# Influence of Dominance and Drift on Lethal Mutations in Human Populations

D. Waxman[1] and A. D. J. Overall[2]

[1]Centre for Computational Systems Biology, ISTBI, Fudan University, 220 Handan Road, Shanghai 200433, PRC

[2]School of Pharmacy & Biomolecular Sciences, Huxley Building, University of Brighton, Brighton, East Sussex, BN2 4GJ, UK

This file provides some of the details underlying the simulations used to obtain the figure in the Supplementary Material.

To simulate the Wright-Fisher model for the conventional case of weak selection we use Eq. (6) of the main text, which we reproduce here as

$$X' = \frac{\mathrm{Bin}(2N, X + F(X))}{2N} \qquad \text{Wright-Fisher model for weak selection.} \tag{1}$$

For the case of a lethal allele we use Eq. (7) of the main text, which we reproduce here as

$$X' = \frac{\mathrm{Bin}(N, 2X + 2F(X))}{2N} \qquad \text{Wright-Fisher model for a lethal genotype.} \tag{2}$$

Here, we just give details for simulation for the conventional case, by directly using Eq. (1) above.

Schematically we proceed as follows.

1. Start with a given value of $X$.

2. Using this value of $X$, calculate $X + F(X)$.

3. Generate a single binomial random number with parameters $2N$ and $X + F(X)$ which we write as $B$

   ($2N$ represents the number of independent trials and $X + F(X)$ represents the probability of success on each trial).

4. Divide the binomial random number, $B$, by $2N$, to obtain $B/(2N)$ which is a relative frequency. This lies in the range 0 to 1.

5. Assign $X$ to be $R/(2N)$

6. Store the value of $X$ and go back to step 1, using the value of $X$ just obtained, and repeat the procedure.

In this way a set of $X$'s is obtained:

$$(X_0, X_1, X_{2,} ....).$$

This is a frequency trajectory.

Repeating the entire process again, with the same initial frequency, will generally yield a different trajectory - a replicate trajectory. Statistics of the frequency can be obtained by averaging over replicate trajectories. For example, the mean frequency at generation 2 is estimated by averaging $X_2$ over a large number of replicate trajectories.

In the file FigureS1.m the implementation of the above in Matlab, for a single replicate trajectory, covering $T$ generations, is given by

```
for t=1:T
    x=X(t);
    F=((1-h)*u-(h+(2-3*h)*u)*x-(1-2*h)*(1-u)*x^2)/((1+(1-2*h)*u)
        +(1-2*h)*(1-u)*x);
    X(t+1)=binornd(2*N,x+F)/2/N;
end
```