Equilibrium and Extinction In a Trisexual Diploid Mating System

Erik Buehler[1], Sanjoy Das[1], Jack F. Cully[2]

[1]Electrical and Computer Engineering
Kansas State University
Manhattan, KS 66506

[2]Division of Biology
Kansas State University
Manhattan, KS 66506

Abstract. In order to study the dynamics of a three-sex (trisexual) mating system, we extend the heterogametic sex-determining mechanism, used in many species, to include three sexes: XX, XY and YY. In this model, non-like types may mate, but like-types may not mate. We compare the dynamics of this system to a Mendelian system under Hardy-Weinberg conditions, and coin the term Trisexual Equilibrium to describe a system state very similar to Hardy-Weinberg Equilibrium. We construct computer simulations and mathematical models in an attempt to quantify the system's dynamics, and conclude that three-sex systems are not stable over time; they are destined to converge to two-sex systems. This conclusion is based on the fact that the less-represented homozygote's frequency variance (between adjacent generations) is positively linearly proportional to the respective frequency its self.

1 Introduction

> *"No practical biologist interested in sexual reproduction would be led to work out the detailed consequences experienced by organisms having three or more sexes; yet what else should he do if he wishes to understand why the sexes are, in fact, always two"* [1]

Our motivation for this investigation stems from the realization that in some species, such as most fish, XX is female, and XY is male [2]. Under certain conditions, a YY individual may be produced, and in the case of fish, this usually develops into a male. However, for some amphibians, XX is male, XY is female, and YY is also female. It should be noted that in many of these species, sex is not purely chromosomally determined. Environmental pressures can force what would normally develop into an XX male, to develop into an XX female, but environmentally-influenced sex determination is beyond the scope of this paper. We are specifically interested in the emergent behavior of a hypothetical system where the YY type is not just another male or female, but is its own distinct sex, or "mating type". In such a system, the terms "male" and "female" do not apply, so we will refer to the sexes by their notations XX, XY and YY. The XY sex is also referred to as the herterozygote, while the XX and YY sexes are sometimes referred to as the homozygotes. In this system, XX can only mate with XY and YY, XY can only mate with XX and YY, and YY can only mate with XX and XY. We are interested in discovering the emergent characteristics of this “diploid trisexual mating system” (DTMS), in an attempt to address why such a system is not observed in nature today. Yeasts and fungi are known to have multiple mating types (sometimes numbering in the thousands), but the mechanics of these sex-determining systems are markedly different from the herterogametic system we are interested in studying [3].

A few authors have attempted to tackle similar problems qualitatively, but their models of multi-sex systems tend to be overly complicated, and are not conducive to reduction [4,5]. To investigate the stability of this three-sex system over time, we must isolate the system itself by using a perfectly flat fitness landscape. This prevents natural selection from acting on the population. In addition, we maintain the system under what we will refer to as "Hardy-Weinberg conditions" [6]:

1) Each individual only lives for one generation.
2) Each individual attempts to mate with exactly one random partner, during each generation.
3) The individuals produce offspring at random, with equal probability.
4) There is no migration out of or into the population.
5) There is no mutation.
6) There is no selection pressure other than that imposed by the mating system itself.

For true Hardy-Weinberg Equilibrium to exist in a Mendelian system, one final assumption needs to be made about the population: its size must be infinite. In this case, all genotypes of subsequent generations have the same frequencies as the current generation. However, because we are interested in real world conditions, the models this investigation develops assume finite populations.

For the rest of this paper, we will refer to a Mendelian system as one dealing with two alleles A and a, which can occupy one loci for a pair of chromosomes. Possible pairings are aa, aA and AA. The frequencies of these genotypes in the population are F_{aa}, F_{aA} and F_{AA}, and their populations are P_{aa}, P_{aA} and P_{AA}. Note that the major difference between the Mendelian system and the trisexual system is that the Mendelian system allows any individual (such as aA) to mate with any other individual (such as aA, aa, or AA).

2 DTMS Simulation

We constructed a computer model (A1) to simulate a DTMS under Hardy-Weinberg conditions.

2.1 Algorithm A1: DTMS Simulation

```
POPULATION_SIZE = 1000
MAXIMUM_GENERATION = 1000

Initialize POPULATION to be a random composition of XX, XY and
```

```
YY

ITERATION = 0
While ITERATION < MAXIMUM_GENERATION
    Choose PARENT_PAIRS at random
    Loop Until NEW_POPULATION is full
        For each PARENT_PAIR, if they are not the same sex
            Mate them and place their offspring in NEW_POPULATION
            Break out of this For-Loop if NEW_POPULATION is full
        EndFor
    EndLoop
    POPULATION = NEW_POPULATION
    ITERATION = ITERATION + 1
EndWhile
```

After much experimentation, we found that on average, one of the homozygotes (XX or YY) tends to go extinct after roughly N/2 generations (N being the population size). Also, the population seemed to hover near an equilibrium of some sort, but not quite the Hardy-Weinberg equilibrium [6]. However, we are still left with the task of answering, "What is the mechanism causing these systems to converge to two-sex systems?" We first must quantify a DTMS more completely in order to determine the mathematical definition of the Trisexual Equilibrium we have observed in our simulations.

3 Computing Trisexual Equilibrium Points

Three mating types exist in the population: XX, XY and YY, each having the following respective frequencies: F_{xx}, F_{xy} and F_{yy}. Of course, Equation E1 must be satisfied at all times:

$$F_{xx} + F_{xy} + F_{yy} = 1 \tag{E1}$$

Knowing the mating type frequencies for the current generation, the expected frequencies for the subsequent generation can be predicted.

$$F_{xx}' = F_{xx}F_{xy} \tag{E2}$$
$$F_{xy}' = F_{xx}F_{xy} + 2F_{xx}F_{yy} + F_{xy}F_{yy} \tag{E3}$$
$$F_{yy}' = F_{yy}F_{xy} \tag{E4}$$

For reference, the equivalent next-generation-frequency equations for a system subject to Hardy-Weinberg Equilibrium are: [Stern, 1943]

$$F_{aa}' = F_{aa}^2 + F_{aa}F_{aA} + \tfrac{1}{4}F_{aA}^2 \tag{E2-H}$$
$$F_{aA}' = F_{aa}F_{aA} + \tfrac{1}{2}F_{aA}^2 + 2F_{aa}F_{AA} + F_{aA}F_{AA} \tag{E3-H}$$
$$F_{AA}' = F_{aA}F_{AA} + \tfrac{1}{4}F_{aA}^2 + F_{AA}^2 \tag{E4-H}$$

F_{aa}', F_{aA}' and F_{AA}' are inherently normalized, and do not need to be corrected by rescaling. However F_{xx}' F_{xy}' and F_{yy}' are *not* normalized, such that their sum is equal to 1. Our model requires a population of fixed size from generation to generation, so these values must be normalized. Frequencies with a subscript of "2" refer to the normalized, next-generation frequencies:

$$F_{xx2} = F_{xx}'/(F_{xx}' + F_{xy}' + F_{yy}') \tag{E5}$$

$$F_{xy2} = F_{xx}'/(F_{xx}' + F_{xy}' + F_{yy}') \qquad \text{(E6)}$$
$$F_{yy2} = F_{xx}'/(F_{xx}' + F_{xy}' + F_{yy}') \qquad \text{(E7)}$$

A Mendelian system under Hardy-Weinberg Equilibrium conditions, when perturbed away from equilibrium, will always reestablish equilibrium in a single generation [Stern, 1943]. For a DTMS, this is not the case. Convergence to ~99% of equilibrium is rather quick, on the order of fiver generations, but only truly converges as the number of generations approaches infinity. This makes it very difficult to compute the final equilibrium state of the system, given a starting state. However, one may note that this is a constrained system. In fact, under equilibrium, given one frequency, the other *two* frequencies are fixed quantities, and should also be computable. Substituting E2-4 into E5:

$$F_{xx} == \frac{F_{xx}F_{xy}}{F_{xx}F_{xy} + F_{xx}F_{xy} + 2F_{xx}F_{yy} + F_{xy}F_{yy} + F_{yy}F_{xy}} \qquad \text{(E8)}$$

$$2 == \frac{F_{xy}}{F_{xx}F_{xy} + F_{xx}F_{yy} + F_{xy}F_{yy}} \qquad \text{(E9)}$$

$$1 == F_{aa}^2 + 2F_{aa}F_{aA} + F_{aA}^2 + 2F_{aa}F_{AA} + 2F_{aA}F_{AA} + F_{AA}^2 \qquad \text{(E9-H)}$$

Therefore, if our “litmus test” equality (E9) is satisfied, then the system is in Trisexual Equilibrium, and all subsequent generations will have the same sex frequencies as the current generation (for a sufficiently large population). We derived the equivalent Hardy-Weinberg test, E9-H, for comparison. In a further continuation of the above development, we can also compute Trisexual Equilibrium points directly. For notational brevity, Let $a = F_{xx}$, $b = F_{xy}$, $c = F_{yy}$, and x = the unknown frequency.

$$2 == \frac{x}{ax + a(1 - a - x) + x(1 - a - x)} \qquad \text{(E10)}$$

$$0 == (-1)x^2 + (\tfrac{1}{2} - a)x + (a - a^2) \qquad \text{(E11)}$$

$$F_{xy} = \frac{1 - 2F_{xx} + (-12F_{xx}^2 + 12F_{xx} + 1)^{\frac{1}{2}}}{4} \qquad \text{(E12)}$$

$$F_{yy} = \frac{3 - 2F_{xx} - (-12F_{xx}^2 + 12F_{xx} + 1)^{\frac{1}{2}}}{4} \qquad \text{Where } 0 \leq F_{xx} \leq \tfrac{1}{2} \qquad \text{(E13)}$$

For comparison, we derived the equivalent equations to E12 and E13, for a Mendelian system under Hardy-Weinberg Equilibrium:

$$F_{aA} = 2F_{aa}^{\frac{1}{2}} - 2F_{aa} \qquad \text{(E12-H)}$$
$$F_{AA} = 1 - (2F_{aa}^{\frac{1}{2}} - F_{aa}) \qquad \text{(E13-H)}$$

It will be useful to us later if we take a moment to compute the bounds for F_{xy}, under Trisexual Equilibrium conditions. First, when $F_{xx} = 0$ and $F_{xx} = \frac{1}{2}$, then $F_{xy} = \frac{1}{2}$.

However, we will need to compute the zero of F_{xy}'s derivative with respect to F_{xx} in order to find the maximum value realized by F_{xy} under Trisexual Equilibrium, which happens to occur at the system's balance point ($F_{xx} = F_{yy}$):

$$\frac{dF_{xy}}{dF_{xx}} = \frac{3 - 6F_{xx} - (-12F_{xx}^2 + 12F_{xx} + 1)^{1/2}}{(-12F_{xx}^2 + 12F_{xx} + 1)^{1/2}} = 0$$

$$F_{xx} = (3 - 3^{1/2})/6 \sim 0.2113 \text{ @ } F_{xy}\text{'s maximum}$$

Plugging this into for F_{xx} in E13 yields: $\max(F_{xy}) = 3^{1/2}/3 \sim 0.5774$, so our bounds for F_{xy} are: $1/2 = F_{xy} = 3^{1/2}/3$. Similarly, for F_{xy} known, find F_{xx}:

$$F_{xx} \text{ or } F_{yy} = \frac{1 - F_{xy} + (1/F_{xy}^2 - 3)^{1/2}}{2} \qquad \text{(E14)}$$

$$F_{yy} \text{ or } F_{xx} = \frac{1 - F_{xy} - (1/F_{xy}^2 - 3)^{1/2}}{2} \quad \text{Where } 1/2 \leq F_{xy} \leq 3^{1/2}/3 \qquad \text{(E15)}$$

Note that equations E14 and E15 are ambiguous as to which frequency they refer. This is due to the symmetry of the system. If an arbitrary value is chosen for F_{xy}, there are two possible values for F_{xx}, and the same two possible values for F_{yy}. Of course, E16 and E17 are analogous to E12 and E13:

$$F_{xy} = \frac{1 - 2F_{yy} + (-12F_{yy}^2 + 12F_{yy} + 1)^{1/2}}{4} \qquad \text{(E16)}$$

$$F_{yy} = \frac{3 - 2F_{yy} - (-12F_{yy}^2 + 12F_{yy} + 1)^{1/2}}{4} \quad \text{Where } 0 = F_{yy} = 1/2 \qquad \text{(E17)}$$

Figures 1 and 2 illustrate the nature of the curves (E12 - E17) graphically. Figure 2 also includes the curves for a Mendelian system under Hardy-Weinberg Equilibrium (computed using E12-H and E13-H).

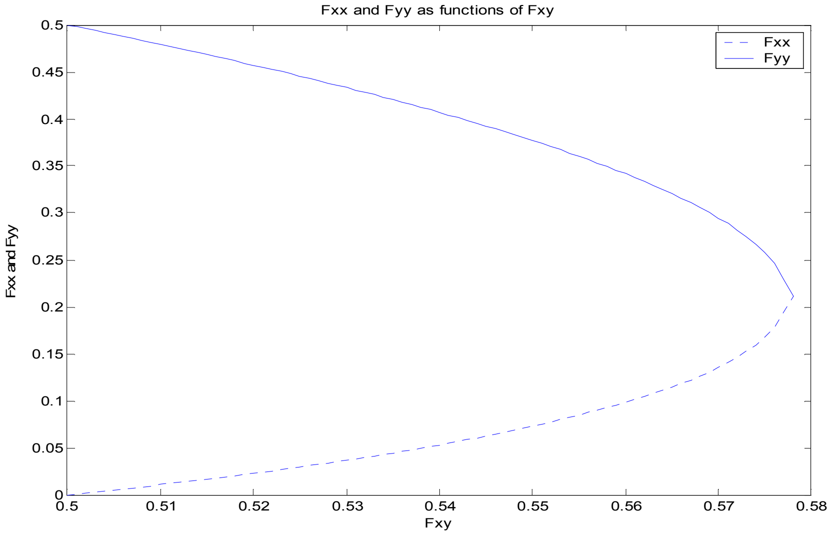

Fig. 1. F_{xx} and F_{yy} plotted as functions of F_{xy}, under Trisexual Equilibrium conditions

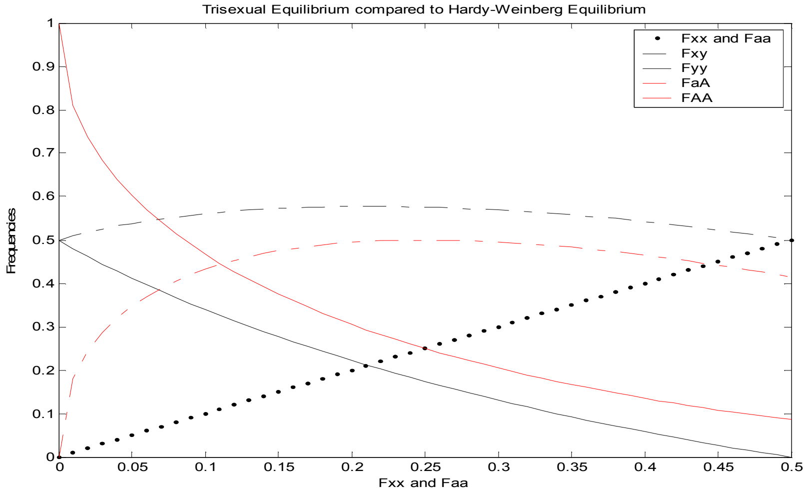

Fig. 2. F_{xx}, F_{xy} and F_{yy} plotted as functions of F_{xx}, under Trisexual Equilibrium conditions, compared to the equivalent frequencies under Hardy-Weinberg Equilibrium

As a quick aside, it is easy to show that the mating efficiency of the trisexual system is ~ 57.7%, significantly more efficient than that of an equivalent two-sex system (50%). Note that as the three-sex system drifts away from the balance point, towards the extinction of one of the homozygotes, the mating efficiency decreases towards 50%. For a trisexual system to have a significant advantage over a two-sex system, a mechanism is needed to push the system back towards the balance point if it drifts away. From our experimentation thus far, it seems that such a mechanism does not exist in the system, and the rest of this paper is dedicated to showing just that.

4 Generation-to-generation bias

Our end goal is to discover, at the ground level, what is the fundamental pressure that forces a DTMS to converge to a two-sex system. One possible mechanism could be that for the less-represented heterozygote, a slight bias causes a higher probability of its frequency to decrease in the next generation, and for the more-represented heterozygote, an opposite bias might exist pushing it towards fixation. To gather evidence to this effect, we have modified with the model used in Algorithm A1 to see if, when started at an arbitrary Trisexual Equilibrium point, there is any such bias in the system. It begins with an F_{xx} of 0, computes the equilibrium F_{xy} and F_{yy}, then constructs the population based on these frequencies. It then allows mating to occur to construct the new population. Then F_{xx} is incremented, and the whole process is repeated until F_{xx} is equal to the maximum allowed value. The goal is to construct a dataset, from which the generation-to-generation statistics can be measured. Upon analyzing the mean generation-to-generation changes, no biases were observed, weighting for or against the less-represented homozygotes.

To illustrate how strongly the DTMS system hovers near Trisexual Equilibrium, Figure 3 shows all three frequencies plotted as the system converges to two sexes (using Algorithm A1). The corresponding equilibrium frequencies are overlaid as dark black lines. The equilibrium frequencies were computed based on the actual F_{xx} for that generation, using equations E12 and E13. This validates our method for computing generation-to-generation statistics by repetitively starting the system at an equilibrium point and observing the change in the next generation.

5 Lone remaining X homozygote

To further explore this idea, we now examine the simplified situation where the less-represented mating type has a frequency of 1/POP_SIZE. In other words, regardless of the population size, only one individual remains of that particular type. We are only referring to the homozygotes in this case, since it should be intuitively obvious that even if no XY members remain in the population, they will immediately become reconstituted by pairings between the XX and YY members. However, if only one XX member remains in a sea of XY and YY members, there is a finite chance that in the next generation, none of its offspring will carry the XX type, and from then on, no XX individuals will arise from subsequent pairings of XY and YY individuals. What, then, is the probability of this extinction, with a sole remaining XX individual?

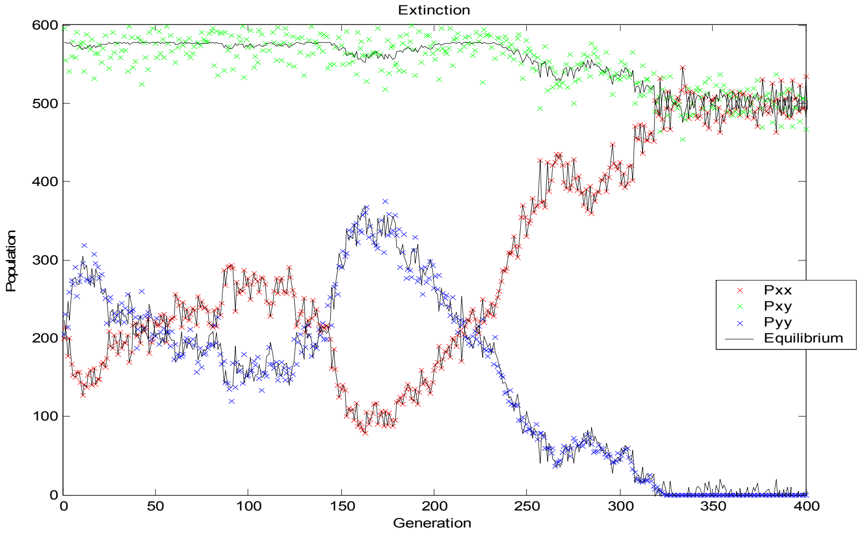

Fig. 3. DTMS hovering near Trisexual Equilibrium, Converging

In such a setup, E12 and E13 tell us that F_{xy} and F_{yy} approach ½ and ½ as F_{xx} approaches 0. Because XX is the only mating type with which XX is incompatible, the lower its frequency, the higher its probability of mating success (as it probably will not be paired with an incompatible type). Therefore, the probability of the last XX being paired with a compatible mate approaches 1, but in order for XX to have a chance of being represented in the subsequent generation, it must be paired with XY, not YY. The probability of the last XX being paired with an XY approaches ½.

There will also be a pairing inefficiency induced by the high concentration of YY and XY mating types in the population. Only ½ of YY's parings will be with XY, resulting in successful mating, and only ½ of XY's pairings will be with YY. The implications of this mean that each successful pairing must produce four children in order for the subsequent population to be of the same size. Therefore, the expected children from a pairing of an XY with an XX are two XXs and two XYs. However, because this pairing can only be expected to occur ½ of the time, the expected number of XX in the subsequent generation becomes 1, and the probability of extinction is therefore ½.

This result also shows no hint of a pressure pushing the lesser homozygote towards, or away from, extinction. For large populations, it seems to be neutral with respect to mating type frequencies. This might seem proof enough that nothing is pushing the system towards the extinction of the less-frequent mating type, but we know from experience that these systems do converge to two-sex systems.

Perhaps a less obvious way of approaching this problem involves the generation-to-generation frequency variance. What we have not shown yet is how easily a loss in frequency can be recovered. For instance, in a scenario where F_{xx} (0.10) decreases by 0.01, with ½ probability, and increases by 0.01 with the same probability, in the next iteration it may have fallen to 0.09. In the following generation, it would again have the same probability of increasing or decreasing, but due to the nature of the probability distribution, it may have a smaller variance, and therefore is likely to change by a value less than the original change of 0.01. So even if it increases in the following generation, it may not completely recoup its losses. This can be thought of as a "two steps back, one step forward" mechanism. We are merely hypothesizing at this point, but a mechanism like this could prove to be the one which causes the eventual extinction of a third sex in a DTMS.

Figure 4 verifies our suspicion of a variance that is proportional to frequency. A conclusion can immediately be drawn from Figure 4: The frequency variance can be very closely approximated by ½•Frequency•PopulationSize for a homozygote. Note the striking resemblance between Figure 4 and Figure 2 (for the homozygotes).

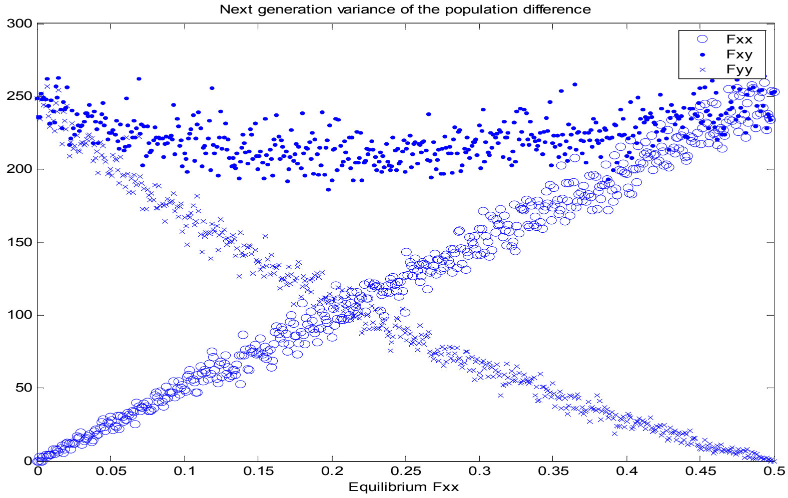

Fig. 4. Frequency change variance from one generation to the next

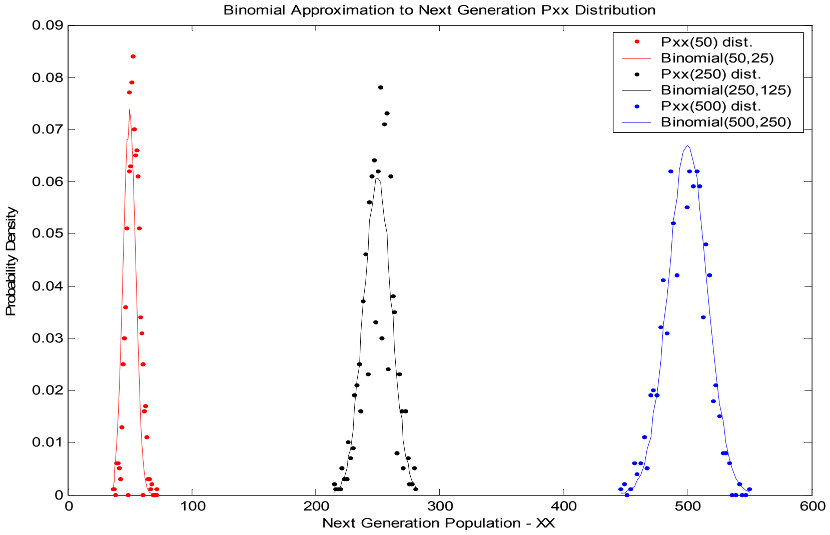

Fig. 5. - Binomial distributions compared to the experimental next-generation change distributions

We hypothesized that the binomial distribution could pose as a reasonable model for the generation-to-generation homozygote frequency probability distribution. Using it, with the knowledge that the homozygote's generation-to-generation variance seemed to be equal to ½•Frequency•PopulationSize, we plotted the actual observed generation-to-generation frequency change distributions (with a bin size of 40) against their respective binomial distributions. The results can be seen in Figure 5, and a pretty close match between the distributions is observed.

All of this certainly indicates that we are on to something, with respect to looking at the system's variance rather than its average generation-to-generation behavior, so to see if this phenomenon is actually capable of producing convergence, we constructed an extremely simple simulation, which merely forces a binomial distribution to use its last value as its mean, and ½ of its last value as its variance. This is outlined under Algorithm A3.

The simulation is started with a P_{xx} that is very close to the theoretical Trisexual Equilibrium midpoint, with an F_{xx} of 0.211, translating into a P_{xx} of 211 for a population of 1000. Its behavior is compared (graphically) to the simulation described in Algorithm A1, with a population of 1000. These comparisons are shown in Figure 6.

5.1 Algorithm A3 - Binomial Distribution With Variance-Feedback

```
POPULATION_SIZE = 1000
MAX_GENERATIONS = 1000

Pxx = 211
GENERATION = 0

While GENERATION < MAX_GENERATIONS
    Pxx = Binomial(Pxx,½Pxx)
    GENERATION = GENERATION + 1
EndFor
```

Note that Binomial(u,v) produces a random number, according to the binomial distribution, with mean u and variance v.

The plots in Figure 6 are just a few random samples of the many runs we performed, and are shown side by side simply to illustrate the strikingly similar behavior between the frequency of the lesser homozygote in Algorithm A1, and the emulated frequency in Algorithm A3. In fact, not only do the behaviors *look* very similar, but A3 even *converges* roughly as quickly as A1 (near N/2 iterations). You will note that no extinction is shown for YY, and this is simply because we chose to only show the plots where XX went extinct. YY and XX went extinct with equal probability, and in fact, Figure 3 shows an example of YY extinction.

The similarity in convergence behavior with our DTMS simulation's homozygote extinction behavior provides a strong piece of evidence supporting our hypothesis. Coupled with the striking similarity between the binomial distribution and the intra-generation change distribution, we are forced to conclude that variance-feedback is, indeed, the extinction-driving mechanism.

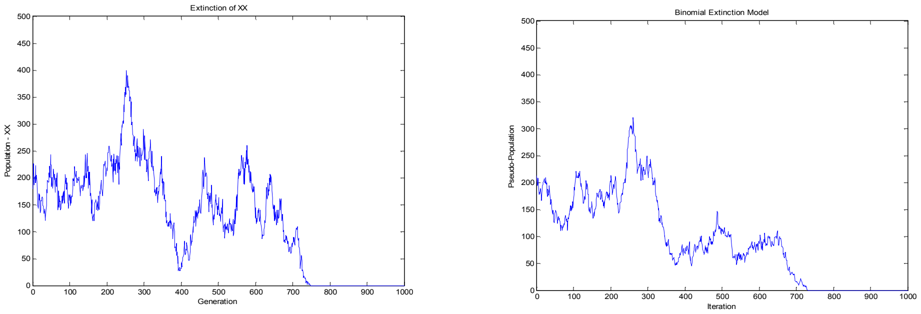

Fig. 6. Binomial Distribution with feedback compared to actual DTMS

6 Conclusion

We began this investigation asking whether or not a three-sex mating system would always converge to two sexes. We chose a diploid, chromosomally-determined mating system as a representative model for a possible three-sex mating system, and constructed a simulation to explore its dynamics. The simulation showed very strong evidence that such a system, with no outside pressures, would always converge to a two-sex system, and we were left to answer the question of "why?". We broke the problem down, first mathematically, by deriving the equations that define Trisexual Equilibrium. We then performed a simulation to determine if there was a higher probability of the lesser homozygote decreasing after each generation, and found no evidence of this. We looked at the statistics of a lone remaining homozygote in the population, and found its expected representation in the population to be 1, thus also showing no bias against lesser homozygotes in subsequent generations. Finally, we looked into the generation-to-generation frequency variance, and found that it is closely approximated by ½•Frequency•PopulationSize. This led us to hypothesize that the decreasing variance with decreasing frequency is the fundamental mechanism driving the extinction of the less-represented homozygote. We compared the binomial distribution to the frequency-change distribution, and found a striking similarity, so we used the binomial distribution in a feedback algorithm to determine if variance-feedback alone could cause a system to converge. All of our simulations showed that this, indeed, is a very reliable phenomenon, and we conclude that variance-feedback is the primary mechanism causing the extinction of one of the homozygotes in a diploid trisexual mating system.

References

1. Fisher, R.A., 1930: *The genetic theory of natural selection.* Oxford: Clarendon Press.
2. Bull, J.J., 1987: *Sex determining mechanisms: An evolutionary perspective*. The Evolution of Sex and Its Consequences. S.C. Stearns, 1987. 93-114. Birkhäuser Verlag Basel.
3. Kothe E. 1996: *Tetrapolar fungal mating types: sexes by the thousands.* FEMS Microbiology Review 18: 65-87.
4. Coker, P., Winter, C., 1997: *N-Sex reproduction in Dynamic Environments.* Proceedings of the Fourth European Conference on Artificial Life.
5. Jaffe, K., 1996: *The dynamics of the evolution of sex: Why the sexes are, in fact, always two?* Interciencia 21(6), 259-267.
6. Stern, C., 1943: *The Hardy-Weinberg Law.* Science 97, 137-138.