File S1

Supporting Information

Calling trajectories The resetting of fixed mutant alleles led to difficulties in calling trajectories which would not be encountered with biological sequence data, particularly in identifying the end of a trajectory. While, in the identification of the initial point of a trajectory, ignoring sampling effects makes for a pragmatic solution, the time of fixation or death of a polymorphism is more difficult to pinpoint. Indeed, errors at this point can lead to the spurious identification of new trajectories, leading to obvious problems in the later analysis. Here, a cutoff number of observations, c, was defined as

$$c = \max\{300/n_g, 8\} \tag{1}$$

In the case of an apparent death of a polymorphism at some locus i, indicated by a zero sample frequency, the frequencies $\hat{q}_i^1(t_k)$ at the subsequent $c-1$ samples were examined. If a non-zero sample frequency was observed in any of these samples, it was assumed that the apparent death was an artefact of limited sampling, and the polymorphism was assumed to be in existence across the intervening time-points. If no non-zero sample frequencies were observed, the apparent observation of the death of the wild-type allele was assumed to reflect a death in the underlying population. The value of 300 used in the definition of c here reflects the number of observations required to be 95% certain that a polymorphism does not exist at a frequency greater than 1% in the underlying population. In the case of an apparent fixation in the population, marked by a sample frequency value of 1, a slightly different process was required, due to the delayed return in the simulation of fixed mutant alleles to the wild type. Subsequent to an apparent fixation, the allele frequencies at the locus were observed

as above. If a polymorphic frequency between 0 and 0.5 was observed at that locus, it was assumed that the fixed mutant allele had been returned to the wild-type, and a fixation was recorded at the first time of apparent fixation in the locus. If a polymorphic frequency greater than 0.5 was observed at the locus, it was assumed that the initial observation of fixation arose through limited sampling, no fixation having occurred in the intervening time. We note that, at large values of dt_s, this method is not error-free in the histories of polymorphisms called at different loci, leading to a potential worsening of the results reported in the main text for long sampling times. However, when the method is extended to a biological system, with binary alleles replaced by codons, the mutation of a fixed allele back to wild-type would become easily distinguishable from a mutation to a new, third allele at the same locus.

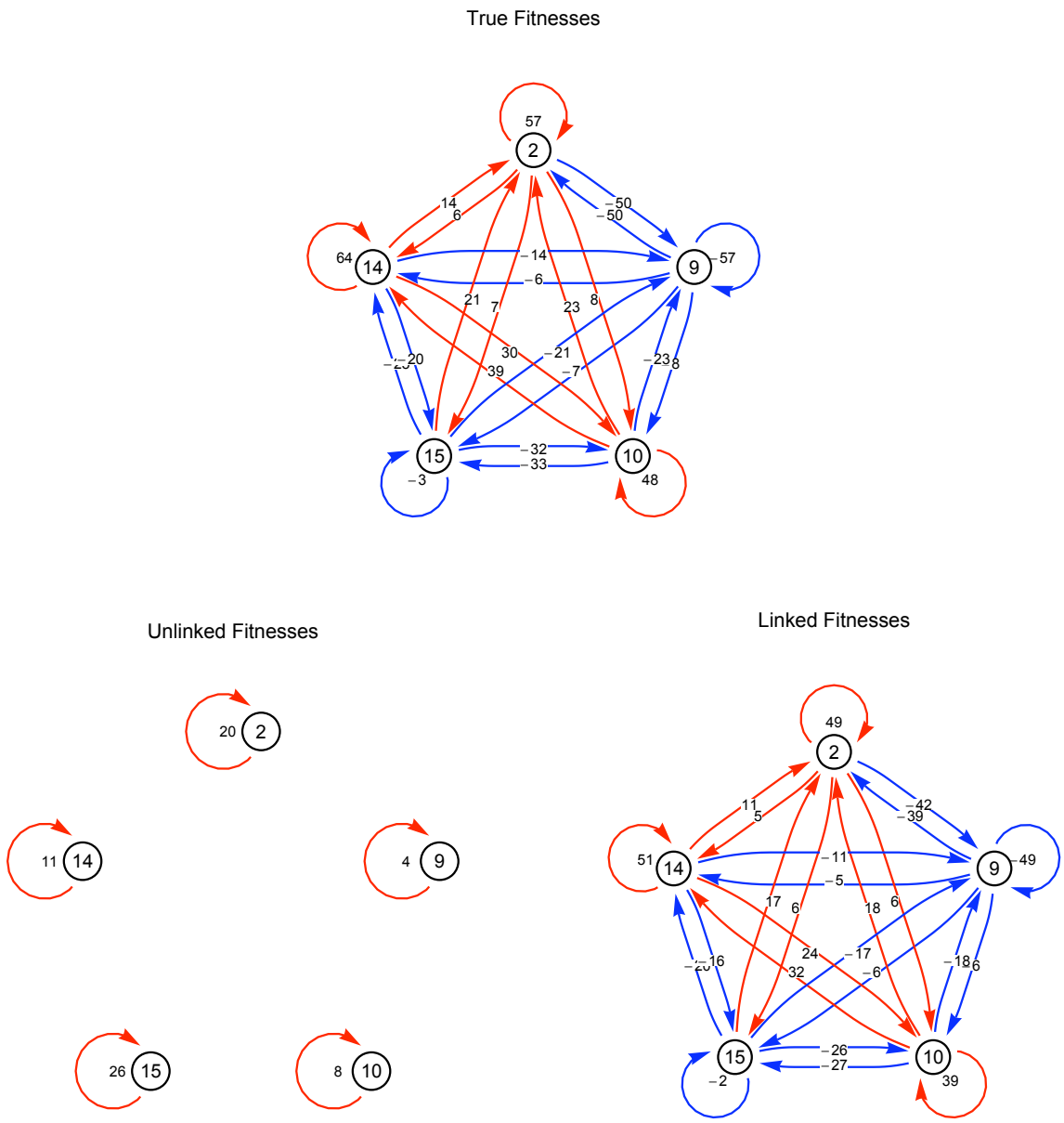

Figure S1 Selection and inter-locus effects in a model system. Details of selection coefficients in the system represented in Figure 2 of the main text. Red edges indicate positive selection and inter-locus effects, while blue edges indicate negative selection and inter-locus effects. The data represented is identical to that in the graphs of Figure 2, albeit with numerical values included. A close fit between the true selection and the inferred selection using linked method can be observed.

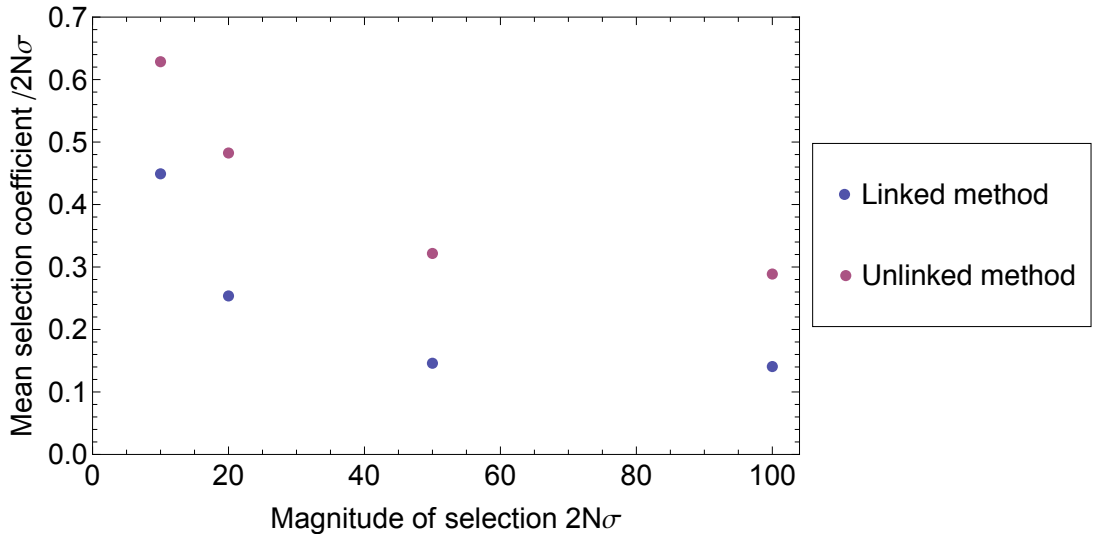

Figure S2 The linked method more accurately measures selection in passenger loci observed to undergo fixation. Mean inferred selection coefficients, in units of $2N\sigma$, assigned to loci from simulations with $T = 500$, $n_g = 100$, and $dt_s = 100$, for various values of the underlying selection coefficient σ, calculated from the models excluding linkage (purple), and including linkage (blue). The correct value in each case is zero. More than 800 inferred selection coefficients are represented by each data point. A similar relationship between selection coefficients is observed for different values of the parameters T, n_g, and dt_s.

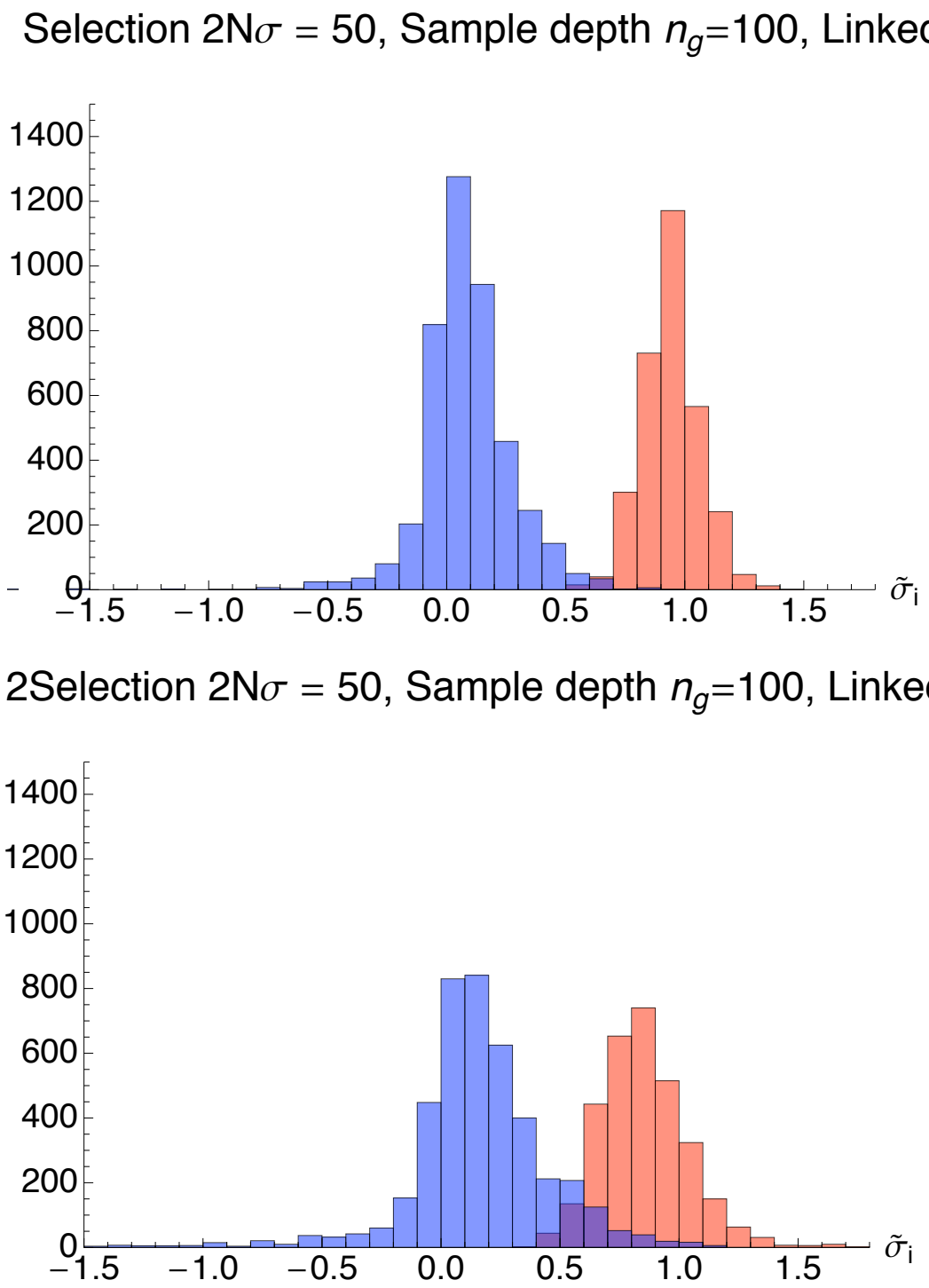

Figure S3 Selection coefficients inferred from lower sampling levels had a greater variance. Inferred selection coefficients, in units of $2N\sigma$, for driver (red), and passenger (blue) selection coefficients for T = 500, dt_s = 100, and n_g equal to 100 and 5 respectively. The greater variance at lower sampling is evident. Outlier passenger coefficients falling lower than the range shown are excluded - these comprise 3.5% of coefficients at n_g 5 and 0.3 % of passenger coefficients at n_g = 100.

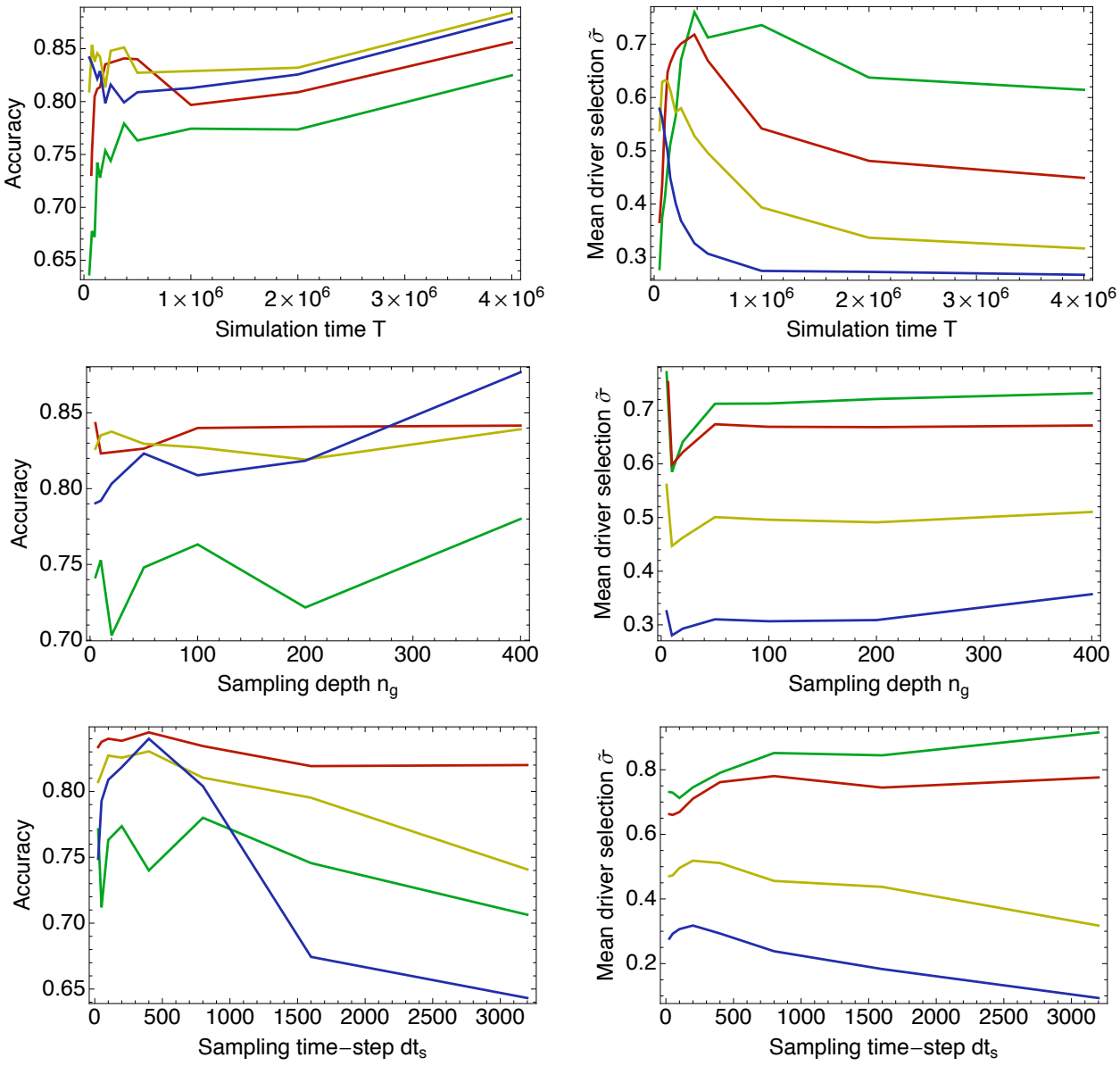

Figure S4 Performance of the unlinked method under varying data collection scenarios. Variation in the accuracy of the method in identifying driver and passenger loci (left column) and in the reproduction of selection coefficients, in units of $2N\sigma$, for driver loci (right column). Default parameters were $T = 5 \times 10^5$, $n_g = 100$, and $dt_s = 100$. Top: Variation in performance under different values of T , the number of generations sampled. Middle: Variation in performance under different values of n_g, the number of individuals sampled at each time point. Bottom: Variation in performance given different values of dt_s, the time between sample points. Data is plotted for values of $2N\sigma$ equal to 100 (blue), 50 (red), 20 (yellow) and 10 (green). Accuracy values are averaged over simulations with 5, 10, 15, 20, and 25 drivers, while mean sigma values are averaged over simulations with 5, 10, 15, 20, 25, and 50 drivers.