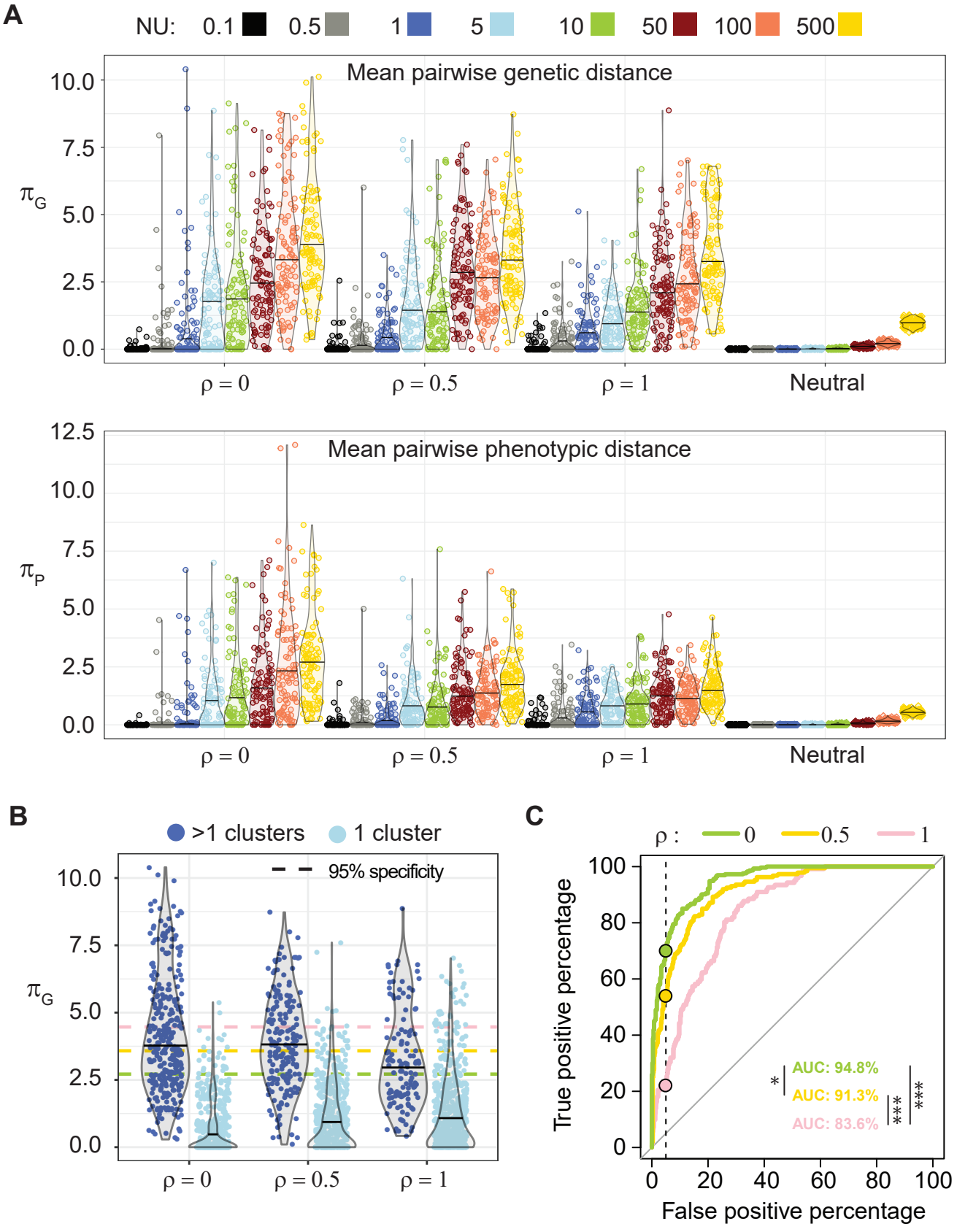
Figure 6. **Tajima's D within and between populations.** The Tajima's D statistics was computed for each population at the end of the adaptation (generation 10000) from a sample of $m = 100$ genotypes. Samples under selection (empty circles) are compared with samples under neutrality (full diamonds). Each data point is an independent population. Those populations whose sample did not present polymorphism, are not represented (e.g. black diamonds are missing) because the Tajima's D is not defined as the number of segregating sites would be zero. The corresponding count of monomorphic samples is given in the lower panel. Other parameters: $N = 10^7, \rho = 0, \sigma = 0.05, R = 2$. The violin plots show the distributions of the data and their median.

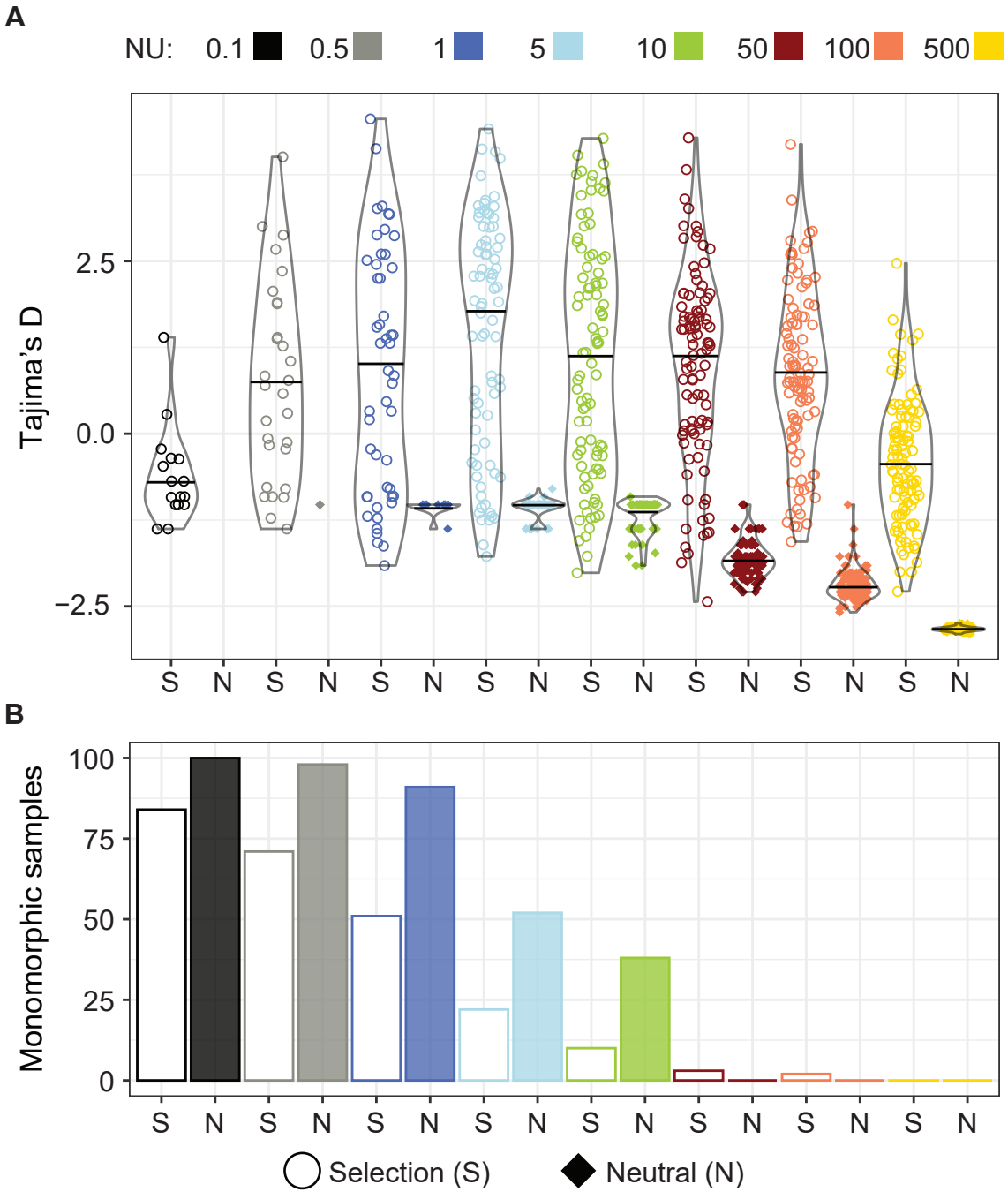

A
t + 1
Resources: r_1 r_2
B
Pleiotropic effect
ρ = 0 ... ρ = 1
~ σ
...
Optimal phenotype
Consumption rate of r_2 (α_2)
Consumption rate of r_1 (α_1)
1
0
0
1
C
Genotypes' abundance
1
200
NU << 1
N: population size
U: mutation rate
NU >> 1
Genotypes' abundance
1
Generations
200


A
ρ: 0 0.5 1
σ: 0.0125 0.025 0.05
Expected [s+]
0.6
0.3
0
0.05
0
0.9
1
0.1
0.9
1
$\alpha_1^{wt} + \alpha_2^{wt}$
B
Observed
Analitical approximation E [s+ | α̂]
Mean (s+)
0.6
0.3
0
0.05
0
0.9
1
0.1
0.9
1
Average trait sum (α̂)
C
ρ: 0 0.5 1
Average trait sum (α̂)
1
0.5
0
0
2000
4000
6000
8000
10000
Generations
D
σ: 0.0125 0.025 0.05
1
0.5
0
$I := \int_{0.1}^{1} E[s^+|\hat{\alpha}] \; d\hat{\alpha}$
$\sigma* := \sigma\sqrt{1+\rho^2}$
0
90
Generations · I · σ*
E
NU: 0.1 1 10 100
Average trait sum (α̂)
1
0.5
0
1
0.99
9900
10000
0
2000
4000
6000
8000
10000
Generations
Expected [s+]
10^{-2}
10^{-3}
10^{-4}
0.1
1
10
100
NU


A
Neutral
NU: 0.1
1
10
100
Number of genotypes (M)
1000
100
10
1
0
2000
4000
6000
8000
10000
Generations
565
M*
555
545
9000
10000
B
Mean number of genotypes (M*)
10000
1000
100
10
1
0.1
0.5
1
5
10
50
100
500
NU
σ = 0.125
σ = 0.025
σ = 0.05


A
Optimal phenotype
Ancestral phenotype
1.0
0.5
0.0
α_2
α_1
$\frac{\alpha_1}{\alpha_1 + \alpha_2}$
B
Population counts
100
75
50
25
0
0.1
0.5
1
5
10
50
100
500
NU
Clusters:
1
2
3
4
C
σ:
0.0125
0.025
0.05
ρ:
0
0.5
1
Neutral
Probability (Clusters > 1)
-2
0
2
4
6
ln(NU)


A
NU: 0.1 0.5 1 5 10 50 100 500
Mean pairwise genetic distance
π_G
10.0
7.5
5.0
2.5
0.0
ρ = 0
ρ = 0.5
ρ = 1
Neutral
Mean pairwise phenotypic distance
π_P
12.5
10.0
7.5
5.0
2.5
0.0
ρ = 0
ρ = 0.5
ρ = 1
Neutral
B
>1 clusters
1 cluster
95% specificity
π_G
10.0
7.5
5.0
2.5
0.0
ρ = 0
ρ = 0.5
ρ = 1
C
ρ : 0 0.5 1
True positive percentage
False positive percentage
0
20
40
60
80
100
AUC: 94.8%
AUC: 91.3%
AUC: 83.6%
*
***
***


A
NU: 0.1 0.5 1 5 10 50 100 500
Tajima's D
2.5
0.0
−2.5
S N S N S N S N S N S N S N S N
B
Monomorphic samples
100
75
50
25
0
S N S N S N S N S N S N S N S N
Selection (S)
Neutral (N)