

Christopher J. R. Illingworth,[◇] Leopold Parts,[◇] Stephan Schiffels,[†] Gianni Liti,[‡] and Ville Mustonen[◇]
[◇]Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA, UK, [†]Institut für Theoretische Physik, Universität zu Köln, Zùlpicher Str. 77, 50937 Köln, Germany, and [‡]Laboratory of Biology and Pathology of Genomes, CNRS UMR 6267, INSERM U998, University of Nice Sophia-Antipolis, 06107 Nice, France.

When selection is acting on a large, genetically diverse population, beneficial alleles increase in frequency. This fact can be used to map quantitative trait loci by sequencing the pooled DNA from the population at consecutive time points, and observing allele frequency changes. Here we present a population genetic method to analyse time-series data of allele frequencies from such an experiment. Beginning with a range of proposed evolutionary scenarios, the method measures the consistency of each with the observed frequency changes. Evolutionary theory is utilized to formulate equations of motion for the allele frequencies, following which likelihoods for having observed the sequencing data under each scenario are derived. Comparison of these likelihoods gives an insight into the prevailing dynamics of the system under study. We illustrate the method by quantifying selective effects from an experiment in which two phenotypically different yeast strains were first crossed and then propagated under heat stress (Parts et al., *Genome Res.* 2011). From these data we discover that about 6% of polymorphic sites evolve non-neutrally under heat stress condition, either because of their linkage to beneficial (driver) alleles or because they are drivers themselves. We further identify 44 genomic regions containing one or more candidate driver alleles, quantify their apparent selective advantage, obtain estimates of recombination rates within the regions, and show that the dynamics of the drivers display a strong signature of selection going beyond additive models. Our approach is applicable to study adaptation in a range of systems under different evolutionary pressures.

Introduction

Fitness differences between individuals enable natural selection to increase the frequency of beneficial variants within a population over time. The specifics of this process, however, are often complex, with the fitness difference conferred by a variant potentially depending on time, space, the genetic background of the individual, the genotypic composition of the population, and other species in the vicinity (Hartl and Clark, 2007). Furthermore, other evolutionary forces such as mutation and genetic drift contribute to allele frequency changes, and their effects can mask those arising from differences in fitness.

For these reasons, the importance of studying the fitness effects of mutations and evolution in controlled laboratory settings is well known. One of the most celebrated of these experiments is the “*E. coli* long-term evolution experiment” which so far covers over fifty thousand generations (Woods et al., 2011). Deep sequencing data from this and other experiments have provided an unprecedented level of insight into the processes of molecular evolution (Barrick et al., 2009; Burke et al., 2010; Hietpas et al., 2011). However, despite this progress in the quantitative study of evolution, fundamental challenges still remain.

A first challenge in studying fitness effects is the time-scale it can take for evolution to increase the frequency of a mutation. Consider for instance a hypothetical variant in *E. coli* with a fitness advantage of $\sigma_0 \equiv f_{\text{mutant}} - f_{\text{wildtype}} = 10^{-5}$. Supposing the variant to have survived genetic drift, and neglecting the effects of other mutations, in a population of size $N = 10^8$ the variant would take $\sim 10^6$ generations to fix, representing hundreds of years of evolution. While in population genetic terms such a variant would be considered strongly beneficial (with the ratio of time-scales

associated with genetic drift and selection, $\sigma_0 N \gg 1$), in a laboratory setting it would be barely detectable within the lifetime of the researcher. For this reason, knowledge of fitness effects of that size is derived via analysis of intra- and interspecies variation (Sawyer et al., 2003; Eyre-Walker, 2006; Eyre-Walker and Keightley, 2007; Mustonen and Lässig, 2007; Sella et al., 2009).

A second challenge is posed by the mutation rate, which is often small, such that variance within a population is created slowly. The low initial frequency of a new mutant renders even strongly beneficial mutations susceptible to elimination through genetic drift, while, as noted above, mutants escaping drift take time to grow to detectable frequencies. Waiting for a well-adapted initial population to “find” beneficial mutations, and for these to become visible, may require long-term experiments. Use of mutagens (Weigand and Sundin, 2009), or increasing the overall number of individuals in the population can alter the number of mutations entering the population (Perfeito et al., 2007).

One means of increasing the rate of adaptation is to apply artificial selection on a system by imposing environmental stress (Kishimoto et al., 2010; Bell and Gonzalez, 2011) or nutritional restriction (Kao and Sherlock, 2008). A recently described method, examining heat tolerance in yeast, combines this approach with the addition of artificial variance generated by genetic crosses (Parts et al., 2011), crossing two strains of yeast, and propagating the resulting population asexually (and sexually) under heat stress conditions. Conceptually, this factorizes the evolutionary dynamics: mutations have accumulated over a long time period since the last common ancestor of the parental strains, recombination during the two-way crossing protocol ensures that variants not in close proximity are reasonably unlinked and at substantial frequencies, and strong selective pressure can be applied. As the population adapts to the selected condition, multiple changes in allele frequencies can be observed. The process can be traced at a single nucleotide resolution by deep sequencing the pooled DNA

Key words: adaptive molecular evolution, fitness effects of mutations, artificial selection

E-mail: vm5@sanger.ac.uk

Submitted to *Mol. Biol. Evol.* 1–10.

© The Author(s) 2011. Published by Oxford University Press on behalf of the *Society for Molecular Biology and Evolution*.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

from the population at consecutive time points.

This concept, of applying artificial selection to an artificially mixed population in order to identify quantitative trait loci (QTLs), has previously been applied to the malaria parasite (Culleton et al., 2005), and to yeast (Segrè et al., 2006; Ehrenreich et al., 2010). As performed in the instance analysed here (Parts et al., 2011), the method offers the advantages of a high genomic resolution and, extra to previous studies, time-resolved sequencing data, with allele frequencies measured at four distinct time-points. In order to fully exploit the time-series aspect of these new data, further methodological development is required.

In this paper we develop an approach, based on population genetic theory, to investigate data from such an experiment. Examining changes in the frequencies of segregating alleles over time, we derive measurements of the fitness effects that different alleles confer for heat tolerance, and identify the presence of non-additive fitness effects. As the data we analyse is from the asexual version of the experiment the qualitative picture of the dynamics is simple. During the crossing a large pool of recombinant genotypes are created. Under selective pressure, the genotypes acquire some unknown haplotype fitness distribution, which leads to a relative proliferation of the fitter, at the cost of the less fit, genotypes. However, to obtain a quantitative picture of the dynamics is challenging: sequencing of the pool gives only data of allele frequencies, not of genotypes. As will be shown, success depends on several factors, the most important being the population size. Another key issue is whether the initial pool contained enough variation in the high fitness part of the population for no single clone to dominate the population by the end of the experiment; if this were the case, changes in allele frequency would not allow for discrimination between selected and non-selected alleles in the clone. Here we perform our analysis from the perspective of alleles but go on to use the allele picture to seek evidence of more complex selective scenarios such as those acting on genotypes. While in this case the evolution is clearly driven by haplotype selection (which we can however analyse starting from an allelic viewpoint), for a sexually propagated population an important consideration is whether an “allele selection” or a “genotype selection” mode of evolution dominates (Neher and Shraiman, 2009).

In focusing on standing variation generated by the crossing of strains, important questions concerning *de novo* mutation processes are not addressed. Simultaneously, we recognize that it is not clear how important such extreme selection pressures are for organisms over macro-evolutionary time-scales. However, experiments such as this undoubtedly provide an exciting opportunity to study strong fitness effects, both quantitatively and systematically, at the molecular level. Here we address this opportunity to quantify the fitness effects acting on the heat tolerance trait in yeast within a set of more than thirty thousand segregating sites.

Overview of the experiment

Two diverged strains of *S. cerevisiae*, North-American (NA) and West-African (WA), were crossed for 12 generations to create a large pool of segregants (Parts et al.,

2011). After the crossing, the pool was put under heat stress (40°C) for a period of $T = 288$ hours (h) with replating (after mixing) of 10% of the pool every 48 hours, DNA from the remainder of the pool being sequenced at time points $t_0 = 0h, t_1 = 96h, t_2 = 192h, t_3 = 288h$. Estimating the number of generations (*gen*) that T corresponds to is difficult, however it should in any case be less than 10^2 *gen*. In order to avoid the need to convert the unit of time from hours to generations (which would be the more natural choice in terms of population genetics) we here measure rates in units of $1/(96h)$. During the selection protocol the population size N varied from $\sim 10^7$ after each replating to $\sim 10^8$ before replating. In the following we denote the NA allele with index 0 and WA allele with index 1. Substantial allele frequency changes were observed over the course of the experiment (Parts et al., 2011).

Modelling the observed time evolution

Time-scales of the processes

A simple calculation suggests the likely role of genetic drift and *de novo* mutations in the selection process to be negligible. We first note that the duration of the experiment, T , is between five to six orders of magnitude smaller than the population size, N . As such, genetic drift, which changes allele frequencies at time-scales of $\sim N$ generations (Kimura, 1964), most likely has very little effect across the course of the selection experiment. Considering next the possibility of *de novo* mutation, the mutation rate is known to be small (Drake et al., 1998; Lynch et al., 2008; Lang and Murray, 2008). Estimating a worst case scenario by modelling the deterministic growth of a mutant at the outset of the experiment, we note that selection pressures of $\sim 1/T \log(N/N^s)$ are required to reach frequency $1/N^s$ within time T (as can be derived from Eq. 5), where N^s is the mean sequencing read depth. This means that the variant would need to have ~ 0.03 growth rate advantage per hour to reach detection threshold during the experiment (for this data $N^s \sim 10^2$). We contend that mutations with so large an effect are likely to be extremely rare and assume that changes in allele frequencies are driven by population variation existing in the initial population (we demonstrate later that *de novo* mutations do not play a substantial role directly from the data). Data from a biological replica of the experiment supported this conclusion (Parts et al., 2011).

In order to analyse the allele frequency changes observed in the experiment, we considered a variety of scenarios described by deterministic evolutionary dynamics in conjunction with a stochastic sampling process resulting from finite sequencing depth. For the reasons argued above, the evolution of the system was taken to be driven by fitness differences between segregating alleles in the initial population.

Driver and passengers

We consider first a system with a single driver at locus i with two possible alleles $a \in \{0, 1\}$, influencing all passenger mutations which are in linkage disequilibrium with it. We recall from deterministic single locus theory that the driver evolves according to the equation (Hartl and Clark,

2007):

$$dq_i^1/dt = \sigma_i q_i^1 q_i^0 \quad (1)$$

where the frequency of the allele 1 is denoted with q_i^1 , ($q_i^0 = 1 - q_i^1$) and the selection coefficient σ_i equals the Malthusian fitness difference $f_i^1 - f_i^0$ between the alleles.

The general mathematical framework to compute the effect of a deterministic driver on linked neutral variation has been introduced in classical work on genetic hitchhiking (Smith and Haigh, 1974; Barton, 2000). Here we use the case with zero recombination because during the artificial selection phase the population evolves asexually and no further recombination takes place. Therefore, given the time evolution of the driver we can write down equations of motion for the passengers:

$$q_j^b(t) = \sum_{a \in \{0,1\}} q_i^a(t) \frac{q_{ij}^{ab}(t_0)}{q_i^a(t_0)}, \quad j \neq i, \quad (2)$$

where q_{ij}^{ab} denotes a two locus haplotype at loci i, j with alleles $a, b \in \{0, 1\}$. Equation 2 follows simply by noticing that the passenger locus has by definition zero selection (its alleles would stay at their initial frequency without linkage to the driver), such that the passenger's initial linkage to the driver dictates its motion. We note that the two locus haplotype frequencies at the initial pool can be written in terms of allele frequencies and linkage disequilibrium:

$$q_{ij}^{ab}(t_0) = q_i^a(t_0) q_j^b(t_0) + (-1)^{a+b} D_{ij}. \quad (3)$$

As such, the dynamics of the passengers given the driver are fully fixed by D_{ij} . Values of D_{ij} can be inferred directly for each locus, or parametrized in terms of the recombination which took place during the crossing:

$$D_{ij}(\rho, \Delta_{ij}) = D_{ij}'(1 - \rho \Delta_{ij})^{N_{\text{crossing}}} \quad (4)$$

where $D_{ij}' = \min\{q_i^0(0)q_j^1(0), q_i^1(0)q_j^0(0)\}$ is the maximum linkage disequilibrium attainable, Δ_{ij} denotes the distance between the loci in base pairs (bp), ρ measures the recombination rate during the crossing process in units of $1/(bp \times gen)$, and N_{crossing} denotes the number of crossing rounds (for $\rho \Delta_{ij} > 1$ we set $D_{ij} = 0$). Equation 4 assumes an infinite population size but is nevertheless a good description for the system due to the large number of individuals in the initial pool. (See *Supporting Text* for analysis of finite populations by means of computer simulation. The population size required to correctly decide whether a marker moved due to linkage to a nearby selective sweep or just due to drift has been calculated (Logeswaran and Barton, 2011).)

In a case where full sequences from the initial pool were available, the initial linkage pattern could be included directly by measuring the linkage, circumventing Eq. 4, which from the inference framework most critically relies on a large population size.

Liberal-driver and passengers

We also consider models in which drivers are allowed to take any trajectory, that is, in which their dynamics are not parametrized, and hence constrained, by the equation of motion given in Eq. 1. For these ‘‘liberal-drivers’’ the

passenger dynamics again follow Eq. 2. We further evaluate time-dependent selection coefficients for such trajectories; from Eq. 1 we get for each time interval:

$$\sigma_i(t_k) = \frac{1}{\Delta t_k} \left\{ \log \frac{q^1(t_{k+1})}{q^0(t_{k+1})} - \log \frac{q^1(t_k)}{q^0(t_k)} \right\}, \quad (5)$$

where $\Delta t_k = t_{k+1} - t_k$. Such dynamics could result from externally driven, truly time-dependent selective pressures, or as yet unidentified internal interactions. Internal interactions could result from linkage to other drivers, epistatic fitness interactions between the locus and other loci (e.g. genotype selection), frequency-dependent selection, or a combination of all of these factors. We return to the interpretation of these drivers later.

The evolutionary equations outlined in this section were applied within a standard maximum-likelihood framework (see *Methods*) to explore the observed time evolution of the allele frequencies.

Results

Prevalence of non-neutral trajectories

Close to six percent of the segregating sites across the genome were identified as evolving in a significantly non-neutral manner. To detect non-neutral behaviour, we calculated likelihoods for each trajectory under two models, the first assuming that they evolved neutrally, with any motion reflecting noise from the finite depth of sampling, and the second assuming that they evolved independently under selection as described by Eq. 1. For each locus, this gave maximum-likelihood predictions for the allele frequencies, corresponding likelihood scores, and for the second model, a trajectory-specific selection coefficient. Global statistics for the likelihood differences between models, and the identified selection coefficients, are reported in Figure 1. Applying the Akaike information criterion (AIC) (Akaike, 1974), the fraction of loci identified by this comparison as being not consistent with neutral evolution (AIC difference ≥ 10 (Burnham and Anderson, 2002)) was 0.058 (see *Supporting Text*, Figure S1 for analysis of trajectories from a control experiment indicating false positive rate less than 1%, Figures S2, S3 and S5 for chromosome specific breakdown). The average selection coefficient was -0.36 (std. 0.63) implying that the North-American allele, rather than the West-African one, is more often the beneficial allele under the imposed condition of heat stress. The choice of which of the two alleles has frequency q^1 is arbitrary and sets the direction of selection (see Eq. 1).

This simple protocol of assigning a selection coefficient to each locus provided an instant genome-wide view of the data set, allowing for the rapid identification of genomic regions of interest. Using the log-likelihood scores generated under each model, we identified 44 candidate driver foci for further examination (see *Supporting Text*).

Including linkage - driver and passengers

A driver-passenger model (Eqs. 1-4), which takes the linkage between nearby loci into account, gave a substantially better representation of the evolution of the system in

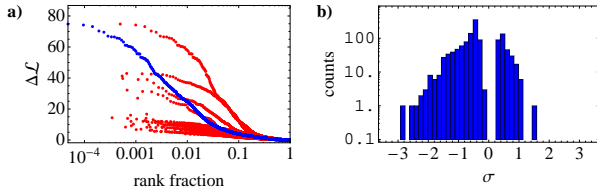


FIG. 1.—**Genome wide view under unlinked analysis.** a) Ordered score differences $\Delta\mathcal{L}$ (see *Methods*) between the unlinked additive drivers and the neutral model. Blue dots show values for all trajectories while sets of red dots correspond to individual chromosomes (the upper most is chromosome XV and the second highest II, see *Supporting Text* Figures S2 and S3 for individual chromosomes). b) Global histogram of selection coefficient calling only loci for which $\text{AIC} \geq 10$. The negative mean selection $\bar{\sigma} = -0.36$ reflects that the North-American allele is more often better adapted for the heat stress condition than the West-African one.

all of the 44 candidate driver foci than the unlinked driver model.

A particularly intuitive example is provided by the driver in chromosome II reported in Figure 2. While the additive unlinked drivers model explains the main part of the sweep region well, it fails to account for the motion observed around 475kb (Figure 2a). This failure is easy to understand - the observed allele frequencies move substantially during $\Delta t_0 = t_1 - t_0$ but stay at almost the same value during Δt_1 and Δt_2 - a motion incompatible with the family of curves parameterized by Eq. 1. While an unlinked driver allele, at frequencies close to 0.5, moves at roughly constant velocity towards fixation or death, an allele frequency changing through linkage may fix at intermediate values. Hence, taking model complexity into account, a model of a single driver and passengers (Figure 2b), explains the changes in this region of the genome better. Most frequencies in the region move due to linkage, rather than through inherent selection.

Distributions of selection coefficients and recombination rates

We inferred estimates for driver selection coefficients for each of the candidate driver foci using the driver and passengers model with linkage parametrized by recombination rate (Figure 3a). Furthermore, inference in each case gave a maximum likelihood estimate of the local recombination rate (Figure 3b). The inferred selection coefficient was negative for 29 of the 44 regions (mean $\bar{\sigma} = -0.2$, std. 0.6), reflecting the advantage conferred by the North-American allele. A mean magnitude of selection of $|\bar{\sigma}| = 0.44$ indicated that the drivers evolve under substantial selection. The maximum likelihood estimates for recombination rates (mean $\bar{\rho} = 1.6 \times 10^{-6} (\text{bp} \times \text{gen})^{-1}$, std. 1.1×10^{-6}) are consistent with estimates from the literature (Ruderfer et al., 2006; Mancera et al., 2008).

Liberal-drivers and passengers

In the majority of cases (38/44 regions) the liberal-driver model gave a significantly improved fit to the observations after allowing for the additional two degrees of

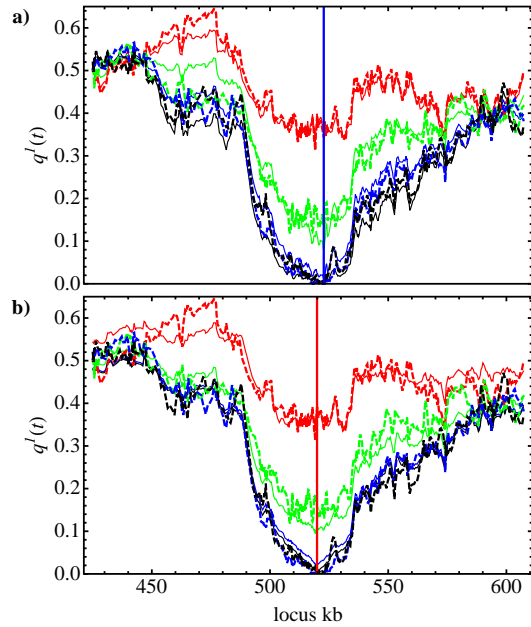


FIG. 2.—**An example of unlinked additive drivers vs. driver and passengers models.** Allele frequencies in the depicted chromosome II region moved substantially during the experiment. a) Unlinked additive drivers model where every trajectory follows Eq. 1: Thick dashed lines (t_0 -red, t_1 -green, t_2 -blue and t_3 -black) show the data and the corresponding solid lines the maximum likelihood predictions for the motion. The blue vertical line denotes the location of the largest σ (for the full selection profile, see *Supporting Text* Figure S5). The model explains the motion close to the sweep focus well, however, it fails qualitatively in the region $\sim 475\text{kb}$. As explained in the text, the motion in that region is not compatible with modes provided by Eq. 1. b) Driver and passengers model as parameterized by Eqs. 1-4. The red vertical line denotes the inferred driver location and as can be seen the region $\sim 475\text{kb}$ is much better explained than by the unlinked additive drivers model. The difference between the AIC scores was 466 in favour of the driver and passenger model, the large difference indicating the importance of including linkage between nearby loci in the model. Data shown is averaged over a sliding window of 5, however all inferences are done with the raw data.

freedom (having three σ_i values compared to one, see *Supporting Text* Table S1). In the example of chromosome II, discussed above, a single driver with a constant recombination rate appeared to explain the observed changes in frequencies very well. Nevertheless, the liberal-driver model explained the motion much better (gain of 166 units of log-likelihood, see *Supporting Text* Figure S4). However, in chromosome XIII, the second example in which the identified driver (almost) fixed within the experimental time frame, a gain of only 3 units of score was achieved, favouring the standard driver and passengers model once degrees of freedom were taken into account.

Apart from these events, other driver alleles were found at intermediate frequencies by the end of the experiment. This observation has two possible explanations. Firstly, all candidate drivers could evolve according to trajectories defined by Eq. 1, but with values of σ_i too low to observe fixation during the length of the experiment. Secondly, candidate drivers could evolve according to some alternative equations of motion. Application of the liberal-

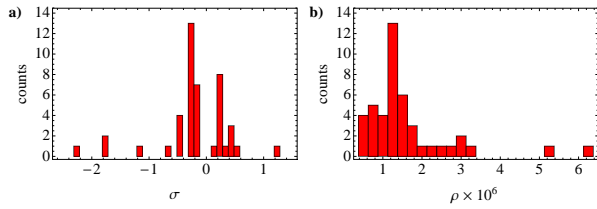


FIG. 3.—Statistics of selection and recombination under the driver and passengers model. a) Histogram of selection coefficients for the driver foci (mean $\bar{\sigma} = -0.2$ std. 0.6). b) Histogram of inferred recombination rates for these regions (mean $\bar{\rho} = 1.6 \times 10^{-6}$, std. 1.1×10^{-6}).

driver model suggested the latter explanation to be correct for the majority of candidate drivers. While the liberal-driver model introduces two additional degrees of freedom, use of this model produced a mean score gain per region of 26 (see *Supporting Text Table S1*), more than compensating for the gain in parameters. Figure 4 shows an example within chromosome XV where the liberal-driver model gave a substantial improvement.

Averaged across the duration of the experiment, the selection coefficients inferred for the liberal-drivers, expressed as in Eq. 5, were very similar to those shown in Figure 3 from the driver and passengers model (mean -0.19 std. 0.68). Estimates of local recombination rates obtained using liberal-drivers were also very similar (mean 1.5×10^{-6} std. 1.9×10^{-6}). To gain an insight to the underlying reason for the liberal drivers' superior log-likelihood scores, we studied the process at the level of individual time intervals Δt .

Time-dependent selection coefficients were used to obtain point estimates of fitness flux, ϕ , a measure of the amount of ongoing adaptation in the system at a given time point expressed as $\phi(t_k) = \sum_{i \in \text{drivers}} \sigma_i(t_k) \Delta q_i^1(t_k) / \Delta t_k$ (Mustonen and Lässig, 2007, 2010). The estimates of fitness flux over the three measured time intervals differ noticeably between the two models (Figure 5a). Whereas, under the standard-driver model, most adaptation takes places in the first time interval, for liberal-drivers the second time interval dominates. Most interestingly, under the liberal-driver model, the fitness flux almost vanishes for the last time interval, suggesting that by the end of the experiment the system had almost equilibrated. This point is further demonstrated by the distribution of the inferred motion of the drivers under both standard and liberal scenarios during the last time interval, shown in Figure 5b.

We suggest that the apparent equilibration observed here should be understood in the sense of separation of time-scales, in that it represents the completion, within sequencing resolution, of the first phase of adaptation (due to the finite sequencing depth we would not observe movement slower than $\lesssim (\text{read depth})^{-1} / \Delta t$). Given further propagation of the system, deviation from this equilibrium would be observed through the arrival of new mutations, however, as discussed later, we do not find evidence for these at substantial frequencies within the time frame of the experiment.

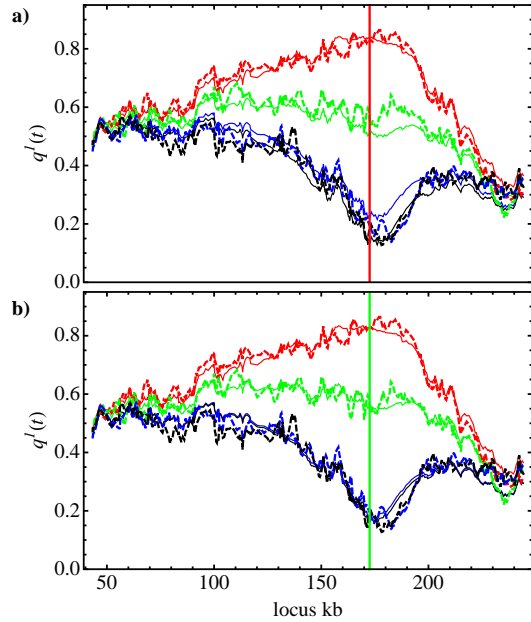


FIG. 4.—An example of standard-driver vs. liberal-driver models. Allele frequencies in the depicted chromosome XV region moved substantially during the experiment. Thick dashed lines (t_0 -red, t_1 -green, t_2 -blue and t_3 -black) show the data and the corresponding solid lines the maximum likelihood predictions for the motion, vertical lines denote the inferred driver location. a) Driver and passengers model: The model explains the motion overall quite well, however, it fails qualitatively near the sweep focus at 172kb. The motion, which, as is evident by the visible overlap between the blue and black dashed lines, seems to reach equilibrium at an intermediate frequency, is not compatible with models provided by Eq. 1. b) Liberal-driver and passengers model: The trajectories near the focus are much better explained than by the standard-driver model. The liberal-driver interpretation gives a gain of 79 units in log-likelihood at a cost of two extra degrees of freedom and is thus strongly supported statistically. Data shown is averaged over a sliding window of 5, however all inferences are done with the raw data.

The observation of candidate driver alleles reaching equilibrium at intermediate frequencies suggests the presence of interactions between drivers in different locations of the genome. An alternative scenario exists, in which drivers do not interact with one another, but instead evolve with genuinely time-dependent selection coefficients. However, while behaviour of this kind might be observed in response to changes in the external environment, the consistency of the experimental conditions suggests its occurrence here to be unlikely, interactions between drivers being the most likely source of the observed allele frequency changes. Different effects leading to interactions between drivers are discussed later.

New beneficial mutations

We contended earlier, based on a simple calculation, that *de novo* mutations are unlikely to play a substantial role for the allele frequency dynamics during the experiment because they would need to carry fitness advantage $\gtrsim 3\%/h$ to be detectable during the experiment. However, in the measured spectrum of fitness effects in Figure 3, we

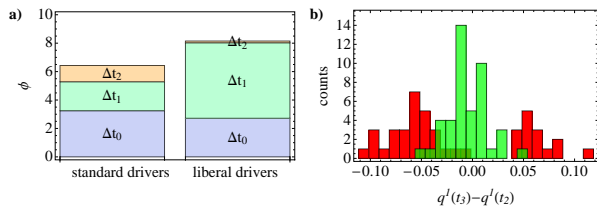


FIG. 5.—**Adaptation under drivers vs. liberal-drivers interpretation.** a) Estimates of fitness flux: in the standard-driver interpretation (Eqs. 1-4) most adaptation took place during the first time interval, Δt_0 , whereas liberal-drivers adapt most during Δt_1 and have a higher total fitness flux. It is also apparent that while standard-drivers continue to sustain a fitness flux during the last interval Δt_2 , the flux generated by liberal-drivers has all but vanished. b) Inferred motion of the drivers (red) and liberal-drivers (green) during the last time interval supports the observation that under the liberal-driver interpretation the system has almost reached equilibrium (in the sense discussed in the text).

observe selection coefficients as large as to be $\approx 2.4\%/h$, raising the possibility that novel mutations might have beneficial effects strong enough to have a measurable effect. We thus took further steps to rule out this possibility.

As there is no recombination during the selection protocol any *de novo* mutation would lead to a global perturbation of the allele frequency dynamics, one of the two alleles at each segregating site being fully linked to the new mutation. This would lead to a large effect, increasing as a function of time, on the movement of the allele frequencies. We thus calculated genome-wide statistics of absolute changes in allele frequency at the three time intervals (i.e. $|q(t_k) - q(t_{k-1})|$) from the trajectories inferred under the unlinked driver model. The statistics showed that the absolute movement of allele frequencies slightly decreased throughout the course of the experiment, with mean values of 0.032, 0.032, 0.030 respectively, reflecting our inference that 94% of allele frequencies evolved in a manner compatible with neutral (no motion) evolution with the remaining 6% evolving as discussed. This decrease suggests that *de novo* mutations did not significantly affect the allele frequency dynamics during the time interval reported.

Spatial uncertainty of identified drivers

Likelihood calculations suggested that, using the (liberal) driver passenger models, the precise location of a driver could be identified on average to within an accuracy of 12kb. Due to linkage, alleles at nearby loci are likely to move in a similar manner, leading to uncertainty in the precise choice of location of a driver locus. This uncertainty was quantified by taking the maximum likelihood drivers inferred under the standard and liberal-driver models and checking the next twenty segregating sites both up and downstream, assigning each in turn to be the driver, re-optimising the remaining degrees of freedom, and calculating the resulting likelihoods. Figure 6a shows for standard-drivers that, using a log likelihood cutoff of three units (95% confidence interval), the driver allele can on average be located within a window of 12kb centred on the maximum likelihood location. This means that on average 38 segregating sites remain to be further studied for

each driver focus (see *Supporting Text Pre-processing data - identifying regions of interest* for details how this number of segregating sites was calculated). However, region specific uncertainty strongly depends on the selective strength of the driver, local recombination rate and on the allele read numbers from sequencing - variability of these factors leads to substantial variance in the accuracy of the inferred driver location (Figure 6a). Nevertheless, for regions with strong drivers, the small number of segregating sites to be further studied reflects the efficiency of recombination breaking linkage when advanced inter-cross lines are used in the experimental design (Darvasi and Soller, 1995). Comparison of driver locations under each of the two models, shown in Figure 6b, revealed differences that were mostly within the uncertainty range, with mean absolute distance ~ 2 kb.

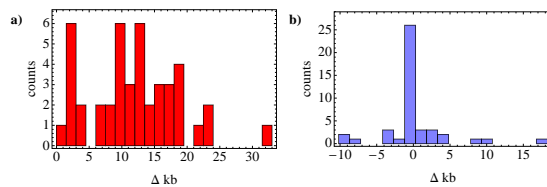


FIG. 6.—**Driver location uncertainty.** a) Inferred sizes of the regions containing loci which when fixed to be the driver are within a distance of 3 units of score (95% confidence interval) from the maximum likelihood value (with re-optimization of the remaining degrees of freedom). Data shown is for standard-drivers. Variability in the magnitude of selection and local recombination rates leads to substantial variation in the uncertainty of the driver locations. b) Distances between the maximum likelihood driver locations inferred under the standard-driver and liberal-driver models. For most foci the two models gave predictions agreeing within variability identified in panel a).

Analysis of a biological replicate experiment and estimates of false positive rates

One great advantage of artificial selection protocols is the opportunity to analyse biological replicates to gauge robustness of inferences made. We analysed data from a biological replicate experiment and show that inferred statistics of selection replicate well (see *Supporting Text* for details). However, as the replicate data set had fewer time-points, and was thus not fully comparable to the one discussed so far, we performed computer simulations to estimate false positive rates for the analysed 44 driver set. This resulted in an estimated false positive rate for our driver region detection to be less than $\sim 2\%$ for populations of size 10^7 . In these simulations we chose the number of segregating sites, number of drivers and their estimated selection strengths and recombination rates to mimic the ones found from the experimental data (see *Supporting Text* for details).

The 44 regions called in our analysis as containing driver loci is significantly larger than the more conservative 21 (all these regions are in our list) obtained by Parts et al. (2011), who, for example, only call QTLs observed in both the primary and the replica datasets. The existence of a higher number of regions under selection than previously

reported is supported by the false positive rate identified from simulated data.

Discussion

Quantifying selection from time-resolved allele frequencies

We performed the first, single nucleotide resolution, genome wide population genetic analysis of time-series allele frequency data from an evolving outbred population under selection. The analysis revealed several insights into the dynamics of the population at the molecular level. First, we estimated that close to 6% of the over thirty thousand segregating sites are affected by the selection pressure as identified via deviation from a neutral null model. Second, we demonstrated the main force causing this motion to be genetic linkage which causes passenger alleles to hitchhike with driver alleles. Third, using a driver and passenger model we quantified the selective advantage of the found drivers, the amount of linkage disequilibrium within the regions that they reside and the uncertainty in their locations. The method we describe offers a first-order approach to examining selection in such an experiment.

Extending the above, by allowing driver alleles to evolve arbitrarily, we found substantial statistical evidence of fitness effects going beyond those compatible with standard additive models of selection, such that many of the driver alleles evolve under an effectively time-dependent fitness seascape (Mustonen and Lässig, 2009). We next discuss possible scenarios underlying the observed liberal-driver dynamics.

Linkage pattern in the initial pool

One explanation for the inferred liberal-drivers (or indeed the standard drivers) signal would be the existence of linkage disequilibrium between candidate drivers. Such linkage could result from stochastic effects during the crossing, arising due to the finite population size, a scenario which Eq. 4 would not capture. Numerical simulations (reported in *Supporting Text*) showed that population sizes from $\sim 10^6$ upwards are sufficient to rule out distortion to the global selection statistics arising from such noise, on the assumption that the numbers of segregating sites and drivers, driver selection strengths and recombination rates are close to those inferred from the real data. Furthermore, analysis of data from a biological replicate gave overall statistics of selection consistent with those inferred, supporting the conclusion that drift-generated linkage disequilibrium is not causing the signal (see *Supporting Text* for more details).

From our simulations, we do note, however, that a large number of false positives would be generated due to drift-generated linkage in population of size 10^5 . Therefore, in smaller populations our method should be applied only for identifying the largest fitness effects to avoid being swamped by false positives (which would reflect mistaken measurements of selection rather than any underlying fitness land- or seascape), or in combination with full sequence data from the initial pool to fix the linkage. The population size required to correctly decide whether a marker moved due to linkage to a nearby selective sweep

or just due to drift has been calculated (Logeswaran and Barton, 2011).

Another possible reason for a non-trivial linkage pattern in the initial pool could come from unknown selection pressures during the crossing protocol. Given that such selection could in principle be arbitrarily complex, our analyses of the data here remain to some extent vulnerable to this. However, there is no statistical evidence for inter-chromosomal pairwise linkages between genotype data of 19 QTL loci (all part of our candidate driver foci list) from 960 individuals at time-point $(t_2 + t_3)/2$ as reported in (Parts et al., 2011). Therefore, worst case scenarios where all but one of the candidate drivers were linked to, and thus would merely hitchhike with, one of the two nearly fixing drivers are not supported by the data. Similarly to non-local linkage disequilibrium generated by small populations, possible effects of selection during the crossing to allele dynamics under the artificial selection phase could be easily included to the present method given suitable data. Extending the method to cover simultaneous recombination and selection is possible, but remains a topic for further investigation.

Clonal interference

In asexually evolving populations recombination cannot combine beneficial mutations that are on different haplotypes. This leads to so-called clonal interference where some beneficial mutations will be removed from the population due to them being interfered with stronger ones in different backgrounds (Fisher, 1930; Muller, 1932). The dynamics of clonal interference are complex and both its experimental and its theoretical study has long been an active research topic (for recent reviews see (Park et al., 2010; Sniegowski and Gerrish, 2010)).

For the results presented here, the large number of drivers make it unlikely to have a large number (or indeed any) individuals with all the beneficial alleles in a population of size $\sim 10^7 - 10^8$, so that such interference seems inevitable. However, simulation and sequencing data both suggest that clonal interference does not account for the liberal-driver signal observed here.

Analysing simulated data with the number of drivers (additive in fitness) and their strengths comparable to those inferred from the real experiment we concluded that the observed liberal-driver dynamics does not reflect effects primarily caused by clonal interference. We reached this conclusion by analysing the simulated data under both standard and liberal-driver models (see *Supporting Text*). With both models we were able to discover the “true” driver set and reproduce the correct selection strengths for the drivers. However, using a liberal-driver model to infer selection from the simulated data, we did not see the decrease in fitness flux observed for the real system (contrast Figure 5 and Supporting Figure S14). For this reason we believe that clonal interference is not the primary explanation for the stronger performance of the liberal-driver model for most candidate driver regions in the real data. Supporting this conclusion, genotyping data of 24 loci from 960 individuals from time-point $(t_2 + t_3)/2$ showed that no single genotype (clone) dominated the population at that time

(maximum count observed was 6), a considerable diversity of 787 unique haplotypes being detected and the data being consistent with the expectation under a hypothesis of independent sites (Parts et al., 2011).

Given the absence of recombination, it is inevitable that clonal interference would, after sufficient time had passed, have a significant influence on the evolution of driver allele frequencies. However, given additive drivers of the size and number that we find from the real system, that point would not have been reached by the end of the experiment.

Clonal interference between (possible) multiple beneficial alleles within a driver focus remains a possibility but would not stop the dominant combination of these mutations from fixing and as such is not consistent with the liberal-driver signal. Extension of the method to cover multiple drivers will be required to study such effects and is beyond the scope of this manuscript.

Epistasis and other more complex selective scenarios

Given that linkage or clonal interference between driver foci do not satisfactorily explain the observed dynamics, we believe the liberal-driver signal is likely to be caused by fitness effects going beyond additive models of selection. For instance, one such complex selection scenario could be that of epistatic interactions between drivers (Weinreich et al., 2005), encompassing combinations of pairs or potentially multiple drivers, while the possibility of genuinely time-dependent or frequency dependent selection cannot be definitively ruled out. Models could be developed to examine the likelihood of each of these potential scenarios. However, inference of a correct model based on the data available presents a substantial challenge, such that the precise nature of the observed time-dependency remains to be shown.

Comparable events at macro-evolutionary time-scales

The adaptive dynamics studied here consisted of driver alleles at loci under strong selection for heat resistance and neutral passenger alleles linked to the drivers. As can be seen in the frequency profiles in Figures 2, 4, and S5, genomic regions around the drivers have substantially reduced allelic variance. The experiment and the observed dynamics took place over time-scale of days. Nevertheless, a similar pattern of the hitchhiking effect has been described theoretically (Kim and Stephan, 2002) and found, among others, in the case of *Drosophila*, where individual selective sweeps in the recent past leave remnant genomic valleys of reduced variability (Svetec et al., 2009; Macpherson et al., 2007). However, while in the case of selective sweeps variability is removed completely from the genomic locality, in many of the cases studied here a substantial degree of variance remains in the immediate vicinity of the sweep locus after the adaptive event: the genomic signature resembles that of a soft sweep (Hermisson and Pennings, 2005), where the beneficial driver allele is initially linked to many different background genotypes.

From driver locations to biology

In the data considered here, the average spatial resolution which can be attained in identifying driver locations is about 12kb (see Figure 6). As such, identification of individual variants responsible for the inferred fitness advantages would require further experimental or bioinformatic analysis, such downstream analysis being very important in terms of the biological insight that can be collected from this data. Analysis of some of the large events observed in the data has been carried out in earlier work (Parts et al., 2011), showing them to be biologically interesting with respect to heat tolerance. While mapping the precise locations of the drivers are important for the biology of the trait in question, we note that inference of selection as was done here is not sensitive in this respect and thus the resulting estimates for selection are robust.

Conclusions

Our analysis here has been of time-resolved allele frequency data from a yeast population under heat stress. However, the method presented has potential application to a range of organisms under a variety of selection pressures. Several conditions were necessary for our analysis to be sensible: efficient recombination during the crossing to break up linkage and create variation, a large population size and high initial allele frequencies to reduce the effect of genetic drift and clonal interference, the duration of the experiment, which does not allow *de novo* mutations to start interfering, and finally, the application of strong selection. The method can be extended to sexually propagated populations by considering selection and recombination simultaneously, and in such a setting the key conditions enumerated above will become somewhat more lenient.

The focus on quantifying selection and utilizing multiple consecutive time points distinguishes our method from a standard QTL mapping design although such a mapping would be a natural biological application. We see the combination of the experiment with multiple time points and the approach presented here as a model for the quantitative study of the fitness effects of mutations and their possible interactions. Time resolution is key to our analyses; if data had been collected only from the initial and final pool, we would not have been able to show that majority of the drivers are not compatible with a standard model of additive selection. Under the framework of experiment and data analysis described here, alterations to study other variables would be possible; for example, the effect of drift could be examined through modulating the population size.

With the imminent arrival of deep population samples of time-resolved full genome sequences from evolution experiments, the possibility emerges of scaling up our measurements of the fitness effects of mutations to include complex multi-locus interactions. Application of population genetic theory will be central to making the best possible use of such data.

Methods

Trajectory probabilities

Given observations of allele counts at a locus i we can write the corresponding trajectory probability under some evolution model \mathcal{M} :

$$\mathcal{P}(\mathbf{n}_i|\mathcal{M}) = \prod_{t=t_0}^T P(n_i(t)|N_i^g(t), \mathcal{M}) \quad (6)$$

where P is binomial probability distribution, $N_i^g(t)$ denotes number of draws i.e. sequence read depth at locus i at time t and the true underlying population frequency of the allele $q_i^a(t)$ depends on the model \mathcal{M} which fixes the time evolution of the system. The evolution of $q_i^a(t)$ can be influenced by alleles at other loci depending on the specifics of \mathcal{M} . We can then write the probability for the total observation given \mathcal{M} which we, to serve as an example, take here to be the driver and passengers model defined in Eqs. 1-4:

$$\log \mathcal{P}(\mathbf{n}|\mathcal{M}_i) = \log \mathcal{P}(\mathbf{n}_i|\mathcal{M}_i^{\text{dri.}}) + \sum_{j \neq i} \log \mathcal{P}(\mathbf{n}_j|\mathcal{M}_j^{\text{pass.}}, \mathcal{M}_i^{\text{dri.}})$$

where $\mathcal{M}_i = \{\mathcal{M}_i^{\text{dri.}}, \mathcal{M}_i^{\text{pass.}}\}$, $\mathcal{M}_i^{\text{dri.}} = \{\sigma_i, \rho, q_i^a(t_0)\}$ and $\mathcal{M}_j^{\text{pass.}} = \{q_j^b(t_0)\}$. Thus, we can form a log-likelihood score of the model \mathcal{M}_i given the observation:

$$\mathcal{L}(\mathcal{M}_i|\mathbf{n}) = \log \mathcal{P}(\mathcal{M}_i^{\text{dri.}}|\mathbf{n}_i) + \sum_{j \neq i} \log \mathcal{P}(\mathcal{M}_j^{\text{pass.}}, \mathcal{M}_i^{\text{dri.}}|\mathbf{n}_j). \quad (8)$$

This is the function underlying our maximum likelihood inference. The structure of the log-likelihood is such that the driver i (index i also to be learned) influences every passenger trajectory j , but given the driver dynamics we can independently maximize likelihoods of the passenger trajectories which only have their initial conditions (if so desired also linkage to driver D_{ij}) as free parameters.

Acknowledgments

We would like to thank the anonymous referees for helpful comments. CI and VM would like to acknowledge the Wellcome Trust for support under grant reference 098051. LP was supported by the Wellcome Trust [grant number WT077192/Z/05/Z]. GL was supported by CNRS and ATIP-AVENIR. This research was also supported in part by the National Science Foundation under Grant No. NSF PHY05-51164 during a visit (CI, SS and VM) at the Kavli Institute of Theoretical Physics (KITP, Santa Barbara, CA). We would like to thank Richard Durbin and participants of the KITP program on Microbial and Viral Evolution for discussions.

Literature Cited

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723.

Barrick, J. E., Yu, D. S., Yoon, S. H., Jeong, H., Oh, T. K., Schneider, D., Lenski, R. E., and Kim, J. F. (2009).

Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature*, 461(7268):1243–7.

Barton, N. H. (2000). Genetic hitchhiking. *Philos Trans R Soc Lond, B, Biol Sci*, 355(1403):1553–62.

Bell, G. and Gonzalez, A. (2011). Adaptation and evolutionary rescue in metapopulations experiencing environmental deterioration. *Science*, 332(6035):1327–1330.

Burke, M. K., Dunham, J. P., Shahrestani, P., Thornton, K. R., Rose, M. R., and Long, A. D. (2010). Genome-wide analysis of a long-term evolution experiment with *Drosophila*. *Nature*, 467(7315):587–90.

Burnham, K. P. and Anderson, D. R. (2002). *Model selection and multimodel inference: a practical information-theoretic approach*. Springer, 2 edition.

Culleton, R., Martinelli, A., Hunt, P., and Carter, R. (2005). Linkage group selection: rapid gene discovery in malaria parasites. *Genome Res*, 15(1):92–7.

(7) Darvasi, A. and Soller, M. (1995). Advanced intercross lines, an experimental population for fine genetic mapping. *Genetics*, 141(3):1199–207.

Drake, J. W., Charlesworth, B., Charlesworth, D., and Crow, J. F. (1998). Rates of spontaneous mutation. *Genetics*, 148(4):1667–86.

Ehrenreich, I. M., Torabi, N., Jia, Y., Kent, J., Martis, S., Shapiro, J. A., Gresham, D., Caudy, A. A., and Kruglyak, L. (2010). Dissection of genetically complex traits with extremely large pools of yeast segregants. *Nature*, 464(7291):1039–42.

Eyre-Walker, A. (2006). The genomic rate of adaptive evolution. *Trends Ecol Evol (Amst)*, 21(10):569–75.

Eyre-Walker, A. and Keightley, P. D. (2007). The distribution of fitness effects of new mutations. *Nature Reviews Genetics*, 8(8):610–8.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.

Hartl, D. and Clark, A. (2007). *Principles of population genetics*.

Hermisson, J. and Pennings, P. S. (2005). Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics*, 169(4):2335–52.

Hietpas, R. T., Jensen, J. D., and Bolon, D. N. A. (2011). Experimental illumination of a fitness landscape. *Proc Natl Acad Sci USA*, 108(19):7896–901.

Kao, K. C. and Sherlock, G. (2008). Molecular characterization of clonal interference during adaptive evolution in asexual populations of *Saccharomyces cerevisiae*. *Nat Genet*, 40(12):1499–1504.

Kim, Y. and Stephan, W. (2002). Detecting a local signature of genetic hitchhiking along a recombining chromosome. *Genetics*, 160(2):765–77.

- Kimura, M. (1964). Diffusion models in population genetics. *Journal of Applied Probability*.
- Kishimoto, T., Iijima, L., Tatsumi, M., Ono, N., Oyake, A., Hashimoto, T., Matsuo, M., Okubo, M., Suzuki, S., Mori, K., Kashiwagi, A., Furusawa, C., Ying, B.-W., and Yomo, T. (2010). Transition from positive to neutral in mutation fixation along with continuing rising fitness in thermal adaptive evolution. *PLoS Genet*, 6(10):e1001164.
- Lang, G. I. and Murray, A. W. (2008). Estimating the per-base-pair mutation rate in the yeast *saccharomyces cerevisiae*. *Genetics*, 178(1):67–82.
- Logeswaran, S. and Barton, N. H. (2011). Mapping mendelian traits in asexual progeny using changes in marker allele frequency. *Genetics Research*, 93(3):221–32.
- Lynch, M., Sung, W., Morris, K., Coffey, N., Landry, C. R., Dopman, E. B., Dickinson, W. J., Okamoto, K., Kulkarni, S., Hartl, D. L., and Thomas, W. K. (2008). A genome-wide view of the spectrum of spontaneous mutations in yeast. *Proc Natl Acad Sci USA*, 105(27):9272–7.
- Macpherson, J. M., Sella, G., Davis, J. C., and Petrov, D. A. (2007). Genomewide spatial correspondence between nonsynonymous divergence and neutral polymorphism reveals extensive adaptation in *drosophila*. *Genetics*, 177(4):2083–99.
- Mancera, E., Bourgon, R., Brozzi, A., Huber, W., and Steinmetz, L. M. (2008). High-resolution mapping of meiotic crossovers and non-crossovers in yeast. *Nature*, 454(7203):479–85.
- Muller, H. J. (1932). Some genetic aspects of sex. *The American Naturalist*, 66(703):118–138.
- Mustonen, V. and Lässig, M. (2007). Adaptations to fluctuating selection in *drosophila*. *PNAS*, 104(7):2277–82.
- Mustonen, V. and Lässig, M. (2009). From fitness landscapes to seascape: non-equilibrium dynamics of selection and adaptation. *Trends Genet*, 25(3):111–9.
- Mustonen, V. and Lässig, M. (2010). Fitness flux and ubiquity of adaptive evolution. *Proc Natl Acad Sci USA*, 107(9):4248–53.
- Neher, R. A. and Shraiman, B. I. (2009). Competition between recombination and epistasis can cause a transition from allele to genotype selection. *Proc Natl Acad Sci USA*, 106(16):6866–71.
- Park, S.-C., Simon, D., and Krug, J. (2010). The speed of evolution in large asexual populations. *J Stat Phys*, 138(1-3):381–410.
- Parts, L., Cubillos, F. A., Warringer, J., Jain, K., Salinas, F., Bumpstead, S. J., Molin, M., Zia, A., Simpson, J. T., Quail, M. A., Moses, A., Louis, E. J., Durbin, R., and Liti, G. (2011). Revealing the genetic structure of a trait by sequencing a population under selection. *Genome Res*.
- Perfeito, L., Fernandes, L., Mota, C., and Gordo, I. (2007). Adaptive mutations in bacteria: high rate and small effects. *Science*, 317(5839):813–5.
- Ruderfer, D. M., Pratt, S. C., Seidel, H. S., and Kruglyak, L. (2006). Population genomic analysis of outcrossing and recombination in yeast. *Nature Genetics*, 38(9):1077–81.
- Sawyer, S. A., Kulathinal, R. J., Bustamante, C. D., and Hartl, D. L. (2003). Bayesian analysis suggests that most amino acid replacements in *drosophila* are driven by positive selection. *Journal of molecular evolution*, 57 Suppl 1:S154–64.
- Segrè, A. V., Murray, A. W., and Leu, J.-Y. (2006). High-resolution mutation mapping reveals parallel experimental evolution in yeast. *PLoS Biol*, 4(8):e256.
- Sella, G., Petrov, D. A., Przeworski, M., and Andolfatto, P. (2009). Pervasive natural selection in the *drosophila* genome? *PLoS Genet*, 5(6):e1000495.
- Smith, J. M. and Haigh, J. (1974). The hitch-hiking effect of a favourable gene. *Genet Res*, 23(1):23–35.
- Sniegowski, P. D. and Gerrish, P. J. (2010). Beneficial mutations and the dynamics of adaptation in asexual populations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1544):1255–63.
- Sveteć, N., Pavlidis, P., and Stephan, W. (2009). Recent strong positive selection on *drosophila melanogaster* *hdac6*, a gene encoding a stress surveillance factor, as revealed by population genomic analysis. *Molecular Biology and Evolution*, 26(7):1549–56.
- Weigand, M. R. and Sundin, G. W. (2009). Long-term effects of inducible mutagenic dna repair on relative fitness and phenotypic diversification in *pseudomonas citrorii* 302959. *Genetics*, 181(1):199–208.
- Weinreich, D., Watson, R., and Chao, L. (2005). Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. *Evolution*, 59(6):1165–1174.
- Woods, R. J., Barrick, J. E., Cooper, T. F., Shrestha, U., Kauth, M. R., and Lenski, R. E. (2011). Second-order selection for evolvability in a large *escherichia coli* population. *Science*, 331(6023):1433–6.