

## FILE S1

**ms commands for demographic models used in this manuscript****TRUE demographic model for Pop A, when  $N_{AA} = 0.1N_A$ :**

```
./ms 40 100000 -t 20.8 -r 20.8 52000 -I 3 0 0 40 0 0 -en 0 2 2 -en 0 3 0.2 -cs 5e-4 3 0.8 -ej 5e-4 3 2 -ej 5e-4 4 1 -en 0.0275 1 0.055 -
en 0.0375 1 1 -en 0.06 2 1 -ej 0.1 2 1
```

**TRUE demographic model for Pop AA, when  $N_{AA} = 0.1N_A$ :**

```
./ms 40 100000 -t 20.8 -r 20.8 52000 -I 3 0 0 40 0 0 -en 0 2 2 -en 0 3 0.2 -cs 5e-4 3 0.8 -ej 5e-4 3 2 -ej 5e-4 4 1 -en 0.0275 1 0.055 -
en 0.0375 1 1 -en 0.06 2 1 -ej 0.1 2 1
```

**EST demographic model for Pop A when  $N_{AA} = 0.1N_A$ :**

```
./ms 40 100000 -t 41.6 -r 41.6 52000 -eN 0.0275 0.5
```

**EST demographic model for Pop AA when  $N_{AA} = 0.1N_A$ :**

```
./ms 40 100000 -t 39.52 -r 39.52 52000 -eN 0.05 0.55
```

**TRUE demographic model for Pop A, when  $N_{AA} = N_A$ :**

```
./ms 40 100000 -t 20.8 -r 20.8 52000 -I 3 0 0 40 0 0 -en 0 2 2 -en 0 3 2 -cs 5e-4 3 0.8 -ej 5e-4 3 2 -ej 5e-4 4 1 -en 0.0275 1 0.055 -
en 0.0375 1 1 -en 0.06 2 1 -ej 0.1 2 1
```

**TRUE demographic model for Pop AA, when  $N_{AA} = N_A$ :**

```
./ms 40 100000 -t 20.8 -r 20.8 52000 -I 3 0 0 40 0 0 -en 0 2 2 -en 0 3 2 -cs 5e-4 3 0.8 -ej 5e-4 3 2 -ej 5e-4 4 1 -en 0.0275 1 0.055 -
en 0.0375 1 1 -en 0.06 2 1 -ej 0.1 2 1
```

**EST demographic model for Pop A when  $N_{AA} = N_A$ :**

```
./ms 40 100000 -t 41.6 -r 41.6 52000 -eN 0.0275 0.5
```

**EST demographic model for Pop AA when  $N_{AA} = N_A$ :**

```
./ms 40 100000 -t 45.76 -r 45.76 52000 -eN 0.03863636 0.45
```

**Growth model estimated from the NIEHS YRI population in Lohmueller *et al.* (2010):**

```
./ms 20 10000 -s $$ -r tbs $LENGTH_SEQ <$RHO_FILE -eN 0.09259575 0.5
```

**Growth model estimated from the NIEHS AA population in Lohmueller *et al.* (2010):**

```
./ms 20 10000 -s $$ -r tbs $LENGTH_SEQ <$RHO_FILE -eN 0.08276125 0.46,
```

where “\$\$” is the observed number of segregating sites for the locus, “\$LENGTH\_SEQ” is the number of bases sequenced for the locus, and “\$RHO\_FILE” is a file with the values of  $\rho$  to use.

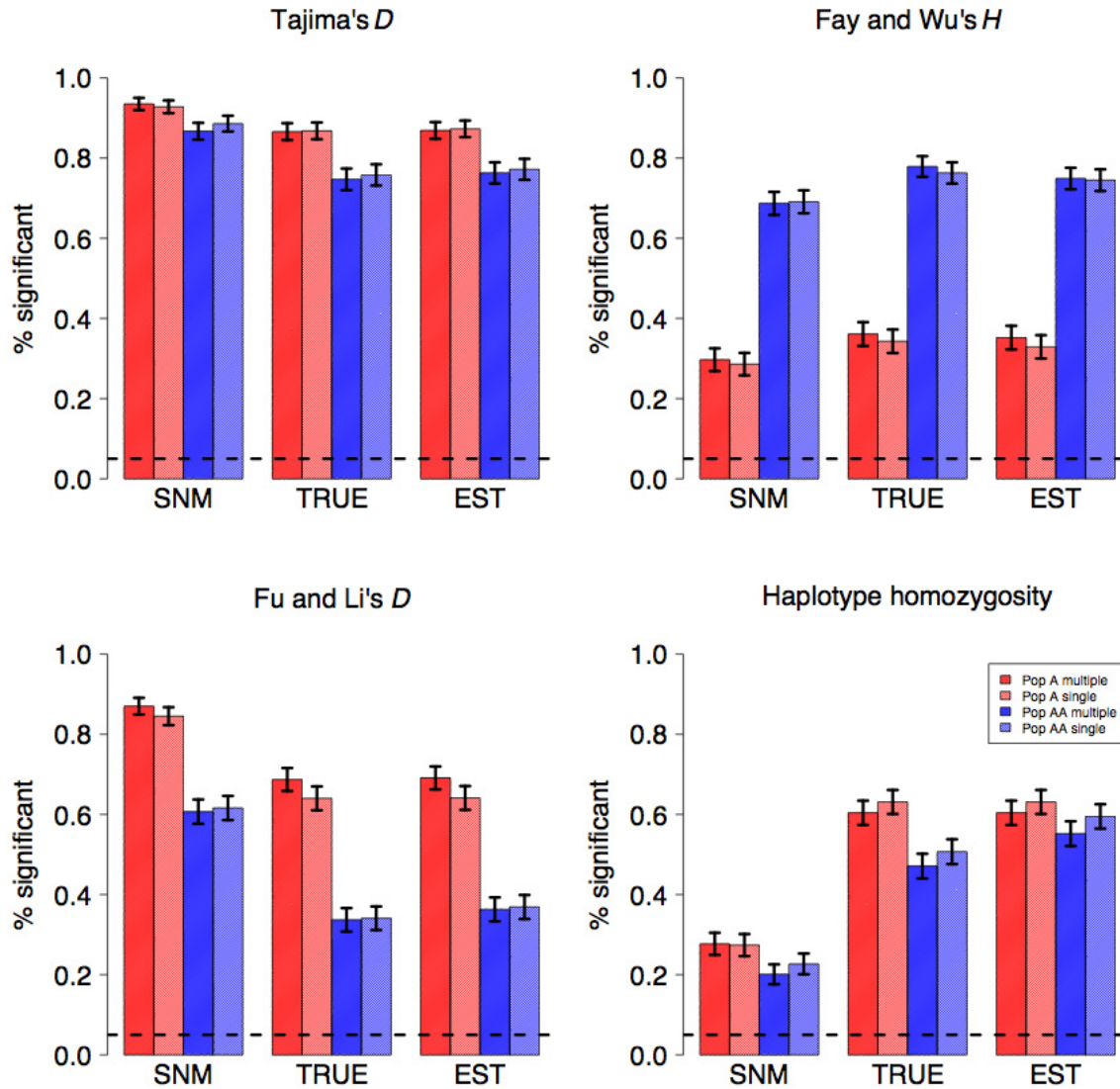


FIGURE S1.—Starting a single mutation at a given time (labeled “single”) and allowing multiple mutations to occur in a selected region for five generations (labeled “multiple”) show similar proportions of selected datasets rejecting neutrality. The fact that both simulation approaches yield similar results supports the validity of our approach. SNM denotes the rejection region defined by the standard neutral model, TRUE the rejection region defined by the true demographic model for each population, and EST the rejection region defined by a growth model where the parameters were estimated from the SFS of neutral data (see

Methods). Error bars denote approximate 95% CIS on the binomial proportion and were calculated from  $\hat{p} \pm 1.96 \sqrt{\frac{\hat{p}(1-\hat{p})}{N}}$ , where  $\hat{p}$  is the proportion of simulation replicates rejecting neutrality, and  $N$  is the total number of simulation replicates (1000 in our case).

**TABLE S1**  
**Demographic model parameters used for simulated datasets**

| Parameter   | Description  | Value                |
|-------------|--|----------------------|
| $N_B$       | Ancestral human population size  | 10,000               |
| $t_{split}$ | Time of the split between Pop A and Pop E (African and European populations) | 4000 generations ago |
| $N_A$       | Current size of Pop A (African population)                                   | 20,000               |
| $t_{cur}$   | Time Pop A expanded  | 2400 generations ago |
| $N_{AA}$    | Current size of Pop AA (African American population)                         | 20,000 or 2,000      |
| $N_{mid}$   | Size of Pop E (European population) during the bottleneck                    | 550                  |
| $t_{mid}$   | Duration of the bottleneck in Pop E (European population)                    | 400 generations      |
| $t_{curE}$  | Time that Pop E recovered from the bottleneck                                | 1100 generations ago |
| $p_{admix}$ | Proportion of European ancestry in Pop AA (African American population)      | 20%                  |
| $t_{admix}$ | Time before present when Pop AA (African American) population was founded    | 20 generations ago   |

**TABLE S2**

**Proportion of SNP segregating at different frequencies in simulated datasets as a function of the time the selected mutation arose ( $t_{sel}$ )<sup>a</sup>.**

| Frequency <sup>b</sup> | $t_{sel} = 2400$ generations |        | $t_{sel} = 3200$ generations |        | $t_{sel} = 4000$ generations |        |
|------------------------|------------------------------|--------|------------------------------|--------|------------------------------|--------|
|                        | Pop A                        | Pop AA | Pop A                        | Pop AA | Pop A                        | Pop AA |
| 0-50%                  | 0.029                        | 0.037  | 0.012                        | 0.012  | 0.015                        | 0.014  |
| >50-80%                | 0.029                        | 0.206  | 0.007                        | 0.193  | 0.008                        | 0.189  |
| >80-97.5%              | 0.219                        | 0.745  | 0.014                        | 0.785  | 0.006                        | 0.784  |
| Fixed                  | 0.723                        | 0.012  | 0.968                        | 0.010  | 0.972                        | 0.013  |

a. For these simulations,  $N_{AA} = 0.1N_A$  and  $\gamma = 100$  when  $t_{sel} = 3200$  and  $t_{sel} = 4000$ . When the selected mutations occur at the population expansion ( $t_{sel} = 2400$  generations),  $\gamma = 200$  to account for the larger population size while keeping  $s$  the same as before.

b. Frequency bins in a sample size of 40 chromosomes.