# Compound Tests for the Detection of Hitchhiking Under Positive Selection

*Kai Zeng,*[*][1] *Suhua Shi,* *and Chung-I Wu*†

*State Key Laboratory of Biocontrol and Key Laboratory of Gene Engineering of the Ministry of Education, Sun Yat-sen University, Guangzhou, China; and †Department of Ecology and Evolution, University of Chicago

Many statistical tests have been developed for detecting positive selection. Most of these tests draw conclusions based on significant deviations from the patterns of polymorphism predicted by the neutral model. However, many non-equilibrium forces may cause similar deviations, and thus the tests usually have low statistical specificity to positive selection. The main challenge is hence to construct test statistics that are reasonably powerful in detecting positive selection, but are relatively insensitive to other forces. Recently, Zeng et al. (2006) proposed a new test, *DH*, which is a compound of Tajima's *D* and Fay and Wu's *H*, and showed that *DH* has reasonably high statistical specificity to positive selection. In this report, we expand the idea of a compound test by combining Fay and Wu's *H* or *DH* with the Ewens-Watterson (*EW*) test. We refer to these 2 new tests as *HEW* and *DHEW*, respectively. Compared to the *DH* test, *HEW* and *DHEW* are more robust against the presence of recombination, and are also more powerful in detecting positive selection. Furthermore, the *DHEW* test, similar to *DH*, is also relatively insensitive to background selection and demography. The *HEW* test, on the other hand, tends to be somewhat less conservative than *DH* and *DHEW* in some cases.

## Introduction

Clarifying the roles of genetic drift and natural selection in the course of evolution is an issue of fundamental importance in the study of evolutionary biology (Kimura 1983; Gillespie 1991). A widely adopted approach is to derive statistical tests under neutrality (the null hypothesis), and use these tests to detect whether the data is significantly different from the neutral expectation. A significant deviation is usually considered to be a consequence of positive selection. (These tests have been reviewed extensively by many authors: Fay and Wu 2003; Nielsen 2005; Biswas and Akey 2006; Sabeti et al. 2006.) However, the neutral theory is a highly simplified mathematical model with many assumptions (e.g., random mating, constant population size). Thus, a significant deviation can usually be explained by more than one alternative model (Simonsen et al. 1995; Przeworski 2002; Zeng et al. 2006). To make the problem even more difficult, different processes may sometimes leave very similar footprints on genetic variation, and are therefore hard to distinguish. For example, after the fixation of an advantageous allele, the population tends to have an excess of low-frequency variants in linked neutral regions (Braverman et al. 1995; Zeng et al. 2006). However, a similar pattern can also be produced by a recent population expansion (Fu 1997; Zeng et al. 2006). With these complications, it is usually hard to interpret results obtained by these statistical tests. This raises the need for statistical tests that have high specificity to positive selection.

Many methods have been proposed for specifically detecting positive selection. Broadly speaking, these methods are derived by using one of the 4 general approaches described below. The first approach, loosely speaking, is to develop methods in the HKA framework (Hudson et al. 1987; Innan 2006). The underlying principle is that positive selection tends to affect only a small number of loci that are closely linked to the advantageous allele, but demographic changes affect all loci in the genome equally. Usually, demographic changes are directly modeled, and genomic data are then used to estimate the parameters. Genes are considered as targets of positive selection if their patterns of genetic variation are incompatible with those predicted by the estimated demographic parameters (Wall et al. 2002; Akey et al. 2004; Li and Stephan 2006; Thornton and Andolfatto 2006). The difficulty with this approach is that it requires the prior knowledge of the demographic history of the species. But this information is usually not available, or is known only vaguely. Furthermore, it is not clear whether we can obtain the same results if we assume a different demographic model.

The second approach is more empirical. It scans patterns of variation over many loci, and considers loci in the tails of the empirical distribution as candidate targets of selection (Akey et al. 2002; Carlson et al. 2005; Voight et al. 2006; Wang et al. 2006). This approach is appealing in that it avoids specifying a complex demographic model. Nonetheless, the support for positive selection is usually obtained by comparing the empirical distribution with distributions obtained by simulating various demographic models. Furthermore, a recent study suggests that this kind of analysis may suffer from high rates of false positives and false negatives (Teshima et al. 2006).

The third approach relies on a composite likelihood method, which treats nucleotide sites as though they were statistically independent. Methods derived by this approach differ from those mentioned above in that the null hypothesis is not a specific population genetic model, but is derived from the background pattern of variation in the data itself (Jensen et al. 2005; Nielsen et al. 2005).

The fourth approach, which is the focus of this paper, has a simple structure. It requires no estimation of parameters, and is most suitable to data from a single locus. The first test derived by this approach is called the *DH* test (Zeng et al. 2006), which is a compound test combining Tajima's *D* (Tajima 1989) and Fay and Wu's *H* (Fay and Wu 2000). The underlying idea is that Tajima's *D* and Fay and Wu's *H* are both powerful in detecting positive selection, but are sensitive to other processes in a mutually exclusive manner (Zeng et al. 2006). Therefore, by properly combining these

---

2 tests, the resulting test (*DH*) obtains relatively high statistical specificity to positive selection.

The *DH* test, Tajima's *D*, and Fay and Wu's *H* only use the site-frequency spectrum (Fu 1995) for detection. It is known that these methods tend to be too conservative in the presence of intragenic recombination (Wall 1999; Przeworski et al. 2001; Zeng et al. 2007). In the companion study (Zeng et al. 2007), we have shown that these site-frequency methods are usually not as powerful in detecting positive selection as methods based on the haplotype-frequency spectrum such as the Ewens-Watterson (*EW*) test (Watterson 1978). Furthermore, we have shown that the *EW* test is insensitive to recombination, and is usually very powerful in detecting positive selection especially when recombination rate is high. In the light of these recent results, we expand the idea of a compound test by constructing new test statistics which make use of both site- and haplotype-frequency spectra.

In this paper, we propose two new tests by combining Fay and Wu's *H* or the *DH* test with the Ewens-Watterson (*EW*) test. We refer to these two tests as *HEW* (*H* + *EW*) and *DHEW* (*DH* + *EW*), respectively. The effect of positive selection is studied using the model first introduced by Maynard Smith and Haigh (1974). By doing extensive computer simulations, we try to address the following questions: (1) Are *HEW* and *DHEW* more robust against the presence of intragenic recombination than is the *DH* test? (2) Are *HEW* and *DHEW* more powerful than *DH* in detecting positive selection? (3) Are *HEW* and *DHEW* sensitive to background selection and demographic changes?

## Methods
### Statistical Tests

We first describe 4 previously developed tests (Tajima's *D*, Fay and Wu's *H*, Ewens-Watterson's test, and the *DH* test), and then we describe 2 new tests, *HEW* and *DHEW*.

#### Tajima's D and Fay and Wu's H

Tajima's well-known test statistic is defined as:

$$D = \frac{\theta_\pi - \theta_W}{\sqrt{Var(\theta_\pi - \theta_W)}} \qquad (1)$$

(Tajima 1989). In equation (1), $\theta_\pi$ (Tajima 1983) and $\theta_W$ (Watterson 1975) are 2 unbiased estimators of the population mutation rate $\theta$ ($= 4Nu$, where $N$ is the effective population size and $u$ is the neutral mutation rate of the locus).

Fay and Wu's *H* contrasts the abundance of high-frequency variants with the abundance of intermediate-frequency ones (Fay and Wu 2000). Here we use the normalized *H* statistic, defined as:

$$H = \frac{\theta_\pi - \theta_L}{\sqrt{Var(\theta_\pi - \theta_L)}} \qquad (2)$$

(Zeng et al. 2006). $\theta_L$ above is a new unbiased estimator of $\theta$, and by using $\theta_L$, the variance term $Var(\theta_\pi - \theta_L)$ can be derived (Zeng et al. 2006).

#### The Ewens-Watterson test

The test statistic of the Ewens-Watterson (*EW*) test is:

$$F|K = n^{-2} \sum_{i=1}^{K} f_i^2, \qquad (3)$$

where $F$ is the sample homozygosity, $K$ is the number of haplotypes in a sample of size $n$, and $f_i$ is the frequency of the *i*th haplotype (Watterson 1978). Under the no-recombination infinite-allele model, Ewens (1972; see also Karlin and McGregor 1972) showed that the distribution of $F|K$ is independent of $\theta$.

#### The DH test

We denote an observed sample with $S$ ($\geq 1$) segregating sites as $X_{S,obs}$. We say $X_{S,obs}$ is rejected by the *DH* test at the significance level $p$ if

$$D(X_{S,obs}) \leq D_{S,cri} \text{ and } H(X_{S,obs}) \leq H_{S,cri}, \qquad (4)$$

where $D(X)$ and $H(X)$ are, respectively, values of Tajima's *D* and Fay and Wu's *H* for a sample *X*, and $D_{S,cri}$ and $H_{S,cri}$ are critical values satisfying

$$p\{D(X_{S,neu}) \leq D_{S,cri} \text{ and } H(X_{S,neu}) \leq H_{S,cri}\} = p, \qquad (5)$$

and

$$p\{D(X_{S,neu}) \leq D_{S,cri}\} = p\{H(X_{S,neu}) \leq H_{S,cri}\}$$
$$= p* . \qquad (6)$$

$X_{S,neu}$ in the equations above represents random neutral samples with $S$ segregating sites. The probabilities in equations (5) and (6) can be estimated by coalescent simulation with the number of segregating sites fixed (Hudson 1993; Wall and Hudson 2001).

#### The HEW test

Similar to the *DH* test, for a given significance level *p*, the rejection region (*RR*) for *HEW* is:

$$RR = \{X_{\theta,neu} | H(X_{\theta,neu}) \leq H_{\theta,cri} \text{ and } EW(X_{\theta,neu}) \geq EW_{\theta,cri}\}, \qquad (7)$$

where $X_{\theta,neu}$ represents random samples generated with a given $\theta$, $EW(X)$ is the value of the test statistic of the *EW* test for a sample *X*, and the critical values ($H_{\theta,cri}$ and $EW_{\theta,cri}$) satisfying

$$p\{H(X_{\theta,neu}) \leq H_{\theta,cri} \text{ and } EW(X_{\theta,neu}) \geq EW_{\theta,cri}\} = p, \qquad (8)$$

and

$$p\{H(X_{\theta,neu}) \leq H_{\theta,cri}\} = p\{EW(X_{\theta,neu}) \geq EW_{\theta,cri}\} = p* . \qquad (9)$$

These probabilities can be estimated by coalescent simulation (Hudson 1990). In practice $\theta$ is unknown. In our implementation, we set $\theta = \theta_W$ (Watterson 1975).

*The DHEW test*

The rejection region (*RR*) for *DHEW* for a given significance level *p* is:

$$RR = \left\{ X_{\theta,neu} \middle| D(X_{\theta,neu}) \leq D_{\theta,cri}, H(X_{\theta,neu}) \right.$$
$$\leq H_{\theta,cri} \text{ and } EW(X_{\theta,neu}) \quad (10)$$
$$\left. \geq EW_{\theta,cri} \right\}.$$

Similarly, the critical values satisfy the following conditions:

$$p\left\{ D(X_{\theta,neu}) \leq D_{\theta,cri}, H(X_{\theta,neu}) \right.$$
$$\leq H_{\theta,cri} \text{ and } EW(X_{\theta,neu}) \quad (11)$$
$$\left. \geq EW_{\theta,cri} \right\} = p,$$

$$p\left\{ D(X_{\theta,neu}) \leq D_{\theta,cri} \right\} = p\left\{ H(X_{\theta,neu}) \leq H_{\theta,cri} \right\}$$
$$= p\left\{ EW(X_{\theta,neu}) \geq EW_{\theta,cri} \right\}$$
$$= p* .$$
$$(12)$$

Simulating the null distributions of the tests

We obtain the null distribution of the *EW* test by using Slatkin's algorithm to enumerate all possible haplotype-frequency configurations for a given *K* (Slatkin 1994). For *D*, *H*, *HEW* and *DHEW*, we conduct standard coalescent simulation without recombination (Hudson 1990) and with θ estimated by Watterson's estimator (Watterson 1975). Finally, the critical values of the *DH* test are determined using coalescent simulation with the number of segregating sites fixed (Hudson 1993; Wall and Hudson 2001).

All tests are one-sided and are conducted at the 5% significance level. The tail of the null distribution of a test statistic which can maximize its power to detect positive selection is used. Thus, for the *EW* test, values falling into the upper 5% tail are considered significant, and for Tajima's *D* and Fay and Wu's *H*, values falling into the lower 5% tail are considered significant.

Simulating evolutionary processes

We simulate various evolutionary processes using the coalescent method (Hudson 1983, 1990) and the infinite-site model (Kimura 1969). Intragenic recombination is included in all the simulations. The intensity of recombination is measured by $\rho = 4Nr$, where *r* is the recombination rate of the locus per generation. We use the coalescent process with a selective phase to generate random samples under the hitchhiking model. This model assumes that a beneficial mutation arises on a single chromosome, and that the fitnesses of the 3 genotypes (*BB*, *Bb*, and *bb*) are $1 + s$, $1 + h \cdot s$, and 1, respectively, where *B* is the advantageous allele, *b* is the wild type allele, and *s* and *h* are the selection and dominance coefficients. The behavior of the selected allele is mainly determined by the scaled selection coefficient $\alpha \ (= 2Ns)$ and *h*. The trajectory of the advantageous allele is obtained by using the time-reversal property of the diffusion process (Ewens 2004), and the pseudo-sampling device (Kimura 1980). This procedure takes into account the stochastic effects on the change in frequency of the selected allele. It also allows us to explore selected alleles with different levels of dominance. (Only results obtained by assuming $h = 0.5$ have been presented. Other results obtained by assuming $h = 0.1$ or 0.9 are qualitatively similar.) The genealogy is then generated using the structured coalescent method (Braverman et al. 1995; Zeng et al. 2006). Background selection is simulated using the 2-locus model described in Hudson and Kaplan (1994). This model assumes that the deleterious locus is in mutation-selection equilibrium and is not recombining, but that recombination can occur between the neutral locus and the deleterious locus. Here we extend the Hudson and Kaplan model by allowing intragenic recombination within the neutral locus. Finally, we use the algorithms implemented in the software package *ms* (Hudson 2002) to simulate various demographic scenarios.

## Results

In the companion study, we have shown that the *EW* test is more powerful than Tajima's *D* in detecting selection before and around the fixation of an advantageous allele, while under background selection or other demographic perturbations these 2 tests behave similarly (Zeng et al. 2007). Thus, we do not present results of the *EW* test, but will simply mention its contribution to the 2 new compound tests.

Null distributions of the compound tests

Statistically, a compound test transforms a sequence sample into a vector, and uses this vector as the test statistic. The elements of the vector are *p*-values of the component tests. An example of such a transformation is shown in figure 1 for the *DH* test, generated with $n = 50$ and $S = 20$. By definition, the 5% rejection region (the shaded area) is in the lower-left corner of the plot. Interestingly, under neutrality the vectors are largely uniformly distributed, suggesting that the p-values of *D* and *H* may be only weakly correlated (the left panel of figure 1). In contrast, when positive selection is in operation, the vectors tend to cluster in the lower-left corner, resulting in rejections of neutrality (the right panel of figure 1). In table 1, we show the significance levels of the component tests (i.e., *p**'s defined by equations (6), (9) and (12)) of the 3 compound tests under various combinations of parameter values. In all cases, *p** is several-fold higher than *p*, the nominal significance level of a compound test. The *DHEW* test has the highest *p** value, because this test is constrained by 3 different tests. The actual rejection rates of the 3 compound tests under neutrality are presented in table 2. When there is no recombination ($\rho/\theta = 0$), the actual rejection rates of the tests are very close to the nominal significance level (5% in table 2),
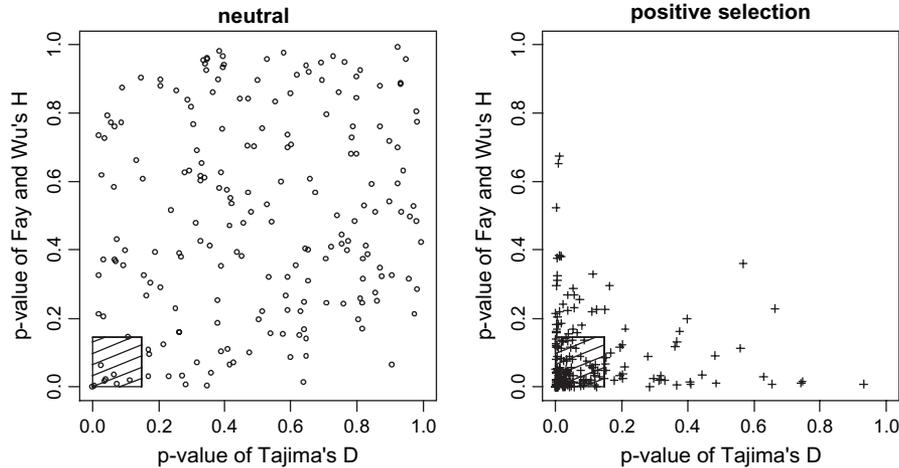
Fig. 1.—The distribution of the *DH* statistic under neutrality or positive selection. Sample size is 50, and each random sample has 20 segregating sites. Each sample is transformed to a vector by calculating the *p*-values of Tajima's *D* and Fay and Wu's *H*. Neutral samples are generated using the standard coalescent method. Positive selection is simulated using the coalescent process with a selective phase (see Methods). Here we assume that a selected site with scaled selection coefficient $\alpha = 300$ and dominance coefficient $h = 0.5$ is located in the middle of a neutral region, and that the frequency of the selected allele at the time of sampling is 80%. No intragenic recombination is included in either model. The shaded region indicates the 5% rejection region of the *DH* test (see also table 1).

suggesting that they are indeed legitimate statistical tests. When intragenic recombination is included, the *DH* test becomes overly conservative, in agreement with previously reports (Wall 1999; Zeng et al. 2007). However, the other 2 tests (and especially *HEW*) are affected by recombination to a much lesser extent. This gain in robustness is due to the *EW* test whose rejection rate has been shown to be insensitive to a change in local recombination rate (Zeng et al. 2007).

Power of the compound tests to detect positive selection

In this section we study the power of the compound tests to detect positive selection, and the differences between compound tests and simple site-frequency tests (i.e., Tajima's *D* and Fay and Wu's *H*). In the simulations, we incorporate intragenic recombination, and assume that $\alpha = 300$, $h = 0.5$. Results obtained by using other parameter values are qualitatively comparable (Supplemental Figure S1). Figure 2 shows the power of the tests as a function

of the size of the neutral region when the advantageous allele is still segregating in the population. In this plot, the selected site is located in the middle of the neutral region. It is clear that *HEW* and *DHEW* are more powerful than the other 3 tests which rely solely on the site-frequency spectrum. This suggests that before fixation the haplotype-frequency spectrum contains useful information about the hitchhiking process, and that ignoring this information would cause a loss of power. At this stage of a sweep, the power of a test to detect positive selection decreases in the following order: *HEW* (the most powerful), *DHEW*, *H*, *DH*, and *D*. However, the difference between *H* and *DH* is not significant. Increasing the local recombination rate does not change this order, but does increase the advantage of *HEW* and *DHEW* over the other tests (figure 2 versus Supplemental figure S2). In general, recombination reduces the size of the region that a selective sweep affects, and lowers the power of the tests, especially that of Tajima's *D* (e.g., compare the fourth panels in figures 2 and S2).

**Table 1**
**Significance Levels, *p\**, of the Component Tests of the Three Compound Tests**

| | $p = 5\%$ | | | $p = 2.5\%$ | | | $p = 1\%$ | | |
|---|---|---|---|---|---|---|---|---|---|
| *S* | *DH* | *HEW* | *DHEW* | *DH* | *HEW* | *DHEW* | *DH* | *HEW* | *DHEW* |
| 5 | 0.196 | 0.207 | 0.268 | 0.112 | 0.135 | 0.167 | 0.067 | 0.081 | 0.097 |
| 10 | 0.173 | 0.174 | 0.234 | 0.104 | 0.113 | 0.152 | 0.051 | 0.067 | 0.089 |
| 20 | 0.147 | 0.180 | 0.242 | 0.084 | 0.115 | 0.160 | 0.038 | 0.064 | 0.095 |
| 30 | 0.136 | 0.180 | 0.252 | 0.077 | 0.121 | 0.172 | 0.033 | 0.070 | 0.099 |
| 40 | 0.131 | 0.190 | 0.264 | 0.073 | 0.123 | 0.188 | 0.033 | 0.074 | 0.110 |
| 50 | 0.121 | 0.195 | 0.270 | 0.068 | 0.131 | 0.186 | 0.031 | 0.077 | 0.114 |

NOTE.—Sample size is 50. *S* is the number of segregating sites, and *p* is the nominal significance level of a compound test. Definitions of *p\** for the 3 compound tests are given in equations (6), (9) and (12), respectively. For *HEW* and *DHEW*, we transform different *S* values into estimates of θ by using Watterson's estimator.

**Table 2**
**Actual Rejection Rates of the Compound Tests Under the Neutral Model**

| | | ρ/θ | | | | | | |
|---|---|---|---|---|---|---|---|---|
| θ | Tests | 0 | 1 | 5 | 7.5 | 10 | 12.5 | 15 |
| 2.5 | *DH* | 0.048 | 0.038 | 0.026 | 0.032 | 0.025 | 0.018 | 0.020 |
| | *HEW* | 0.046 | 0.046 | 0.052 | 0.058 | 0.058 | 0.048 | 0.052 |
| | *DHEW* | 0.051 | 0.048 | 0.044 | 0.040 | 0.037 | 0.029 | 0.034 |
| 5.0 | *DH* | 0.055 | 0.036 | 0.010 | 0.008 | 0.005 | 0.006 | 0.006 |
| | *HEW* | 0.050 | 0.044 | 0.044 | 0.044 | 0.040 | 0.038 | 0.038 |
| | *DHEW* | 0.046 | 0.040 | 0.024 | 0.022 | 0.014 | 0.018 | 0.015 |
| 7.5 | *DH* | 0.056 | 0.020 | 0.004 | 0.004 | 0.002 | 0.002 | 0.002 |
| | *HEW* | 0.049 | 0.046 | 0.030 | 0.036 | 0.040 | 0.026 | 0.025 |
| | *DHEW* | 0.046 | 0.032 | 0.014 | 0.012 | 0.008 | 0.008 | 0.009 |

NOTE.—Sample size is 50. $\theta = 4Nu$ and $\rho = 4Nr$, where *N* is the effective population size, *u* is the neutral mutation rate, and *r* is the recombination rate. The tests are conducted at the 5% significance level.
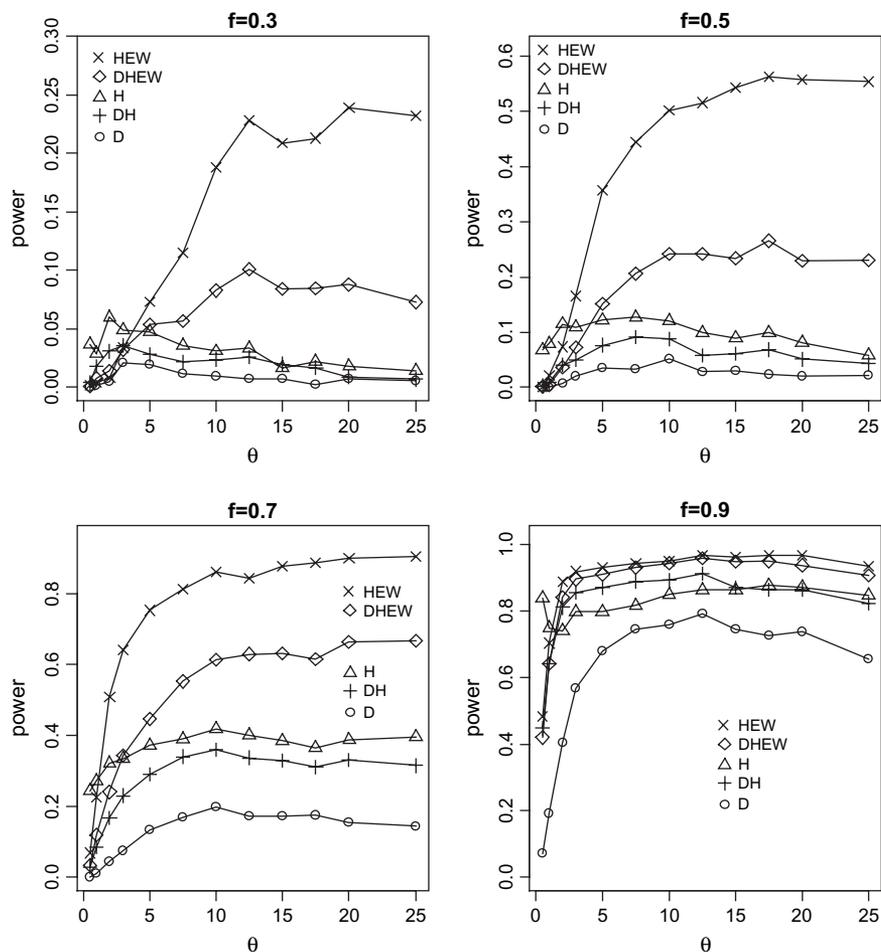
F IG. 2.—Power of the tests to detect positive selection as a function of the size of the neutral region when the advantageous allele is at various frequencies (0.3, 0.5, 0.7 and 0.9 for each panel). Size of the neutral region is measured by θ, the population mutation rate. Positive selection is simulated with α = 300, $h$ = 0.5, and sample size 50. The selected site is located in the middle of the neutral region. Intragenic recombination is included with intensity ρ/θ = 1, where ρ is the population recombination rate of the neutral region. Note that the scales on the y-axes of the plots are different.

*HEW* is less powerful than the *EW* test when the selected allele is at low to intermediate frequency (e.g., less than 70%), but otherwise outperforms it.

After the advantageous allele has been fixed, relative performances of the tests are not always the same as those before fixation. When recombination rate is of the same order as the mutation rate (ρ/θ = 1; figure 3*A*), Tajima's *D* tends to be the most powerful test. In this case, the *DHEW* test outperforms the other 2 compound tests, and the *HEW* test is somewhat better than *DH*. However, if recombination rate is high (ρ/θ = 5; figure 3*B*), the relative performances of the tests resemble those before fixation, i.e. *HEW* tends to be the most powerful, followed by *DHEW*, and then the site-frequency tests. Although the power of the *H* test decreases quickly after fixation (Kim and Stephan 2002; Przeworski 2002; Zeng et al. 2006) results in figure 3 suggest that the rates of decrease in power of the 3 compound tests are much lower than that of the *H* test. As τ (time after fixation, measured in units of 4*N* generations) becomes larger, the special patterns of polymorphism and LD (linkage disequilibrium) due to hitchhiking disappear, and the population exhibits an excess of low- and intermediate-frequency mutations (Zeng et al. 2006). At this stage,

Tajima's *D* is usually the most powerful test (Kim and Stephan 2002; Przeworski 2002; Zeng et al. 2006). However, this pattern of polymorphism is hard to distinguish from that due to population growth, or more generally, from that observed in the recovery phase of a process that causes a loss of polymorphism (Fu 1997; Zeng et al. 2006, 2007). As a result, the compound tests have little power at this stage. (As we shall see later, the compound tests are sensitive primarily to positive selection, and are relatively insensitive to other perturbations.) For this reason, we do not further discuss the post-fixation stage here; some exploratory results are shown in Supplemental figure S3.

It has been shown that the patterns of polymorphism and LD caused by hitchhiking are highly dependent on the distance between the neutral region and the site under selection (Kim and Stephan 2002; Kim and Nielsen 2004; Stephan et al. 2006). To get a better understanding of the performances of the tests, we simulate a neutral locus with scaled mutation rate θ = 10, and place this locus in regions at various distances from the site under selection (figure 4). In this plot, the frequency of the selected allele is 90%, and the distance between the neutral locus and the selected locus is measured by $C_{bet}/s$, where $C_{bet}$ is the
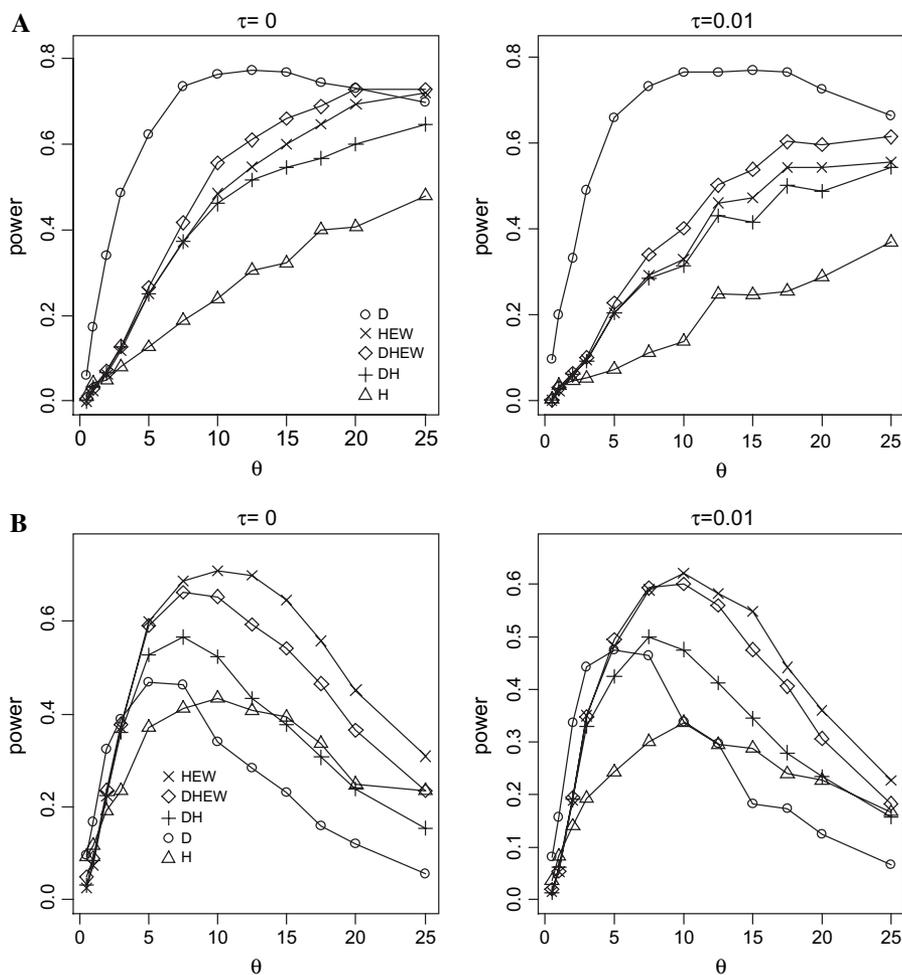
Fig. 3.—Power of the tests to detect positive selection as a function of the size of the neutral region at different time points τ (measured in units of 4N generations) after fixation. Size of the neutral region is measured by θ, the population mutation rate. Positive selection is simulated with α = 300, h = 0.5, and sample size 50. The selected site is placed in the middle of the neutral region. Intragenic recombination is included with 2 different intensities: ρ/θ = 1 in plot A, and ρ/θ = 5 in plot B, where ρ is the population recombination rate of the neutral region. Note that the scales on the y-axes of the plots are different.

recombination distance between the 2 loci, and *s* is the selection coefficient. The 2 new compound tests (i.e., *HEW* and *DHEW*) are in most cases more powerful than *DH*, Tajima's *D*, and Fay and Wu's *H*. In practical terms, *HEW* and *DHEW* are able to detect a sweep further away than are the other tests. A similar picture is observed shortly after fixation (Supplemental figure S4).

Sensitivity to background selection

Purifying selection against linked deleterious mutations maintained by recurrent mutation is referred to as background selection (Charlesworth et al. 1993). It is argued that this kind of selection may be prevalent in many organisms; for example, *D. melanogaster* (Hudson and Kaplan 1995; Charlesworth 1996). Furthermore background selection can skew the frequency spectrum and result in significant Tajima's *D* and Ewens-Watterson's test (Charlesworth et al. 1995; Fu 1997; Zeng et al. 2006, 2007). In table 3, we show the rejection rates of the tests under the background selection model. It is clear that none of the 3 compound tests are affected, despite the high levels of sensitivity of Tajima's *D* and the *EW* test (table 3; Zeng et al. 2007). This insensitivity is due to Fay and Wu's *H*, which is not sensitive to background selection, and therefore counteracts the effects of *D* and *EW*. Although background selection may be a common feature in the genome, it may not cause serious problems in practice. Several studies and the results from our own simulations suggest that background selection can produce significantly negative Tajima's *D* only under some special combinations of parameter values (Golding 1997; Przeworski et al. 1999). Furthermore, in some organisms, for example Drosophila, purifying selection may be too strong to generate significant Tajima's *D* (Andolfatto and Przeworski 2001).

The effects of population demography

Demographic changes may be the most visible violation of the neutral assumptions. Therefore, discrimination between rejections due to demography and those due to positive selection is very important.
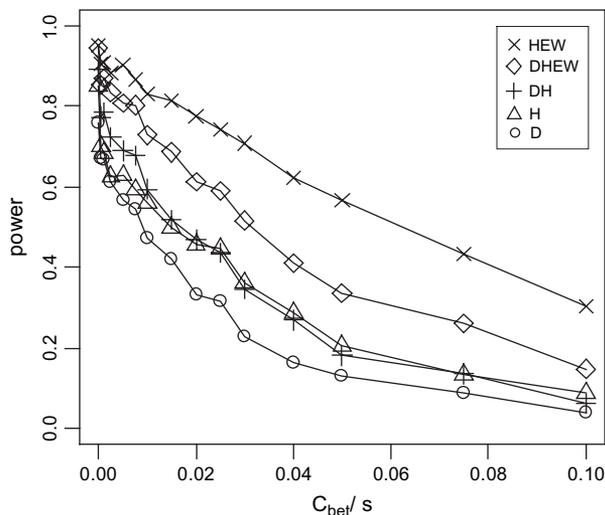
FIG. 4.—Power to detect positive selection as a function of the scaled distance between the selected site and the neutral region ($C_{bet}/s$) when the frequency of the advantageous allele is 90%. Positive selection is simulated with $\alpha = 300$, $h = 0.5$. Sample size is 50. The population mutation ($\theta$) and recombination ($\rho$) rates of the neutral region are both 10. We assume the selected site is on the left-hand side of the neutral region. $C_{bet}$ is the recombination distance between the left end of the neutral region and the selected site. Note that $C_{bet}/s = 0$ means that the selected site is in the middle of the neutral region.

## Population growth

We use the following model to study the effect of population growth: the population size at present is $N$; going backward in time, the population size decreases instantly

**Table 3**
**Sensitivities of the Tests to Background Selection**

| | | | Rejection rate | | | | |
|---|---|---|---|---|---|---|---|
| N | U | $C_{bet}$ | D | H | DH | HEW | DHEW |
| 5,000 | 0.01 | $5.0 \times 10^{-5}$ | 0.036 | 0.053 | 0.048 | 0.068 | 0.052 |
| | | $2.5 \times 10^{-4}$ | 0.034 | 0.047 | 0.046 | 0.053 | 0.047 |
| | | $5.0 \times 10^{-4}$ | 0.037 | 0.030 | 0.036 | 0.070 | 0.052 |
| | 0.10 | $5.0 \times 10^{-5}$ | 0.184 | 0.031 | 0.053 | 0.054 | 0.057 |
| | | $2.5 \times 10^{-4}$ | 0.172 | 0.025 | 0.036 | 0.046 | 0.045 |
| | | $5.0 \times 10^{-4}$ | 0.180 | 0.052 | 0.049 | 0.047 | 0.053 |
| | 0.15 | $5.0 \times 10^{-5}$ | 0.257 | 0.038 | 0.028 | 0.029 | 0.028 |
| | | $2.5 \times 10^{-4}$ | 0.256 | 0.021 | 0.021 | 0.022 | 0.025 |
| | | $5.0 \times 10^{-4}$ | 0.254 | 0.038 | 0.029 | 0.029 | 0.029 |
| 10,000 | 0.01 | $5.0 \times 10^{-5}$ | 0.027 | 0.039 | 0.031 | 0.065 | 0.046 |
| | | $2.5 \times 10^{-4}$ | 0.024 | 0.044 | 0.031 | 0.062 | 0.038 |
| | | $5.0 \times 10^{-4}$ | 0.024 | 0.035 | 0.035 | 0.068 | 0.053 |
| | 0.10 | $5.0 \times 10^{-5}$ | 0.109 | 0.045 | 0.029 | 0.031 | 0.028 |
| | | $2.5 \times 10^{-4}$ | 0.115 | 0.025 | 0.041 | 0.032 | 0.038 |
| | | $5.0 \times 10^{-4}$ | 0.118 | 0.044 | 0.024 | 0.024 | 0.028 |
| | 0.15 | $5.0 \times 10^{-5}$ | 0.161 | 0.030 | 0.032 | 0.026 | 0.032 |
| | | $2.5 \times 10^{-4}$ | 0.171 | 0.027 | 0.014 | 0.011 | 0.014 |
| | | $5.0 \times 10^{-4}$ | 0.165 | 0.021 | 0.015 | 0.013 | 0.013 |

NOTE.—Background selection is simulated using the Hudson and Kaplan model with some modification (see Methods). $N$ is the population size, and $U$ is the deleterious mutation rate per generation per diploid genome. The product of selection and dominance coefficients $sh$ is set to the conventional value of 0.02 (e.g., Charlesworth et al. 1993, 1995; Fu 1997). We assume the neutral locus is on the right-hand side of the deleterious region. $C_{bet}$ is the recombination distance between the deleterious region and the left end of neutral locus. The population mutation rate of the neutral locus is $\theta = 4Nu = 5$. Intragenic recombination at the neutral locus is incorporated with rate $\rho = 4Nr = 5$. Sample size is 50.
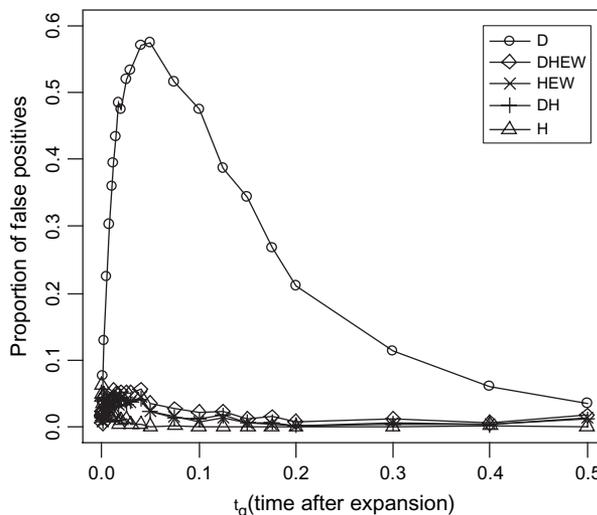


FIG. 5.—The sensitivities of the tests to population growth. We assume that the population mutation rate $\theta$ at the time of sampling is 5. Going backwards in time, $\theta$ decreases instantly at time $t_g$ (in units of $4N$ generations), where $N$ is the population size at the time of sampling) to $\theta_0 = 0.25$ (i.e., the population size decreases 20-fold). Intragenic recombination is incorporated with rate $\rho = 4Nr = 5$. Sample size is 50.

to $N_0$ at time $t_g$ (in units of $4N$ generations). In figure 5, we show the results obtained by assuming $N_0/N = 0.05$. It is clear that the compound tests are largely unaffected. In contrast, the power of the $D$ test can be as high as 60%. The $EW$ test and Tajima's $D$ behave similarly in this case (unpublished data).

## Population bottleneck

Here we model bottleneck in the following way. The population size at present is $N$. Going backward in time the population size reduces exponentially to $\gamma N$. The length of this size-reducing period is $t_b$ (in units of $4N$ generations). Then the population size is restored to $N$ instantaneously (i.e., at time $t_b$ in the past). Summarizing the results of our simulations, we find the following general trends. (1) The $DH$ and $DHEW$ tests behave similarly, and they are in most cases the most conservative tests. (2) When the reduction in population size is mild (e.g., $\gamma \geq 0.25$), none of the tests have a rejection rate higher than 10%. (3) For an intermediate level of size contraction (e.g., $0.001 < \gamma < 0.25$), the compound tests may show limited sensitivity. In this case, $HEW$ tends to be less conservative than $DH$ and $DHEW$ (e.g., figure 6A). However, Tajima's $D$, and to a lesser extent, Fay and Wu's $H$, are usually much more sensitive than the compound tests (figure 6A). (4) Severe bottlenecks (e.g., $\gamma \leq 0.001$) have no significant effect on the compound tests, although Fay and Wu's $H$ and Tajima's $D$ are both highly sensitive (e.g., figure 6B). It has been argued that patterns of variation caused by a bottleneck may be hard to distinguish from those caused by hitchhiking (Barton 1998; Depaulis et al. 2003). The above results suggest that, under the model we assume, the compound tests, especially $DH$ and $DHEW$, are more capable of discriminating between positive selection and bottleneck than are the other tests.
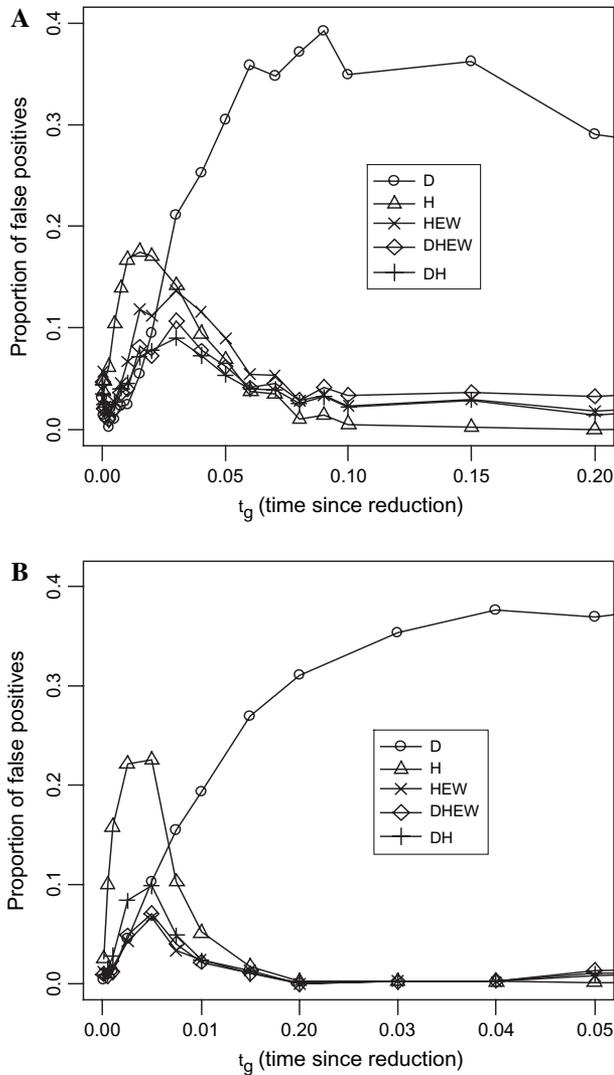
FIG. 6.—The sensitivities of the tests to population bottleneck. We assume that the population mutation rate $\theta$ at the time of sampling is 5. Going backwards in time, $\theta$ decreases exponentially to $\theta_0 = 0.1$ (i.e., $\gamma = \theta_0/\theta = 0.01$; plot A) or $\theta_0 = 0.01$ (i.e., $\gamma = \theta_0/\theta = 0.001$; plot B). The length of this size-reducing period is denoted as $t_b$ (measured in $4N$ generations, where $N$ is the population size at the time of sampling). Intragenic recombination is incorporated with rate $\rho = 4Nr = 5$. Sample size is 50. Note that the scales on the $x$-axes are different.

## *Population subdivision*

In this section, we study the effects of population structure on the tests using the simple symmetric island model (Wright 1931) with different numbers of subpopulations (demes). In table 4, we assume all sequences are taken from 1 deme in a subdivided population with different levels of differentiation. In general, *DH* and *DHEW* tend to behave similarly, but *DHEW* is slightly more conservative than *DH*. The rejection rates of these 2 tests are not affected by subdivision when $F_{ST} \leq 0.25$ (this is also true under other sampling schemes; unpublished data). *HEW*, on the other hand, is not as robust as *DH* and *DHEW*, though it is affected to a lesser extent than Fay and Wu's *H*. Usually *HEW*'s rejection rate is reasonably low when $F_{ST} \leq 0.2$. In

**Table 4**
**Sensitivities of the tests to population subdivision when all samples are taken from one deme**

| No. of demes | $F_{ST}$ | $4N_s m$ | Rejection rate | | | | |
| | | | *D* | *H* | *DH* | *HEW* | *DHEW* |
|---|---|---|---|---|---|---|---|
| 2 | 0.30 | 0.583 | 0.059 | 0.187 | 0.104 | 0.109 | 0.080 |
| | 0.25 | 0.750 | 0.045 | 0.152 | 0.087 | 0.105 | 0.072 |
| | 0.20 | 1.000 | 0.043 | 0.135 | 0.074 | 0.095 | 0.066 |
| | 0.15 | 1.417 | 0.032 | 0.097 | 0.057 | 0.082 | 0.054 |
| 3 | 0.30 | 1.037 | 0.061 | 0.195 | 0.105 | 0.109 | 0.084 |
| | 0.25 | 1.333 | 0.045 | 0.168 | 0.079 | 0.099 | 0.070 |
| | 0.20 | 1.778 | 0.028 | 0.122 | 0.060 | 0.093 | 0.056 |
| | 0.15 | 2.519 | 0.021 | 0.104 | 0.047 | 0.078 | 0.046 |
| 4 | 0.30 | 1.313 | 0.050 | 0.215 | 0.097 | 0.114 | 0.082 |
| | 0.25 | 1.688 | 0.037 | 0.163 | 0.071 | 0.109 | 0.065 |
| | 0.20 | 2.250 | 0.031 | 0.135 | 0.065 | 0.101 | 0.062 |
| | 0.15 | 3.188 | 0.024 | 0.105 | 0.050 | 0.075 | 0.045 |

NOTE.—We simulate the symmetric finite island model with different numbers of demes. Sample size is 50. All samples are taken from 1 deme. We fix the total population mutation rate $\theta_T = 4N_s u d = 5$, where $N_s$ is the number of breeding individuals in a deme, and $d$ is the number of demes. Recombination occurs at rate $\rho_s = \theta_s$ in all cases ($\rho_s = 4N_s r$ and $\theta_s = 4N_s u$).

the Supplement, we explore the effects of pooling samples taken from different demes on the rejection rates of the tests (figure S5). We find that even when sampling is quite biased, the type I error rates of the tests are usually unaffected.

## Differences between compound tests and composite likelihood methods

The composite likelihood test due to Nielsen et al. (2005; referred to as the *CL* test) can be seen as an extension of several other similar methods (Kim and Stephan 2002; Kim and Nielsen 2004; Jensen et al. 2005). This test treats nucleotide sites as though they were statistically independent, and uses the background pattern of variation in the data itself as the null model (Nielsen et al. 2005). Exploratory simulation results (Table 5) suggest that the *CL* test tends to have similar sensitivity to background selection and to several demographic models as the compound tests. However, the *CL* test is not as powerful as the compound tests are in detecting positive selection unless $\theta$ becomes sufficiently large (e.g., $\theta = 50$). Two factors may account for this difference in power. First, the performance of the *CL* test may rely heavily on the accurate inference of the background pattern of variation. However, this inference may require a large amount of data which may not be always available. Second, the lower power of the *CL* test could also be explained by our not having incorporated recombination when simulating *CL*'s null distribution. However, we have been cautious about including recombination in the simulation, since it may make the *CL* test less robust against demographic changes (Nielsen et al. 2005).

## Discussion

In this study, we develop two new compound tests—*HEW* and *DHEW*—by using the novel properties of the Ewens-Watterson (*EW*) test reported in the companion paper (Zeng et al. 2007). By carrying out extensive

**Table 5**
**Comparisons Between the Compound Tests and the Composite Likelihood (*CL*) Test**

| Test | SW1[a] | SW2[a] | SW3[a] | BS[b] | BN[c] | SU[d] |
|------|--------|--------|--------|-------|-------|-------|
| *CL* | 9% | 22% | 61% | 3% | 13% | 8% |
| *DH* | 45.95% | 64.50% | 58.4% | 4.3% | 11.8% | 8.3% |
| *HEW* | 48.35% | 71.80% | 43.2% | 5.1% | 26% | 10.2% |
| *DHEW* | 55.56% | 72.60% | 43.9% | 5.7% | 13.6% | 5.8% |

NOTE.—Sample size is 50. The composite likelihood (*CL*) test is conducted by using the program *SweepFinder* kindly provided by M. Hubisz and R. Nielsen. The power of the *CL* test to detect a specific process is estimated by 100 rounds of simulation. To obtain critical values for the *CL* test, we carry out 500 replicates of coalescent simulation with no recombination and $\theta = \theta_W$ (Watterson's estimator).

[a] Selective sweep is simulated with $\alpha = 300$, $h = 0.5$. The selected site is in the middle of the neutral region. Samples are taken at the time of fixation. In model SW1, $\theta = \rho = 10$; in SW2, $\theta = \rho = 25$; in SW3, $\theta = \rho = 50$.

[b] Background selection is simulated using the following parameters: $N = 10,000$; $U = 0.1$; $sh = 0.02$; $C_{bet} = 2.5 \times 10^{-4}$; $\theta = \rho = 10$.

[c] The population bottleneck model used in Figure 6 is simulated. Here $\gamma = 0.01$, $t_b = 0.03$, and $\theta = \rho = 10$.

[d] A symmetric two-deme model is simulated. $F_{ST} = 0.24$, $\rho_s = \theta_s = 2.5$. All samples are taken from one deme.

**Table 6**
**The effect of incorrect inference of the ancestral state on the type I error rates of the tests**

| | | Rejection rate | | | |
|---|---|---|---|---|---|
| *S* | $P_{mis}$ | *H* | *DH* | *HEW* | *DHEW* |
| 10 | 0.01 | 0.059 | 0.058 | 0.053 | 0.059 |
| | 0.05 | 0.078 | 0.079 | 0.074 | 0.081 |
| | 0.1 | 0.147 | 0.112 | 0.104 | 0.104 |
| | 0.15 | 0.213 | 0.125 | 0.115 | 0.112 |
| 40 | 0.01 | 0.052 | 0.051 | 0.053 | 0.050 |
| | 0.05 | 0.077 | 0.070 | 0.084 | 0.072 |
| | 0.1 | 0.136 | 0.090 | 0.129 | 0.091 |
| | 0.15 | 0.237 | 0.117 | 0.152 | 0.096 |

NOTE.—*S* is the number of segregating sites in a sample. $P_{mis}$ is the probability that the derived allele at a polymorphic site is incorrectly treated as the ancestral allele. Sample size is 50.

simulations, we find that the 3 compound tests (i.e., *DH*, *HEW* and *DHEW*) are reasonably powerful in detecting on-going or recently completed selective sweeps, and are relatively insensitive to background selection and a number of demographic perturbations. In other words, the compound tests have relatively high specificity to positive selection. Among the 3 compound tests, *HEW* is usually the most powerful in detecting positive selection, whereas *DH* and *DHEW* are usually the most insensitive to background selection and demographic changes. Furthermore, by using the fact that the *EW* test is robust against recombination (Zeng et al. 2007), *HEW* and *DHEW* are less sensitive to recombination than is the *DH* test. Having tests that are robust against recombination is desirable for 2 reasons. First, recombination usually reduces the power of the tests to detect positive selection. Second, although it is possible to incorporate recombination into the null model, estimating recombination rate is usually difficult and the estimates are usually not very accurate (Wall 2000).

Compared with Fay and Wu's *H*, the compound tests are better in the following ways: (1) *HEW* and *DHEW* are more robust against the presence of intragenic recombination. (2) *HEW* and *DHEW* are more powerful in detecting positive selection. (3) The power of all 3 compound tests decreases at a much lower rate after fixation. (4) The compound tests, especially *DH* and *DHEW*, are less sensitive to population bottleneck and population subdivision. Another important issue is that the use of Fay and Wu's *H* requires the identification of ancestral/derived alleles at each polymorphic site. This is usually done by using sequences from a closely related species. However, the inference may be inaccurate due to the violation of the infinite-site model. Incorrect inference of the ancestral state usually results in a significantly negative *H* statistic (Baudry and Depaulis 2003). The effect of an incorrect inference on the 3 compound tests is reduced due to the inclusion of Tajima's *D* and/or the *EW* test, both of which are not dependent on the separation of ancestral/derived alleles (table 6). Among the compound tests, the *DHEW* test seems to be the most

robust. In general, if the probability of an incorrect inference per site is lower than 5%, none of the tests tend to be seriously affected.

In the companion study (Zeng et al. 2007), we have shown that the *EW* test is insensitive to the presence of recombination, and is usually very powerful in detecting positive selection especially when recombination is high. However, this test is also very sensitive to background selection, population growth, and population bottlenecks. The results regarding *HEW* and *DHEW* reported above suggest that by using of the idea of a compound test we are able to use the merits of the *EW* test without being affected much by its sensitivity to other driving forces.

It is of some interest to compare the compound tests with tests based on patterns of linkage disequilibrium (e.g., Sabeti et al. 2002). In the companion paper (Zeng et al. 2007), we have shown that when the site under positive selection is known and is used as the core, the extended haplotype homozygosity (*EHH*) test (Sabeti et al. 2002) is very powerful; whereas when this prior information about the selected site is not available, the test has little power. Therefore, the compound tests should be more powerful than the *EHH* test, unless the site under selection is known. Recently, Voight et al. (2006) proposed the integrated haplotype score (*iHS*) test, which can be seen as an extension of the *EHH* test. Unlike the *EHH* test, the *iHS* test has power to detect positive selection when the site under selection is not included in the data set (Voight et al. 2006). However, some preliminary results suggest that the *iHS* test can be less conservative than the compound tests under certain demographic schemes. We are currently preparing a manuscript to further investigate the differences between LD-based methods and the compound tests.

## Conclusion

We have proposed 2 new compound tests, *HEW* and *DHEW*, by combining Fay and Wu's *H* or *DH* with the Ewens-Watterson (*EW*) test. Compared to the *DH* test, these 2 new tests are: (1) more robust against recombination, (2) more powerful in detecting positive selection. Summarizing the results, we conclude that: (1) the 3 compound tests (*DH*, *HEW* and *DHEW*) are most suitable for

detecting ongoing or recently fixed selective sweeps; (2) *DH* and *DHEW* tend to have higher specificity to positive selection than *HEW*. In practice, when only site-frequency data is available, we suggest using the *DH* test due to its high specificity. If haplotype phase is also known, the *DHEW* test may be a good choice since it automatically takes into account the effect of recombination, and also has high specificity to positive selection. The *HEW* test, however, should be used when the population does not show much geographical structure and is not greatly affected by recent bottleneck events.

## Supplementary Material

Supplementary figures S1 through S5 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org).

## Acknowledgments

## Literature Cited

Akey JM, Eberle MA, Rieder MJ, Carlson CS, Shriver MD, Nickerson DA, Kruglyak L. 2004. Population history and natural selection shape patterns of genetic variation in 132 genes. PLoS Biol. 2:e286.

Akey JM, Zhang G, Zhang K, Jin L, Shriver MD. 2002. Interrogating a high-density SNP map for signatures of natural selection. Genome Res. 12:1805–1814.

Andolfatto P, Przeworski M. 2001. Regions of lower crossing over harbor more rare variants in African populations of *Drosophila melanogaster*. Genetics. 158:657–665.

Barton NH. 1998. The effect of hitch-hiking on neutral genealogies. Genet. Res. 72:123–133.

Baudry E, Depaulis F. 2003. Effect of misoriented sites on neutrality tests with outgroup. Genetics. 165:1619–1622.

Biswas S, Akey JM. 2006. Genomic insights into positive selection. Trends Genet. 22:437–446.

Braverman JM, Hudson RR, Kaplan NL, Langley CH, Stephan W. 1995. The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. Genetics. 140:783–796.

Carlson CS, Thomas DJ, Eberle MA, Swanson JE, Livingston RJ, Rieder MJ, Nickerson DA. 2005. Genomic regions exhibiting positive selection identified from dense genotype data. Genome Res. 15:1553–1565.

Charlesworth B. 1996. Background selection and patterns of genetic diversity in *Drosophila melanogaster*. Genet Res. 68:131–149.

Charlesworth B, Morgan MT, Charlesworth D. 1993. The effect of deleterious mutations on neutral molecular variation. Genetics. 134:1289–1303.

Charlesworth D, Charlesworth B, Morgan MT. 1995. The pattern of neutral molecular variation under the background selection model. Genetics. 141:1619–1632.

Depaulis F, Mousset S, Veuille M. 2003. Power of neutrality tests to detect bottlenecks and hitchhiking. J Mol Evol. 57(Suppl. 1):S190–200.

Ewens WJ. 1972. The sampling theory of selectively neutral alleles. Theor Popul Biol. 3:87–112.

Ewens WJ. 2004. Mathematical population genetics. Berlin: Springer-Verlag. pp. 188–191.

Fay JC, Wu C-I. 2000. Hitchhiking under positive Darwinian selection. Genetics. 155:1405–1413.

Fay JC, Wu C-I. 2003. Sequence divergence, functional constraint, and selection in protein evolution. Annu Rev Genomics Hum Genet. 4:213–235.

Fu Y-X. 1995. Statistical properties of segregating sites. Theor Popul Biol. 48:172–197.

Fu Y-X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics. 147:915–925.

Gillespie JH. 1991. The cause of molecular evolution. Oxford: Oxford University Press.

Golding B. 1997. The effect of purifying selection on genealogies. In: Donnelly P, Tavaré S, editors. Progress in population genetics and human evolution. New York, NY: Springer-Verlag. p. 271–285.

Hudson RR. 1983. Properties of a neutral allele model with intragenic recombination. Theor Popul Biol. 23:183–201.

Hudson RR. 1990. Gene genealogies and the coalescent process. In: Futuyma D, Antonovics J, editors. Oxford surveys in evolutionary biology. New York, NY: Oxford University Press. p. 1–44.

Hudson RR. 1993. The how and why of generating gene genealogies. In: Takahata N, Clark AG, editors. Mechanisms of molecular evolution. Sunderland, MA: Sinauer. p. 23–36.

Hudson RR. 2002. Generating samples under a Wright-Fisher neutral model of genetic variation. Bioinformatics. 18:337–338.

Hudson RR, Kaplan NL. 1994. Gene trees with background selection. In: Golding B, editor. Non-neutral evolution: theories and molecular data. London: Chapman & Hall. p. 140–153.

Hudson RR, Kaplan NL. 1995. Deleterious background selection with recombination. Genetics. 141:1605–1617.

Hudson RR, Kreitman M, Aguade M. 1987. A test of neutral molecular evolution based on nucleotide data. Genetics. 116:153–159.

Innan H. 2006. Modified Hudson-Kreitman-Aguade test and two-dimensional evaluation of neutrality tests. Genetics. 173:1725–1733.

Jensen JD, Kim Y, DuMont VB, Aquadro CF, Bustamante CD. 2005. Distinguishing between selective sweeps and demography using DNA polymorphism data. Genetics. 170:1401–1410.

Karlin S, McGregor J. 1972. Addendum to a paper of W. Ewens. Theor Popul Biol. 3:113–116.

Kim Y, Nielsen R. 2004. Linkage disequilibrium as a signature of selective sweeps. Genetics. 167:1513–1524.

Kim Y, Stephan W. 2002. Detecting a local signature of genetic hitchhiking along a recombining chromosome. Genetics. 160:765–777.

Kimura M. 1969. The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. Genetics. 61:893–903.

Kimura M. 1980. Average time until fixation of a mutant allele in a finite population under continued mutation pressure: Studies by analytical, numerical, and pseudo-sampling methods. Proc Natl Acad Sci USA. 77:522–526.

Kimura M. 1983. The neutral theory of molecular evolution. Cambridge: Cambridge University Press.

Li H, Stephan W. 2006. Inferring the demographic history and rate of adaptive substitution in. *Drosophila*. PLoS Genet. 2:e166.

Maynard Smith J, Haigh J. 1974. The hitch-hiking effect of a favourable gene. Genet Res. 23:23–35.

Nielsen R. 2005. Molecular signatures of natural selection. Annu Rev Genet. 39:197–218.

Nielsen R, Williamson S, Kim Y, Hubisz MJ, Clark AG, Bustamante C. 2005. Genomic scans for selective sweeps using SNP data. Genome Res. 15:1566–1575.

Przeworski M. 2002. The signature of positive selection at randomly chosen loci. Genetics. 160:1179–1189.

Przeworski M, Charlesworth B, Wall JD. 1999. Genealogies and weak purifying selection. Mol Biol Evol. 16:246–252.

Przeworski M, Wall JD, Andolfatto P. 2001. Recombination and the frequency spectrum in *Drosophila melanogaster* and *Drosophila simulans*. Mol Biol Evol. 18:291–298.

Sabeti PC, Reich DE, Higgins JM, et al. 2002. Detecting recent positive selection in the human genome from haplotype structure. Nature. 419:832–837.

Sabeti PC, Schaffner SF, Fry B, Lohmueller J, Varilly P, Shamovsky O, Palma A, Mikkelsen TS, Altshuler D, Lander ES. 2006. Positive natural selection in the human lineage. Science. 312:1614–1620.

Simonsen KL, Churchill GA, Aquadro CF. 1995. Properties of statistical tests of neutrality for DNA polymorphism data. Genetics. 141:413–429.

Slatkin M. 1994. An exact test for neutrality based on the Ewens sampling distribution. Genet Res. 64:71–74.

Stephan W, Song YS, Langley CH. 2006. The hitchhiking effect on linkage disequilibrium between linked neutral loci. Genetics. 172:2647–2663.

Tajima F. 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics. 105:437–460.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 123:585–595.

Teshima KM, Coop G, Przeworski M. 2006. How reliable are empirical genomic scans for selective sweeps? Genome Res. 16:702–712.

Thornton K, Andolfatto P. 2006. Approximate Bayesian inference reveals evidence for a recent, severe bottleneck in a Netherlands population of *Drosophila melanogaster*. Genetics. 172:1607–1619.

Voight BF, Kudaravalli S, Wen X, Pritchard JK. 2006. A map of recent positive selection in the human genome. PLoS Biol. 4:e72.

Wall JD. 1999. Recombination and the power of statistical tests of neutrality. Genet Res. 74:65–79.

Wall JD. 2000. A comparison of estimators of the population recombination rate. Mol Biol Evol. 17:156–163.

Wall JD, Andolfatto P, Przeworski M. 2002. Testing models of selection and demography in Drosophila simulans. Genetics. 162:203–216.

Wall JD, Hudson RR. 2001. Coalescent simulations and statistical tests of neutrality. Mol Biol Evol. 18:1134–1135; author reply 1136–1138.

Wang ET, Kodama G, Baldi P, Moyzis RK. 2006. Global landscape of recent inferred Darwinian selection for Homo sapiens. Proc Natl Acad Sci USA. 103:135–140.

Watterson GA. 1975. On the number of segregating sites in genetical models without recombination. Theor Popul Biol. 7:256–276.

Watterson GA. 1978. The homozygosity test of neutrality. Genetics. 88:405–417.

Wright S. 1931. Evolution in Mendelian populations. Genetics. 16:97–159.

Zeng K, Fu Y.-X, Shi S, Wu C-I. 2006. Statistical tests for detecting positive selection by utilizing high-frequency variants. Genetics. 174:1431–1439.

Zeng K, Mano S, Shi S, Wu C-I. 2007. Comparisons of site- and haplotype-frequency methods for detecting positive selection. Mol Biol Evol. 24:1562–1574.