Can background selection explain low Y diversity in *R. hastatulus*?

We have done similations of background selection on the Y chromosome to test whether our observed level of Y/A diversity, $\theta_Y/\theta_A = 0.02$, is consistent with the effects of background selection on this chromosome.

Given our simulation parameters, what does the theory predict?

Here are the input parameters:

- $Ne_Y = 1/4Ne_A$, where the inital $Ne_A = 1000$
- $\mu = 0.005$ (neutral mutation rate, pre gamete, per generation)
- $U = 0.02$ (deleterious mutation rate, per *diploid*; so, twice the per gamete rate)
- $r = 0$ (recombination rate)
- $h = 1$ (dominance)
- $s = -0.01$ (mean s; sampled from a gamma dist)

Under the BGS model, the expected π is $E[\pi] = \pi_0 e^{-\frac{U}{2sh+r}}$, U is the mutation rate to strongly-deleterious alleles, π_0 is the value expected in the absence of BGS (*i.e.* $\pi_0 = \theta = 4N_e\mu$), s and h are the selection and dominance coefficients, and r is the recombination rate.

For our parameters, we have: $\pi_0 = \theta = 4N_e\mu = 5$, so

$$E[\pi_Y] = 5e^{-\frac{0.02}{0.1}} = 1.839397$$

For the autosomes, with $r = 0.5$, we have

$$E[\pi_A] = 20e^{-\frac{0.02}{0.1+0.5}} = 19.24537$$

and for the X chromosome

$$E[\pi_X] = 15e^{-\frac{0.02}{0.1+0.5}} = 14.43403$$

Expected Y/A and X/A diversity

OK, so then the expected ratios of diversity under background selection are:

$$\frac{E[\pi]_Y}{E[\pi]_A} = \frac{1.839397}{19.24537} = 0.095$$

and

$$\frac{E[\pi]_X}{E[\pi]_A} = \frac{14.43403}{19.24537} = 0.75$$

Simulated Y/A and X/A diversity

```
#autosomes
a <- read.csv("/Users/josh/Dropbox/temp/Ydiversity/sims/a_fixed.csv", header = T)
#X-chromosome
x <- read.csv("/Users/josh/Dropbox/temp/Ydiversity/sims/x_fixed.csv", header = T)
#Y-chromsoome
y <- read.csv("/Users/josh/Dropbox/temp/Ydiversity/sims/y_fixed.csv", header = T)
```

Simple box plot of simulated π_Y, π_X, and π_A

```
boxplot(a$thetapi, x$thetapi, y$thetapi, names = c("A","X","Y"),
        ylab = expression(paste("Nucleotide Diversity (", pi, " ) ")))
```
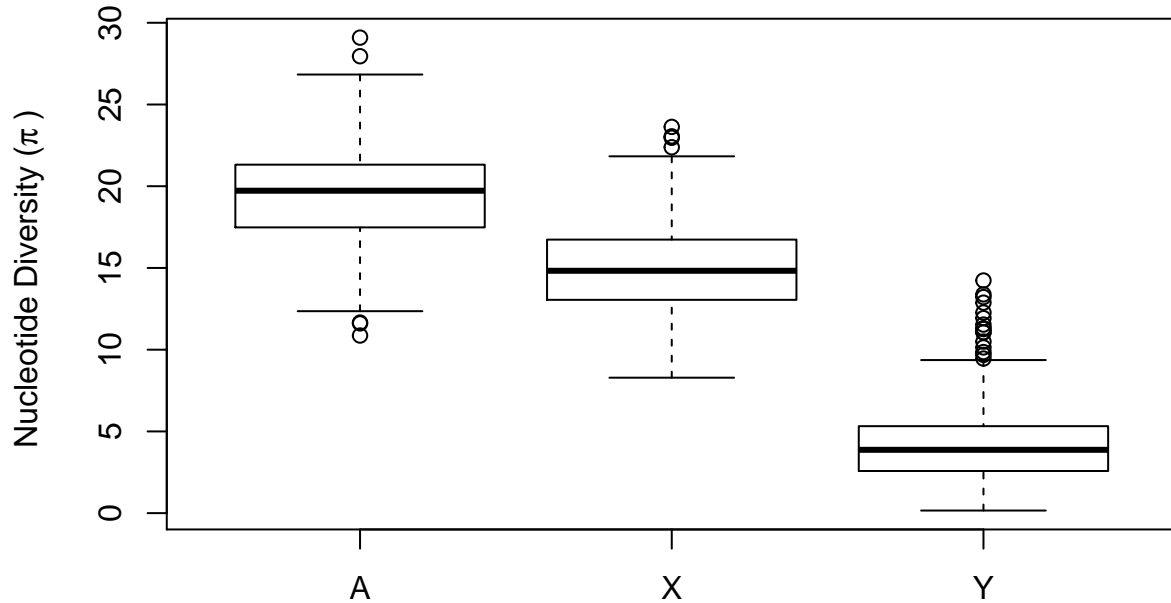

The following π values are the means from all of the simulations, for each chromosome type. You can see for the autosomes and the X chromosome, they are quite close to the theoretically expectd values calculated above.

$$\pi_{A_{sim}} = 19.5$$

$$\pi_{X_{sim}} = 14.9$$

$$\pi_{Y_{sim}} = 4.2$$

However, the estimate for the Y chromosome is much higher than predicted the by model. Moreover, ratios for the simulated data:

$$X/A = \frac{\pi_{X_{sim}}}{\pi_{A_{sim}}} = 0.76$$

and

$$Y/A = \frac{\pi_{Y_{sim}}}{\pi_{A_{sim}}} = 0.21$$

show the same effect: the X/A diversity ratio is not very different from our expectations, but the Y is much higher than expected.

Here is the code I used in fwdpy. And I did 1000 replicates for the 24 populations.

```
nregions = [fp.Region(beg=0,end=1,weight=1)]
```

```
sregions = [fp.GammaS(1,2,1,-0.01,0.1,1)]
```

```
recregions = [fp.Region(beg=-1,end=2,weight=1)]
```

```
N=250
pops = fp.evolve_regions(rng,24,N,nlist[0:],0.005, 0.01,0,nregions,sregions,recregions)
```