A method to infer selective sweeps from reduced diversity at a linked microsatellite locus

Paulo R. A. Campos*, Christoph Adami*†, and Claus O. Wilke*

December 12, 2002

*Digital Life Laboratory, Mail Code 136-93, Caltech, Pasadena, CA 91125
†Jet Propulsion Laboratory, Mail Code 126-347, Caltech, Pasadena, CA 91109

Classification: Biological Sciences/Evolution

Corresponding author: Claus O. Wilke. Phone: 626 395 2258, Fax: 626 395 2944, email: wilke@caltech.edu

Manuscript information: 12 text pages, 5 figures, 114 words in abstract, 31,000 characters total

Abstract

We study the reduction of variability at a microsatellite locus linked to a gene that undergoes a selective sweep. In the literature, this reduction is commonly measured by a decrease in the variance of the microsatellite lengths. However, the variance fluctuates strongly even for a neutrally drifting microsatellite locus, and statistical tests based on a reduction in variance are therefore weak. We propose to measure the reduction of variability with an alternative quantity, the Shannon entropy H. Extensive numerical simulations show that the entropy can detect selective sweeps with high significance, as long as selection is sufficiently strong. We compare the trajectories of variance and entropy in microsatellite data from evolving *Escherichia coli* populations.

Introduction

Microsatellites are regions of non-coding DNA that consist of repeats of short motifs. Microsatellite mutation rates range from 10^{-6} (*Drosophila* [1]) to 10^{-2} (*Escherichia coli* [2]) events per locus per generation. This high variability, combined with their selective neutrality, makes microsatellites an excellent genetic marker [3], which can be used to infer relationships among species [4], population structure [5], demographic parameters [6], or genetic maps in genome projects [7, 8] (see also [9] and references therein).

Many authors have noted that a microsatellite locus (or a neutral locus in general) linked to a gene under positive selection experiences reduced variability within the population [10, 11, 12, 13, 14, 15, 16, 17], studying typically the variance in repeat number as a measure of microsatellite variability. However, even under neutral evolution, the variance fluctuates strongly. Therefore, the confidence intervals around the expected variance at mutation-drift equilibrium are large, which makes statistical tests for reduced microsatellite variance weak [18].

In the context of long-term experimental evolution of bacteria [19, 20], a simple method that could pinpoint the exact generation count at which a selective sweep occurred, and that potentially could even yield the selective advantage realized in that sweep, could be a powerful tool to analyze the evolutionary trajectories of evolving bacteria. In a recent study, Imhof and Schlötterer [21] used a microsatellite marker in the bacterium *Escherichia coli* for exactly this purpose. Instead of studying the variance of repeat numbers, they identified selective sweeps with sudden increases in the frequency of individual repeat lengths. From the rate of increase in frequency, they also determined the selective advantage realized in the selective sweeps.

We show here, through extensive numerical simulations of several different models, that a reduction in the variance of the microsatellite lengths is

not a reliable indicator for selective sweeps. Moreover, we show that a quantity based on information-theoretic considerations, the Shannon entropy H [22], can detect selective sweeps with high significance, as long as the selective advantage of the linked gene under selection is not too small. In comparison to the method employed by Imhof and Schlötterer, which detects even mutations that are later lost to drift, a significant decrease of H occurs only for mutations that reach fixation. Therefore, the entropy based detection of selective sweeps complements their method. We re-analyze the *E. coli* data of Imhof and Schlötterer in the light of these findings.

Materials and Methods

Numerical simulations. We consider genetic sequences that consist of a single microsatellite locus tightly linked to a single gene under selection. We use the fitness of the gene under selection as the fitness of the genetic sequence, and simulate a constant population of N genetic sequences that reproduce in discrete, non-overlapping generations. We employ Wright-Fisher sampling for reproduction, that is, the probability that a sequence i is chosen for a single reproduction event is given by $p_i = w_i / \sum_j w_j$, where w_i is the fitness of sequence i, and the sum runs over all sequences in the population. We use population sizes from $N = 1000$ up to $N = 100,000$.

Simulations of long-term evolution. We use two different models of assigning fitness to the gene under selection. Both are based on an infinite-site model [23, 24], that is, we neglect all back mutations. In the first one, which we call the random energy model (REM) [25, 26], every mutation results in a new gene with random fitness. The fitness is calculated as $w = e^{-\beta X}$, where X is a normally distributed random variable and β is a parameter that governs the average effect of mutations. Throughout this work, we use $\beta = 1$. A detailed discussion of this type of fitness landscape is given in Ref. [27].

In the second model, which we call the staircase model (SM), fitness is a function of the total number of mutations k_i that gene i has accumulated with respect to the wild type. The fitness is given by $w_i = [k_i/K] + 1$, where the square brackets indicate the largest integer smaller than or equal to the expression that they enclose, and K is a parameter that determines how many mutations must have accumulated before fitness is increased by one unit.

In the REM model, most mutations are deleterious, and thus we have strong background selection between selective sweeps of advantageous mutants. In the SM model, on the other hand, all mutations are neutral or beneficial, and background selection is therefore absent.

For both models, we assume Poisson distributed mutations, that is, we introduce k new mutations into each offspring gene, where k is a Poisson

random variable with mean U.

Microsatellites dynamics. We carry out simulations with two standard models of microsatellite evolution, the stepwise mutation model (SMM) [28] and the two-phase model (TPM) [29]. In the SSM, each mutation results in the insertion or deletion of a single repeat at the microsatellite locus. Insertions and deletions are equally likely, and the total probability of mutation is $u_{\rm sat}$ per microsatellite locus and generation.

The TPM is built on top of the SSM. In the TPM, the majority of mutation events are single repeat changes as in the SSM, but occasionally large changes in repeat length can occur. The overall probability of mutation is again $u_{\rm sat}$ per microsatellite locus and generation. A fraction $p_{\rm SMM}$ of all mutations are single repeat changes. The remaining fraction $1 - p_{\rm SMM}$ results in large length changes. The magnitudes of these changes are distributed according to a geometric distribution with variance σ_m^2. Throughout this paper, we use $p_{\rm SMM} = 0.95$ and σ_m^2 between 30 and 50 (these values are comparable to the ones used in Ref. [29]).

Measured quantities. In the simulations, we measure the average fitness, $\langle w \rangle$, the most abundant (dominant) microsatellite length, $n_{\rm dom}$, the variance in microsatellite lengths, ${\rm Var}(n)$, and the Shannon entropy of the microsatellite distribution, H. The Shannon entropy is defined as [22]

$$H = -\sum_n p_n \ln p_n \,, \tag{1}$$

where p_n is the fraction of microsatellites of length n in the population, and the sum runs over all microsatellite lengths that are present. We measure these quantities every 5 generations. For the variance and the entropy, we also record the changes between successive measurements, that is, we calculate $\Delta{\rm Var}(n,t) = {\rm Var}(n,t) - {\rm Var}(n,t-\Delta t)$, with $\Delta t = 5$, and likewise $\Delta H(t) = H(t) - H(t-\Delta t)$.

Sensitivity analysis for H statistic. In our analysis of the influence of the selective advantage s on the magnitude of the entropy reduction, we use a two allele model for the gene under selection, and the TPM model for microsatellite evolution. The wild-type allele of the gene under selection has fitness 1, and the mutant allele has fitness $1 + s$. We seed a population with N identical sequences that have the wild-type allele and microsatellite length of $n = 50$, and let the population equilibrate. (The equilibration time is 1000 generations for $N = 1000$ and 10,000 generations for $N = 100,000$.) Then, we introduce a mutant genotype. This mutant will either go to fixation, with probability $2s$, or will die out. If it does not reach fixation, we re-introduce another advantageous mutant and repeat until the mutant genotype achieves fixation. We repeat this process five times, and measure the maximum drop

in entropy for each selective sweep. If an advantageous mutant does not go to fixation, then it typically does not leave a signature in the entropy, so that we have to measure the entropy drop only for successful selective sweeps.

In order to establish the baseline for the distribution of entropy changes, we measure the evolution of the entropy over 10,000 generations (for $N = 1000$) or 20,000 generations (for $N = 100,000$) for a neutrally drifting microsatellite locus at various mutation rates.

Results

Simulations of long-term evolution. We simulated the long-term evolution of a microsatellite locus linked to a gene under selection as described in the Methods section. Fig. 1 shows a typical example of a simulation run over 10,000 generations, for $N = 1000$. From the evolution of the mean fitness of the population (Fig. 1a), we can see that three selective sweeps occurred, approximately at $t = 2000$, $t = 3000$, and $t = 7500$ generations. In Fig. 1b, we display the dominant microsatellite length, n_{dom}. The dominant microsatellite length fluctuates strongly, with sudden drastic changes that are not temporally correlated to the selective sweeps. From the temporal pattern displayed by n_{dom}, we are not able to identify the selective sweeps.

By contrast, the entropy H (Fig. 1c) shows significant drops at the exact points in time at which selective sweeps occur. This correspondence is even more apparent in the entropy difference between successive measurements, $\Delta H(t)$. The entropy difference (Fig. 1d) lies consistently between -0.2 and 0.2, except for the three time points of the selective sweeps. There, ΔH exceeds -0.5.

Finally, the variance in microsatellite length, $\mathrm{Var}(n)$, and the change in variance between successive measurements, $\Delta\mathrm{Var}(n)$, are again poor indicators of selective sweeps (Fig. 1e and f). Drastic changes in the variance seem to be more closely related to variations in n_{dom} than to selective sweeps. This does not mean that the variance does not decrease substantially when a selective sweep takes place. During the first two selective sweeps in Fig. 1, the variance does undergo a clear reduction. However, drops in variance of similar magnitude occur in the absence of selective sweeps as well. Therefore, a sudden drop in variance does not imply that a selective sweep has taken place.

We carried out a number of simulations of long-term evolution, with different fitness landscapes and microsatellite dynamics. Qualitatively, the pattern we observed was always similar to the one displayed in Fig. 1. We show two examples of this behavior in Figs. 6, 7 in the web supplement. Note that our results seem to be only weakly dependent on the amount of background selection: The trajectories from the REM model, with strong background selection, and from the SM model, with no background selection, are extremely

similar.

While Figs. 1 and 6 show results from simulations with a comparatively small population size of $N = 1000$, Fig. 7 in the web supplement shows a simulation run with $N = 100,000$. The variance behaves more deterministically and fluctuates less for the larger population size. However, substantial fluctuations do build up towards the end of the simulation, and their magnitude exceeds the reductions in variance during the first two selective sweeps. The fluctuations in the entropy, on the other hand, do not increase, and even seem to decrease slightly over time.

Figure 2 shows how H and $\mathrm{Var}(n)$ vary over time in a neutrally drifting microsatellite locus. The variability in the variance is substantial, although the population size in this example is $N = 100,000$. This observation is in agreement with the results of Goldstein et al. [18], who showed that the expected variability in the variance is proportional to N, and that the interval in which the variance lies in 95% of the cases increases rapidly with time. By contrast, the entropy trajectories are much more coherent among the ten runs.

Entropy reduction as a function of the selective advantage. If the selective advantage of a mutant gene is very low, so that the selective sweep takes a long time in comparison to the time scale on which microsatellite variation is generated, then there will be no significant entropy reduction during this selective sweep. Therefore, the entropy-based detection of selective sweeps is only feasible if the selective advantage exceeds a certain magnitude. In order to determine this minimum selective advantage, we studied a model in which the gene under selection can assume only two possible allelic states, wild type with fitness 1, and mutant with fitness $1 + s$.

Figure 3 shows a typical selective sweep in this model. Fig. 3a shows the evolution of the mean fitness, Fig. 3b the change in entropy between successive measurements, ΔH, and Fig. 3c the distribution of ΔH. The distribution of ΔH shows a main Gaussian mode, which stems from the neutral fluctuations in microsatellite lengths, and three outliers. The outliers correspond to the three entropy reductions during the selective sweep. The biggest reduction in entropy lies more than fifteen standard deviations away from the center of the Gaussian mode, and is therefore highly significant.

In Fig. 4, we show the maximum drop in entropy during a selective sweep, $\Delta H_{\max}$, averaged over five independent simulation runs, as a function of the selection strength s. We have plotted $\Delta H_{\max}$ in units of the standard deviation of the Gaussian mode, so that the significance of the entropy reduction is immediately apparent from the graph. For $N = 1000$, $\Delta H_{\max}$ is in the range of two to three standard deviations for $s < 0.1$, which means that approximately 4.3% of the events in the Gaussian mode are of similar magnitude as the maximum entropy reduction during a selective sweep. Therefore, for small populations, we cannot identify selective sweeps from their associated

drops in entropy if s is below 0.1. The entropy becomes much more sensitive for a larger population size. For $N = 100,000$, we can identify selective sweeps down to approximately $s = 0.01$.

The inset of Fig. 4 shows the influence of the microsatellite mutation rate on the entropy reduction during the selective sweep. For $N = 1000$, the largest drops in entropy occur for microsatellite mutation rates of approximately $u_{\text{sat}} = 10^{-2}$. The sensitivity of the entropy to selective sweeps decreases both if the mutation rates are higher and lower. If the microsatellite mutation rate is too high with respect to the duration of the selective sweep, then microsatellite variation is regenerated while the sweep takes place, and a reduction in variability cannot be noticed. On the other hand, when the inverse of the microsatellite mutation rate is on the order of the population size, then the time from the most recent common ancestor of the population to the present is too short to generate sufficient variability in repeat numbers, and therefore the entropy cannot report a reduction in variability. From the data for $N = 100,000$, we see that the sensitivity of H does not decrease with microsatellite mutation rate as long as the population size is sufficiently large.

We also measured the entropy reduction associated with beneficial mutants that got lost to drift, and found that these entropy reductions did not stand out significantly from the main Gaussian mode (data not shown). Thus, only mutants that go to fixation (or at least near fixation) will leave a detectable signature in the entropy.

Application to microsatellite data from *E. coli*. Imhof and Schlötterer propagated ten replicate populations of *E. coli* for 1000 generations, and determined the distribution of microsatellite lengths every 90 generations [21]. We calculated both the entropy H and the variance $\text{Var}(n)$ for all 110 data points (eleven successive samples in ten replicates). Because of the long time span of 90 generations between successive data points, the differences in entropy ΔH, or variance $\Delta\text{Var}(n)$, are not particularly meaningful, and therefore we consider only the absolute quantities H and $\text{Var}(n)$.

In Figs. 8 and 9 in the web supplement, we display H and $\text{Var}(n)$ as functions of time for all ten replicate populations. In some replicates, such as population 1, entropy and variance run in parallel. In others, such as population 3, the two trajectories are very different. We present all data points in a single scatter plot of $\text{Var}(n)$ versus H in Fig. 5. We see that at medium to high entropy, that is, for high diversity in the microsatellite distribution, both small and large values of the variance occur, whereas for small values of the entropy, the variance is typically small. This is in agreement with our assessment from the numerical simulations: The variance will be reduced during a selective sweep, but reductions of a similar magnitude can also be caused by neutral fluctuations in the microsatellite length distribution.

Discussion

We have shown that the entropy H is a reliable indicator for reduced microsatellite variability caused by a selective sweep. Even in the presence of strong background selection, which by itself leads already to a reduction of variability [14, 30], selective sweeps cause a reduction in entropy that clearly exceeds the level of neutral fluctuations. A reduction in the variance in microsatellite length, on the other hand, is a less reliable indicator, because the variance fluctuates substantially for a neutrally drifting microsatellite locus as well. Even for large populations, the variance experiences large fluctuations, and the expected variability of the variance grows over time [18].

Throughout this paper, we have only considered the comparatively simple microsatellite mutation schemes SSM and TPM. More elaborate mutation mechanisms have been proposed, in particular ones in which the mutation probability is proportional to the repeat number [31, 32]. Such mutation mechanisms may change the exact quantitative results that we have reported here, but we are confident that they do not affect our qualitative results. The advantage of the entropy over the variance is that it does not change if the arrangement of frequencies is changed, that is, if the frequency values are reassigned to different repeat numbers, while the variance is extremely sensitive to this transformation. This advantage is unaffected by the mutation scheme of the microsatellites.

From their method of detecting selective sweeps, Imhof and Schlötterer deduced that the selective advantage in their evolving *E. coli* populations was always below 0.06, and had its peak at 0.01 or smaller. At an effective populations size of approx. 5×10^6, the entropy should be sensitive enough to detect these selective sweeps. However, as we have mentioned in the results section, only mutations that go to fixation leave a detectable signature in the entropy trajectory of the population. Therefore, it is not surprising that we find only between 1 and 2 selective sweeps per population, whereas Imhof and Schlötterer detected a total of 65 advantageous mutations.

Our simulation results are valid only for selective sweeps that are well separated in time, that is, we have not considered series of successive selective sweeps or effects of clonal interference [33]. When several selective sweeps happen within a short period of time, then the microsatellite variability is not regenerated inbetween them, and we will not observe additional drops in entropy for the later sweeps. However, the entropy will also not return to its equilibrium value, and therefore successive selective sweeps will result in a reduced entropy for a prolonged period of time. Several of the *E. coli* populations seem to exhibit this behavior, in particular populations 2 and 8 (Fig. 8 in the web supplement).

We conclude that the Shannon entropy H is a useful tool to detect selective sweeps from a reduction in variability at a linked microsatellite locus. With accurate knowledge of the effective population size, the mutation

mechanism of the microsatellites, and their rate of mutation, it may even be possible to determine the selective advantage s from the magnitude of the entropy reduction associated with a particular selective sweep. In a future experiment similar to the one carried out by Imhof and Schlötterer, it might be desirable to determine the distribution of microsatellites more often than every 90 generations, so that the magnitude of the neutral entropy fluctuations can be established accurately. Also, selective sweeps of very large s can conceivably be missed if the spacing between data points is too large, and a higher time resolution will protect against this problem.

Acknowledgments

P. R. A. Campos is supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) under Proj. No. 99/09644-9. C. Adami and C. O. Wilke are supported by the NSF under contract No. DEB-9981397. The work of C.A. was carried out in part at the Jet Propulsion Laboratory, under a contract with the National Aeronautics and Space Administration. We would like to thank M. Imhof and C. Schlötterer for providing us with their *E. coli* data, and C. Schlötterer for helpful comments on an earlier version of this manuscript.

References

[1] Schug, M. D., Mackay, T. F. C. & Aquadro, C. F. (1997) **15**, 99-102.

[2] Levinson, G. & Gutman, G. A. (1987) *Nucleic Acids Research* **15**, 5323-5338.

[3] Jarne, P. & Lagoda, P. J. L (1996) *Trends Ecol. Evol.* **11**, 424–429.

[4] Goldstein, D. B., Linares, A. R., Cavalli-Sforza, L. L. & Feldman, M. W. (1995) *Genetics* **139**, 463-471.

[5] Pritchard, J. K., Stephens, M. & Donnelly, P. (2000) *Genetics* **155**, 945-959.

[6] Pritchard, J. K., Seielstad, M. T., Perez-Lezaun, A. & Feldman, M. W. (1999) *Mol. Biol. Evol.* **16**, 1791-1798.

[7] Weissenbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, J., Millasseau, P., Vaysseix, G. & Lathrop, M. (1992) Nature **359**, 794-801.

[8] Dib, C., Faure, S., Fizames, C., Samson, D., Drouot, N., Vignal, A., Millasseau, P., Marc, S., Hazan, J., Seboun, E., Lathrop, M., Gyapay, G., Morissette, J. & Weissenbach, J. (1996) *Nature* **380**, 152-154.

[9] Goldstein, D. B. & Schlötterer, C. (1999) *Microsatellites: Evolution and Applications*, (Oxford University Press).

[10] Maynard Smith, J. & Haigh, J. (1974) *Genet. Res.* **23**, 23-35.

[11] Ohta, T. & Kimura, M. (1975) *Genet. Res.* **25**, 313-326.

[12] Kaplan, N. L., Hudson, R. R. & Langley, C. H. (1989) *Genetics* **123**, 887-899.

[13] Stephan, W., Wiehe, T. & Lenz, M. W. (1992) *Theor. Popul. Biol.* **41**, 237-254.

[14] Slatkin M. (1995) *Mol. Biol. Evol.* **12**, 473-480.

[15] Wiehe, T. (1998) *Theor. Popul. Biol.* **53**, 272-283.

[16] Barton, N. H. (1998) *Genet. Res.* **72**, 123-133.

[17] Schlötterer, C. (2002) *Genetics* **160**, 753-763.

[18] Goldstein, D. B., Zhivotovsky, L. A., Nayar, K., Linares, A. R., Cavalli-Sforza, L. L. & Feldman, M. W. (1996) *Mol. Biol. Evol.* **13**, 1213–1218.

[19] Lenski, R.E., Rose, M.R., Simpson, S.C. & Tadler, S.C. (1991) *Am. Nat.* **138**, 1315-1341.

[20] Lenski, R. E. & Travisano, M. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 6808-6814.

[21] Imhof, M. & Schlötterer, C. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 1113-1117.

[22] Shannon, C. E. & Weaver, W. (1949) *The Mathematical Theory of Communication*, (University of Illinois Press).

[23] Kimura, M. (1969) *Genetics* **61**, 893-903.

[24] Watterson, G. A. (1975) *Theor. Popul. Biol.* **7**, 256-276.

[25] Franz, S., Peliti, L. & Sellitto, M. (1993) *J. Phys. A: Math. Gen.* **26**, L1195-L1199.

[26] Franz, S. & Peliti, L. (1997) *J. Phys. A: Math. Gen.* **30**, 4481-4487.

[27] Wilke, C. O., Campos, P. R. A. & Fontanari, J. F. (2002) *J. Exp. Zool., Mol. Dev. Evol.* **294**, 274-284.

[28] Kimura, M. & Ohta, T. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 2761-2764.

[29] Di Rienzo, A., Peterson, A. C., Garza, J. C., Valdes, A. M., Slatkin, M. & Freimer, N. B. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 3166-3170.

[30] Charlesworth, B., Morgan, M. T. & Charlesworth D. (1993) *Genetics*, 1289-1303.

[31] Calabrese, P. P., Durrett, R. T. & Aquadro, C. F. (2001) *Genetics* **159**, 839-852.

[32] Webster, M. T., Smith, N. G. C. & Ellegren, H. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 8748-8753.

[33] Gerrish, P. J. & Lenski R. E. (1998) *Genetica* **102/103**, 127-144.

Figure 1: Long-term evolution of a microsatellite locus linked to a gene under selection. The fitness landscape is REM, and the microsatellite dynamics is TPM, with $U = 0.01$, $u_{\mathrm{sat}} = 0.01$, $p_{\mathrm{SMM}} = 0.95$, $\sigma_m^2 = 30$, $\beta = 1.0$, and $N = 1000$. a. Mean fitness in the population. b. Length of the dominant microsatellite allele, n_{dom}. c. Entropy H. d. Change in entropy between successive measurements ΔH. e. Variance of microsatellite lengths $\mathrm{Var}(n)$. f. Change in variance between successive measurements $\Delta\mathrm{Var}(n)$.

Figure 2: Entropy and variance of a neutrally drifting microsatellite locus as a function of time, in ten independent simulations. The microsatellite dynamics is TPM, with $u_{\mathrm{sat}} = 0.01$, $p_{\mathrm{SSM}} = 0.95$, $\sigma_m^2 = 50$. The population size is $N = 100,000$, and the population is seeded with N identical copies of a microsatellite locus with repeat number $n = 50$.

Figure 3: Selective sweep in the two-allele model. The selective advantage of the mutant allele is $s = 0.5$, and $N = 1000$. a. Mean fitness in the population. b. Change in entropy between successive measurements ΔH. c. Frequency distribution of ΔH. Here, ΔH is plotted in units of the standard deviation σ of the frequency distribution of ΔH in the absence of selection, and the solid line is a Gaussian fit to this distribution.

Figure 4: Maximal entropy reduction during a selective sweep ΔH_{max} as a function of the selective advantage s at $u_{\mathrm{sat}} = 0.01$. Circles represent $N = 1000$, and diamonds represent $N = 100,000$. ΔH_{max} is plotted in units of the standard deviation σ of the distribution of ΔH_{max} at the respective microsatellite mutation rate and population size in the absence of selection. The inset shows ΔH_{max} as a function of u_{sat} for $s = 0.2$.

Figure 5: Variance $\mathrm{Var}(n)$ versus entropy H in the ten replicate *E. coli* populations of Imhof and Schlötterer [21]. Each set of symbols represents the data points from one replicate.

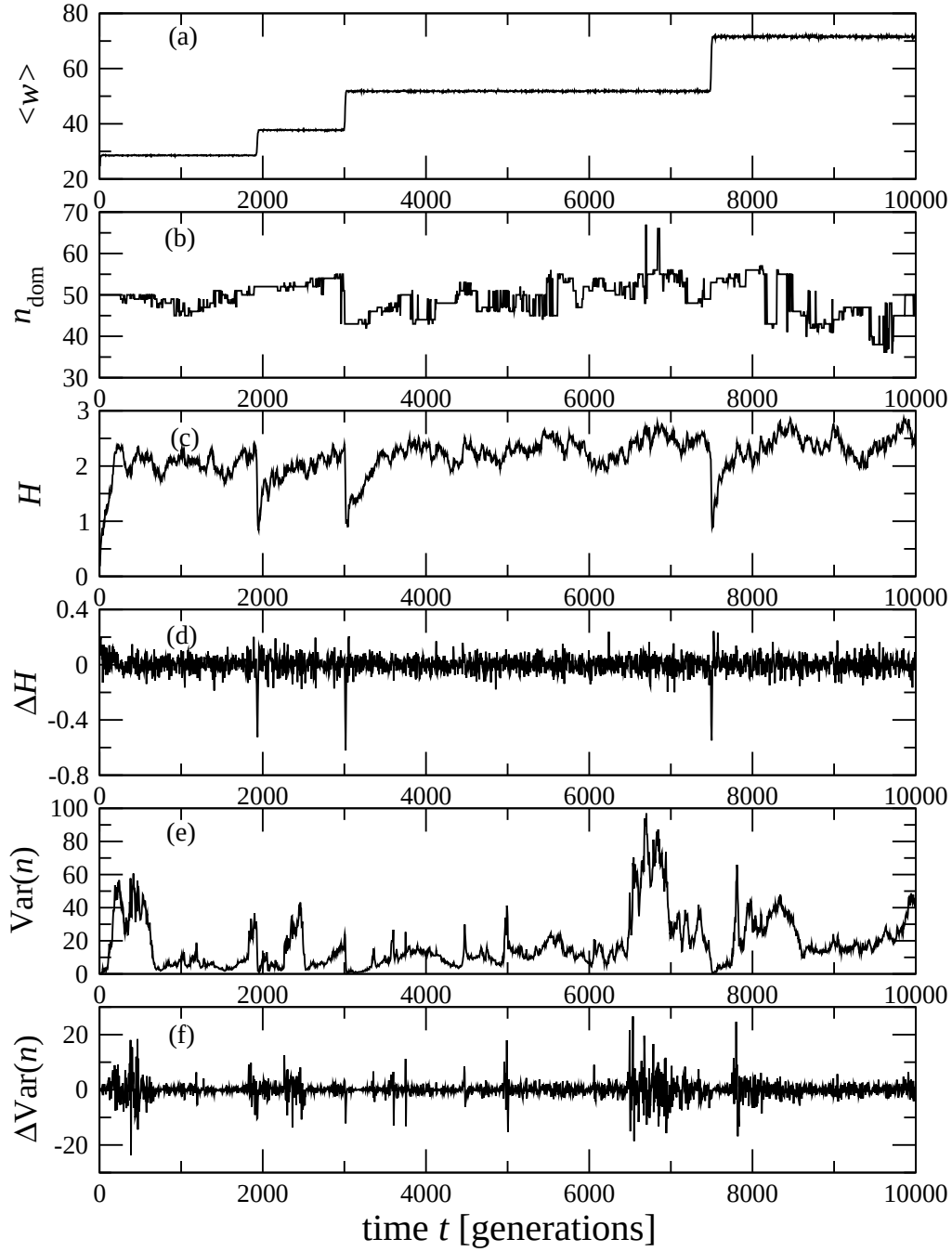

Figure 1

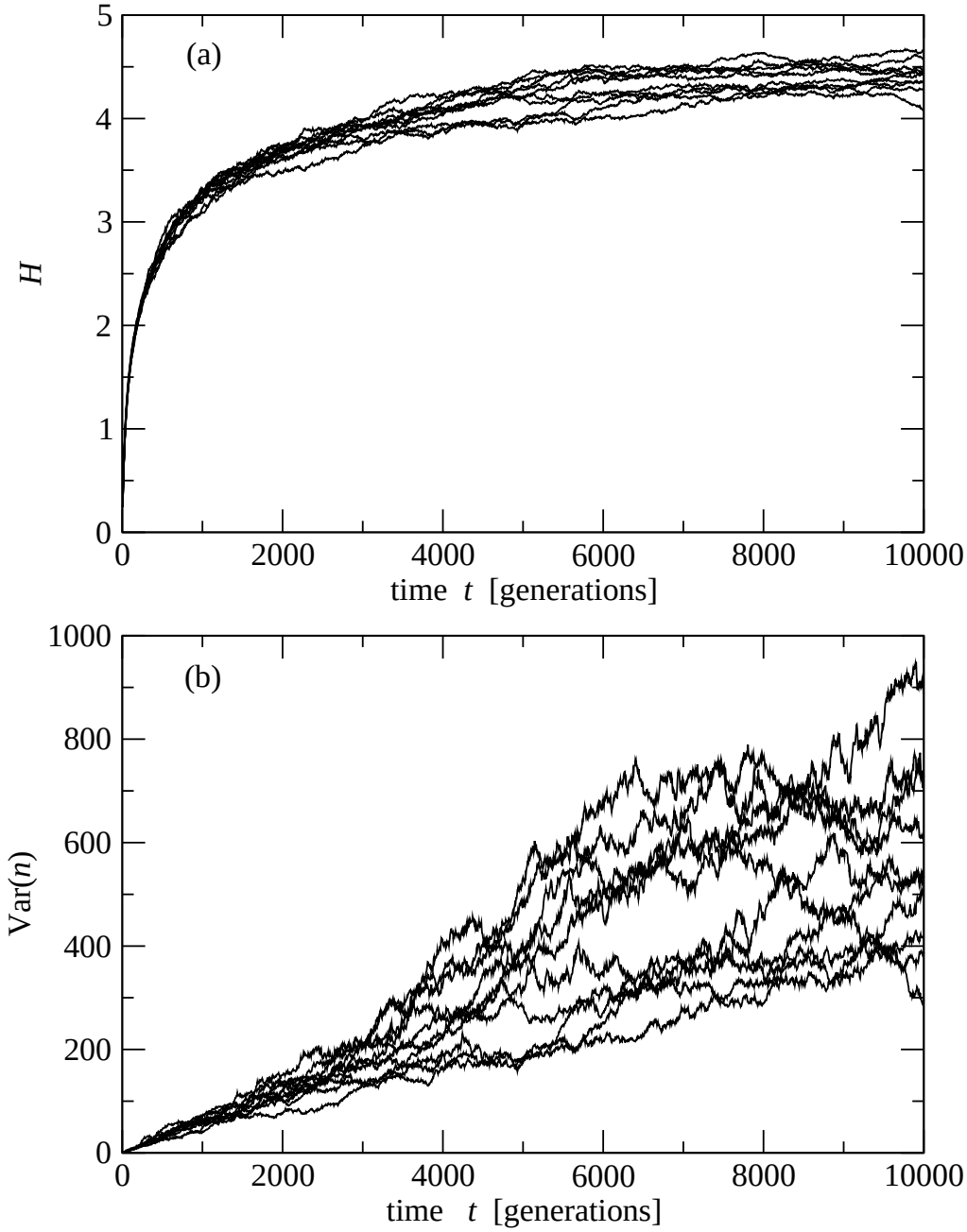

Figure 2

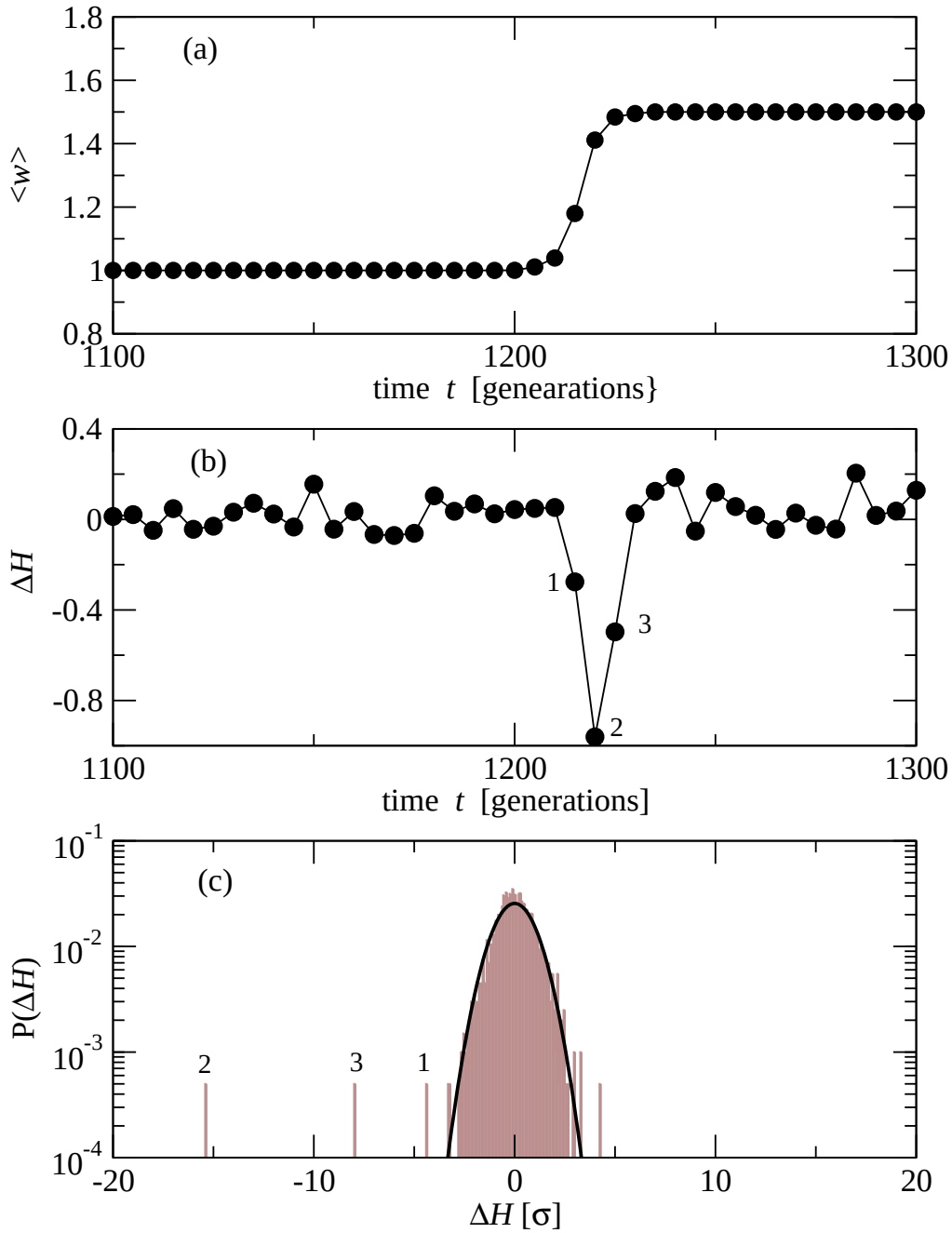
(a)
<w>
1.8
1.6
1.4
1.2
1
0.8
1100
1200
1300
time t [genearations}
(b)
ΔH
0.4
0
-0.4
-0.8
1
2
3
1100
1200
1300
time t [generations]
(c)
P(ΔH)
10^-1
10^-2
10^-3
10^-4
2
3
1
-20
-10
0
10
20
ΔH [σ]

Figure 3

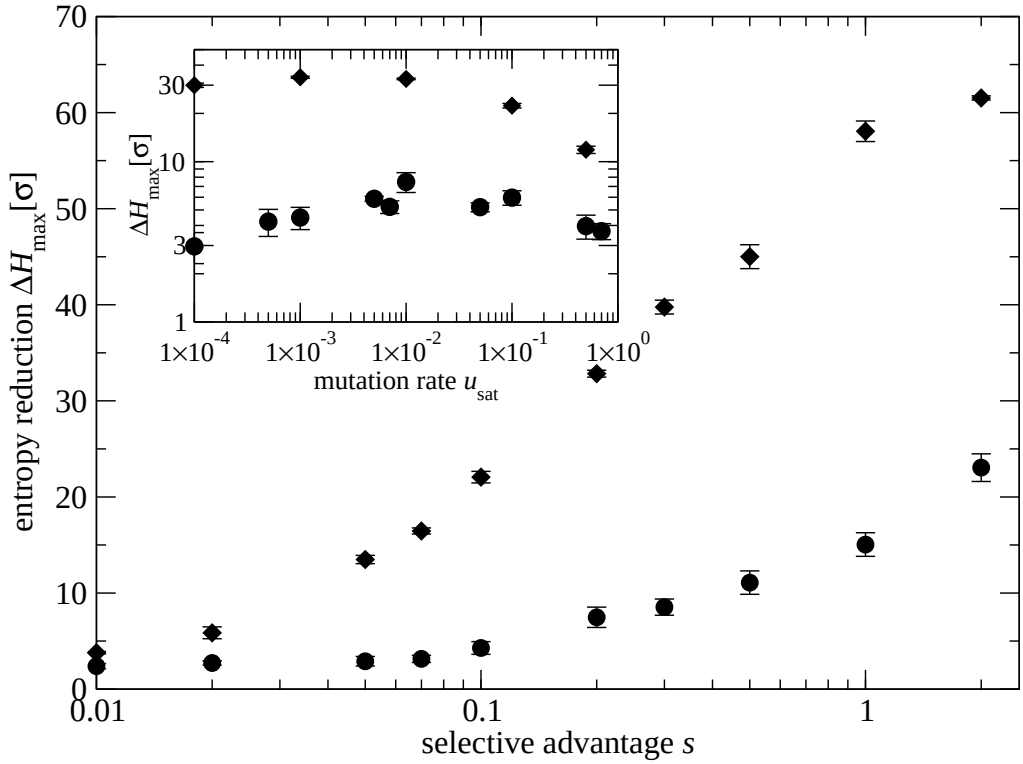
entropy reduction ΔH_{max}[σ]
selective advantage s
0.01
0.1
1
0
10
20
30
40
50
60
70
ΔH_{max}[σ]
1
3
10
30
1×10^{-4}
1×10^{-3}
1×10^{-2}
1×10^{-1}
1×10^{0}
mutation rate u_{sat}

Figure 4

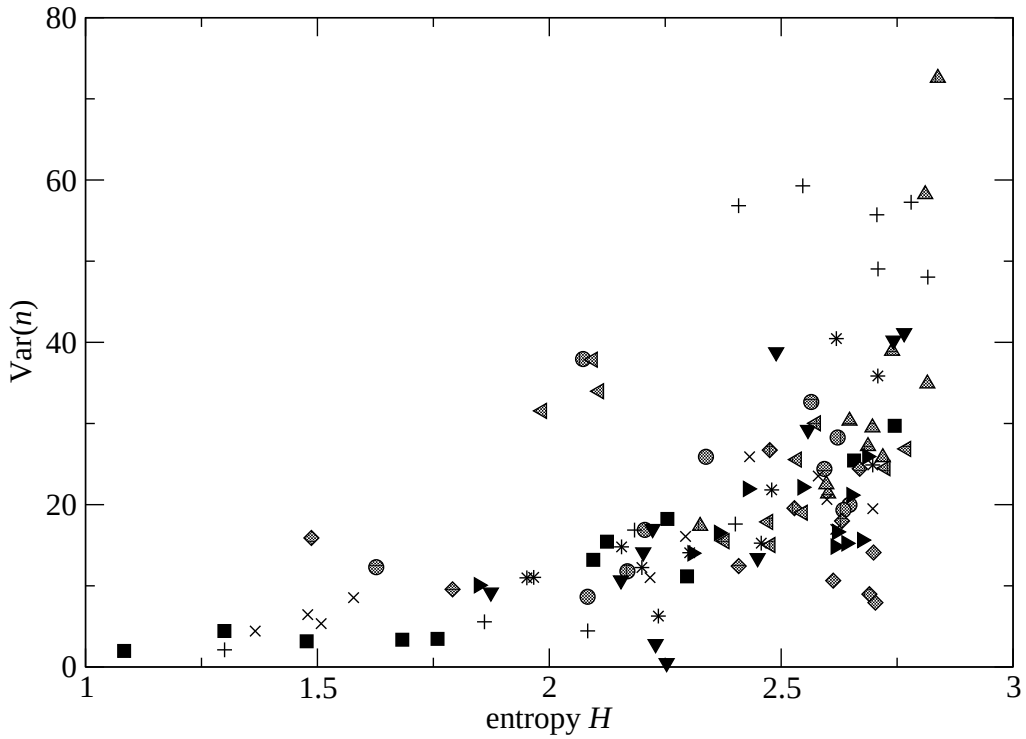

Figure 5