Computer Science
University of Manchester

The Evolution of Polygenic Sex Determination with Potential for Environmental Manipulation

M.J. Hatcher & C. Tofts

Technical Report UMCS-95-4-2

The Evolution of Polygenic Sex Determination with Potential for Environmental Manipulation*†

M.J. Hatcher

Department of Pure and Applied Biology
University of Leeds
Leeds
LS2 9JT
pab6mjh@gps.leeds.ac.uk

C. Tofts

Department of Computer Science
University of Manchester
Oxford Road, Manchester, UK.
M13 9PL
cmnt@cs.man.ac.uk

†M.J. Hatcher is supported by the NERC, C. Tofts is supported by an EPSRC Advanced Fellowship

Abstract

We present a formal account of a genetic mechanism for Environmental Sex Determination (ESD) using a probabilistic process algebra (WSCCS). We exploit the process algebra to demonstrate that a single chromosome pair based polyfactorial system is intrinsically unstable. The system evolves, with a deterministic relationship between an animal's sex and its polyfactor count, to three possible forms of genetic mechanism: pure ESD; pure genetic sex determination (GSD) (both XY and ZW being possible outcomes); or a mixed solution with a 50/50 division between ESD and GSD. With a probabilistic dependence between gene count and sex we observe the first and last of the above solutions (no pure GSD), but with population gene frequencies that can be far from Hardy-Weinberg equilibrium. A model of the systemic behaviour in the presence of recombination is presented, and the effect of the relative recombination and evolution rates is discussed. Our process algebraic model is extended to account for the polyfactors being distributed over many possible chromosome pairs and demonstrates that only one essentially new stable solution is reachable. We demonstrate that within a multichromosome system there will almost always be a 'leading' chromosome pair which determines sex, other chromosome pairs will have identical sex determining information upon both members of the pair, but **not** necessarily none.
Keywords: Process Algebra, Probability, Theory of Computing, Simulation, Evolution, Sex Determination, ESD, Polygenics.
CR categories: F.3.1, I.6, G.3.

1 Introduction

In this paper we attempt to provide a genetic basis for *Environmental Sex Determination* (ESD). Rather than assuming that genetic make up forces gender we shall assume that the underlying genotype of an animal gives a predisposition towards a particular sex, but that other factors may affect the individual's final choice of sex. To that end we will describe sex as being determined by a polygene whose intensity defines the prediposition towards a particular sex (male). Of particular interest to us are two separate populations of *Gammarus duebeni*, one of which shows conventional genetic sex determination and the other ESD [NAG88a, NAG88b, WA94].

The allocation of resources to each sex within a population is a major evolutionary pressure in sexual species ([Fis30, Wri69, Cha82, Bul83, KL86] and many others). In many diploid organisms, sex is determined with respect to a single pair of *sex chromosomes*, which may be structurally distinct from the autosomes (non-sex chromosomes) and from each other (termed heteromorphic *e.g.,* [Ohn67, Ohn79]). Where cytology has not identified heteromorphism, gender appears to be determined by one or more *major sex factors*, and in some cases by the actions of many *minor sex factors*, (the terms are relative; reviewed in [Bul83, Ohn79]). In this paper we develop models relating to major and minor factor control systems, which we term *multifactorial* or *polygenic*. Factors are distributed over a number of sites (multiple loci) on one or more chromosome pairs.We shall refer to the total sum of factors as a *polygene*, and to the factors collectively as polyfactors.

The dynamics of multilocus or polygenic factors for determining organism phenotype have been widely studied [Win34, Wri39, Kos64, Wri68, Wri69, KL86, Bul83, Poe88, TB90, Nag93] and the effect of using such a system for sex determination has been studied in [Win34, Kos64, Wri69, KL86, Poe88]. There are many examples of animals in which it is believed that gender is determined by some underlying multifactor system [Aid36, Gor54, Fre76, BVB82, CVE92]. Necessarily, many of these studies have concentrated on sex determined by genotype at conception, although for many organisms there may be a selective advantage to determination of sex with respect to environmental cues [CB77, WC78, Cha82, BC88]. This advantage may arise whenever the environment is heterogenous in time or space and the relationship between environment quality and adult reproductive success differs between the two sexes. Under these circumstances, selection favours those individuals which produce relatively more of the sex with the greater relative fitness for that patch [TW73, CHJA81]. When patch quality cannot reliably be predicted by the parent, but can be predicted from environmental cues experienced by the offspring, ESD is favoured [CB77]. Among others, ESD occurs in plants [LB84, Zim91]; nematodes [BlC89], crustaceans [NAG88a, AGN87]; fish [CK81, LC93]; and reptiles [Bul80, DF88, Jan92]. It is usually supposed that ESD has multifactorial genetic underpinnings as (1) adaptive interpretations of ESD imply selection, which requires that some aspects of ESD are *heritable*, and (2) it is difficult to envisage how ESD could be encoded under simple chromosomal sex determination [BVB82, CVE92].

Unlike the analysis of [BB82] we shall not assume that the polyfactors are normally distributed and then look for a single control point which induces stability. We shall analyse two forms of decision:

1. **deterministic system**: if the total factor count of the polygene is below a particular value then the animal will become female, if the loading is above a (possibly different) control value the animal will become male, otherwise it will decide its sex in an unbiased random fashion (50/50);

2. **nondeterministic system**: we assume that a particular polygenic load leads to a particular probability of becoming female, for simplicity this distribution will initally be linear.

For our purposes the random region represents that area in which the environment *can potentially* have an effect on the outcome of sex determination, although we do not model ESD explicitly.

The model methodology that we adopt is essentially a **formal** individual based techinique, *Process Algebra*. Process algebra [Mil90, Mil80, BW90, Hoa85, Hen88] is particularly appropriate to the modelling of evolutionary systems as it permits us to track precisely the genetic make up of *every* individual within a population. To express the behaviours of our system of organisms we need to model the effects of time, probability and priority. Hence, our process algebra will need to be able to express all of these behaviours. The *Weighted Synchronous Calculus of Communicating Systems* (WSCCS)[Tof94, Tof90a] is capable of expressing all of these properties and we shall use it to derive the properties of our evolutionary system. An application of these methodologies to the behaviour of social insects can be found in [Tof90b, Tof91, Tof93a, Tof93b] and their biological consequences in [THF92].

Evolutionary systems are increasingly of interest to computer scientists as the methods are often used for exploration of complex systems, via genetic algorithms [Dav87, Gol89, BB91, Dav91]. The behaviour of genetic algorithms is highly complex and providing appropriate formal analysis techniques for such computational methodologies is difficult. From a computing prespective an evolutionary system is a large scale parallel device which optimises its use of resources. The analysis of such natural parallel systems may help in the design and analysis of similar artificial computing systems.

In this initial study of polygenic sex determination we shall concentrate on the stability of the gene pool. That is to say we shall examine what the stable distribution of genes within the population will be after many generations. In particular we shall perform our analysis in the *absence* of any selective pressures. That is we are studying the structural stability of our genetic model of ESD. The analysis of polygenes for other behaviours [TB90] suggests that it is highly unlikely that the distribution of polyfactors will be Guassian as assumed by [BB82]. Thus it is of great importance to know what the underlying distribution of our factors should be in the absence of an enviromental bias giving advantage to particular individuals.

In the succeeding section we describe our basic process algebraic description of the polygenic system and analyse its stability given sex selection systems of both types stated above. This initial analysis is undertaken under the assumption that all of the factors contributing to a polygene are resident on a single chromosome pair. We consider the effect upon our stable states of recombination between the genes. In Section 5 we analyse the effect of distributing the polygene over many chromosome pairs.

The main observation is that under finite population assumptions the deterministic system evolves to one of three basic forms of stability:

1. monogenic purely random sex determination (all individuals genetically identical);

2. two factor pure genetic sex determination;

3. two factor mixed sex determination;

with each showing a 50/50 sex ratio as we should expect [Fis30] in our sexually unbiased system.

The non-deterministic system evolves into one of two solution types; either 1 or 3 of above. When the factors are distributed over many chromosomes then we see essentially the same behaviour, with one pair becoming 'leading' in sex determination and all of the other pairs being monogenic with respect to our polyfactor.

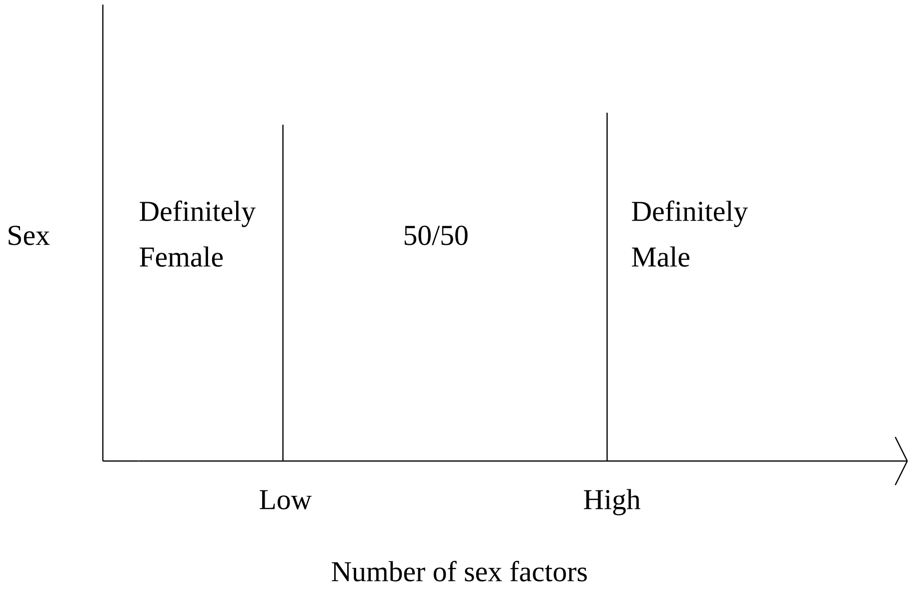

Figure 1: Sex allocation function with deterministic regions.

2 Sex Allocation with Deterministic Regions

We consider the case of stability for a polygenic sex determination system when the sex determining function has the following form: there is some lower limit of factors (Low) below which all animals are **female**; there is an upper limit of factors ($High$), above which all animals become **male**; in between these limits sex is allocated at random on a 50/50 basis. Pictorially such a selection function can be viewed as given in Figure 1. We are modelling a dioecious[1], diploid[2], multiallele[3], multilocus[4] system in which each allele[5] is of unitary value contributing additively[6] to the genotype.

A formal description of the behaviour of our shrimp would be the following collection of WSCCS processes.

The following process describes those animals which are dead:

$$Dead \stackrel{def}{=} \surd : Dead$$

Since we are using a synchronous description system the above process is the identity with respect to parallel composition, as a result adding any such processes to a population will have no effect on the system's behaviours.

[1] Two sexes

[2] Chromosomes occur in pairs

[3] There are many different versions of the gene for each character

[4] There are many places on the chromosomes at which the alleles may occur

[5] a genetic factor which may specify or contribute to the specification of the animal (phenotype)

[6] as opposed to more complicated relationship. In some senses these can be coded up in the response function from number to sex and so we do not need to consider more complicated allelic behaviour

A male process with genotype x, y, which represent the number of polyfactors on each of the two chromosomes, tries to export its genes to a female process and then dies,

$$Male(x,y) \stackrel{def}{=} \omega.\overline{mate}(x,y) : Dead + 1.\surd : Dead$$

Obviously the male animal will prefer to mate if it possibly can, and hence priority is given to achieving that end. Since we are modelling the system as having nonoverlapping generations the male will die unmated if it cannot mate at this point in time.

A female process takes in the genes specified by the male to produce a process with their combined genetic information:

$$Female(u,v) \stackrel{def}{=} \omega.mate(x,y) : Pair(x,y,u,v) + 1.\surd : Dead$$

This may seem an unnatural representation of the actual process of mating where the males genetic information is divided into sperm on a per chromosome basis and *then* passed onto the female. From our point of view as long as the division of genetic information occurs in an unbiased manner then the point in time at which it occurs is irrelevant, provided it occurs before the final allocation of sex to the young. In the above the factors u and v represent the number of polyfactors on the two chromosomes of the female. Again female animals prefer to mate rather than not doing so and if they cannot then they will not influence future generations.

The pair now produces some number of offspring which will later acquire their genetic type, so at this point we retain the information on all four parental chromosomes[7]:

$$Pair(x,y,u,v) \stackrel{def}{=} \sum_{i \in Offspring} off_i.\surd : Young(x,y,u,v)^i$$

In the above off_i represents the probability of having i offspring [8], and $Offspring$ the set of permissible values of number of young. It is simpler to model the production of a number of 'young' with the full parental genetic information and assign an appropriate division of that genetic information at a subsequent point.

The young now acquire their genotype in the correct proportions:

$$\begin{aligned} Young(x,y,u,v) \stackrel{def}{=} \; & 1.\surd : Animal(x,u) \\ & +1.\surd : Animal(x,v) \\ & +1.\surd : Animal(y,u) \\ & +1.\surd : Animal(y,v) \end{aligned}$$

This is the standard pattern of inheritance in a diploid system, following the laws of Mendelian inheritance. That is to say that the members of a homologous pair assort at random independently.

By splitting mating, offspring production and genetic assignment into separate stages we permit the introduction of genetic based biases in each of these life stages, and therefore could permit selection to occur upon any of them. Hence, the process theoretic

[7] in some sense we are performing meiosis on both parents in the young, but formally this is irrelevant

[8] The abbreviation $P^i \stackrel{def}{=} P \times P \times \ldots \times P$ i times.

framework we have provided can be exploited for more general problems of evolution without great alteration, although in this simple system there is no selection based upon genotype.

Now the animals have to eat in order to survive:

$$Animal(x,y) \stackrel{def}{=} \omega.eat : Ani(x,y) + 1.\surd : Dead$$

This stage ensures that the population of adults remains finite, limited by the availability of the food resource. Again we could model selective advantage of a particular genotype in foraging at this point but choose not to do so for simplicity in this analysis. Notice that as a result of the priority if there is a sufficient population to take up all of the available resource then this will happen. Alternatively we have bounded the population to the carrying capacity of its environment.

Finally, the animal's sex is determined:

$$Ani(x,y) \stackrel{def}{=} \begin{cases} \surd : Male(x,y) & \text{if } x+y \geq High \\ \surd : Female(x,y) & \text{if } x+y \leq Low \\ 1.\surd : Female(x,y) + 1.\surd : Male(x,y) & \text{otherwise} \end{cases}$$

which allocates animals a sex on the basis of their genotype given the decision procedure we outlined above.

A process to provide the animals with food:

$$Food \stackrel{def}{=} \textstyle\sum_{i\in[0,N]} 1.\overline{eat}^i : Food$$

The above process limits the total number of *breeding* animals in the population to N; it is possible to have many more *young* in the system, but their numbers are reduced by the limited availability of this resource. Note that $\overline{eat}^0$ is $\surd$.

Letting x_i, y_i, z_j, w_j, with $i \in I$ and $j \in J$ arbitrary index sets, range over some intial sex factor distribution, the complete system is given by:

$$Poly \stackrel{def}{=} \Theta((\textstyle\prod_i Male(x_i,y_i) \times \prod_j Female(z_j,w_j) \times Food)\lceil\{\surd\})$$

The above process describes the intial state of a number of animals under polygenic sex determination given by the sex allocation function we described earlier. It should be noted that sex allocation is the only point in the system at which the genetic make up of the individual is used to make a decision about its future. All other decisions, survival, number of offspring and mating success are *unbiased* with respect to genotype. Thus the system must be selectively neutral other than with respect to an animal's sex.

2.1 Solutions with Deterministic Allocation Function

With the above allocation function the system *Poly* will stabilise in one of the following three states.

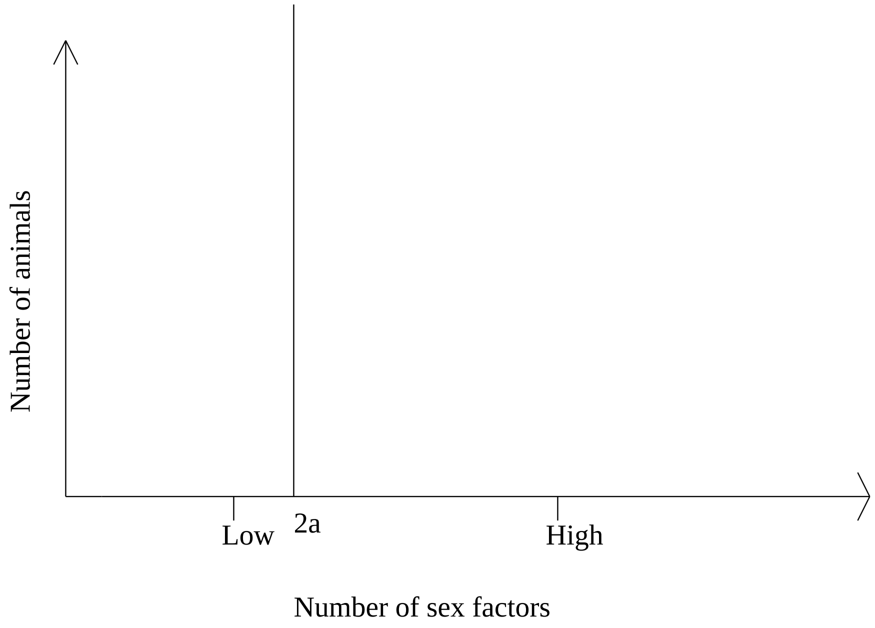

Figure 2: Location of type 1 solution.

Type 1 In this case there is **no** genetic sexual differentiation, all of the animals have the same number of sex factors on both chromosomes. The animals are (a, a) with $Low < a + a < High$. In terms of the allocation function all of the animals are in the undetermined region, see Figure 2. If there is an ESD modifier in this case then it will affect 100% of the animals. Notice that this achieves (on average) a 50/50 sex ratio.

Type 2 In this case the sex of the animals is entirely determined by genotype. Depending upon the precise values of Low and $High$ and the maximum number of copies on either chromosome, the system becomes either an XY or a ZW genetic sex determination (male or female heterogamety). In the former case females are (a, a) and males are (a, b) with $a + a \leq Low < High \leq a + b$, shown pictorially in Figure 3.

Alternatively the system can evolve into a ZW state. In this case female animals are (a, b) and male animals are (b, b) with $a + b \leq Low < High \leq b + b$. Pictorially the distribution of animals is given in Figure 4.

In either case the number of animals of each sex is half of the total population.

Type 3 In this case there is a mixture between animals whose sex is determined genetically and those whose sex is determined randomly. In this case females have genotypes (a, a) or (a, b) and males have genotype (a, b) or (b, b), with $a + a \leq Low < a + b < High \leq b + b$. Pictorially this genotype distribution is given in Figure 5.

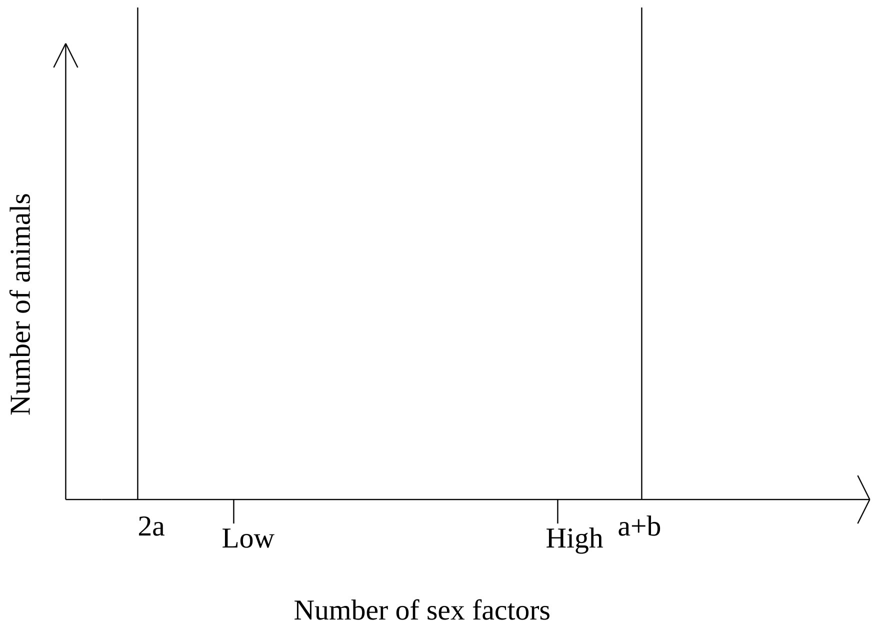

Figure 3: Location of type 2 solution for XY.

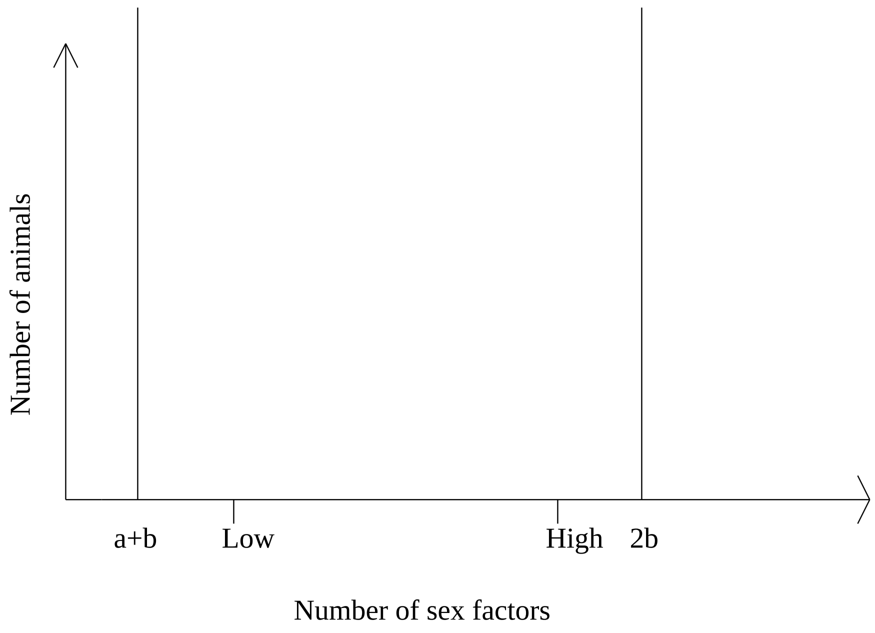

Figure 4: Location of type 2 solution for ZW.

Figure 5: Location of type 3 solutions.

Again this achieves a 50/50 sex ratio. However the populations of each genotype are not in the standard 1:2:1 ratio, see Section 2.4 for details.

There are several observations that can be made about special cases for the limits Low and $High$ and their effects on the solution classes possible.

If $Low + 1 = High$ (for example $Low = 12$ and $High = 13$, see also [BB82]) that is the boundaries are adjacent, then there is no space for type 1 or 3 solutions to occur and we have only XY and ZW type solutions occuring. Similarly if $Low + 2 = High$ and $Low + 1$ is an *odd* number (for example $Low = 12$ and $High = 14$) then there can be no Type 1 solutions.

Assuming that there is some maximum number of copies (max) of the polygene possible on either chromosome in the system and that $High + (Low/2) > max$ (for example $max = 20$, $Low = 4$ and $High = 25$) then there can be no Type 2 XY solutions. Since $2a < 4$ there is no choice for the b value chromosome since $b \leq 20$ with $b + a \geq 25$.

2.2 How stability arises (Informally)

Of note is the longstanding observation that sexual animals should allocate the sex of their offspring on a 50/50 basis[Fis30][9]. This results from the observation that if the population has a biased ratio in favour of either sex then an animal can ensure a greater proportion of its genes will be present in the population in 2 generations time by producing more of the rarer sex.

The stable solutions of the system arise as the result of the interaction of two forces:

[9]More preciseley equal resources should be invested in each sex

- the 'rare female effect' [Fis30];
- probabilistic symmetry breaking.

The first force is the result of departure from Fisherian [Fis30] sex ratio. If this occurs then the rarer sex will get to breed more frequently thus increasing its numbers in the subsequent generations.

The second force results from the absence of selection. Probabilistic drift will cause asymmetries in the numbers of factors on particular chromosomes within the population to arise. In the absence of any sexual (Fisherian) selection the system can continue to drift. As a result of feedback from culling, small biases in population numbers of particular gene counts tend to become exaggerated, and eventually some gene counts will be eliminated. **Only** if there is some selective advantage in keeping a particular gene count will it be retained. Thus the system attempts to reduce to the smallest number of different counts. Notice that for all of our solutions, if at this point any gene count's frequency is reduced then the population's sex ratio will depart from Fisher and rarity effects will restore that particular gene count's frequency in the population.

The precise solution that is reached depends on which gene counts are 'drifted out' in the early period of the system. For instance if gene count 8 is removed then the system cannot reach a solution based upon 8. Counts are removed, at random, until some selective pressure is encountered, at this point those that give a Fisherian ratio are retained and those that do not are eventually lost.

Whilst we can predict that the system eventually arrives in one of our three solution types, it is extremely difficult (and highly parameter dependent) to demonstrate which solutions it will arrive at with what probability.

2.3 Stability (Formally)

In order to demonstrate that the system will eventually reach one of the stable states stated above we have to demonstrate two facts [Tof94]:

1. the sets of states given by those situations are themselves closed (stable);
2. from an arbitrary state one or more of the stable sets is reachable.

Whilst this will demonstrate the eventual solutions of the system are stable it cannot tell us with what frequency each of the solution types will arise.

We proceed by establishing the first fact.

Proposition 2.1 *The following sets of states are closed (with i, j, k, l, m and n free in the positive naturals):*

1. *With $Low < 2a < High$:*

$$\begin{array}{ll} \{ & \Theta((\prod_i Male(a,a) \times \prod_j Female(a,a) \times Food)\lceil\{\surd\}) \\ & \Theta((\prod_i Pair(a,a,a,a) \times Food)\lceil\{\surd\}) \\ & \Theta((\prod_i Young(a,a,a,a) \times Food)\lceil\{\surd\}) \\ & \Theta((\prod_i Animal(a,a) \times Food)\lceil\{\surd\}) \\ & \Theta((\prod_i Ani(a,a) \times Food)\lceil\{\surd\}) \quad \} \end{array}$$

2. With $2a \leq Low$ and $High \leq a+b$:

$$\{ \quad \Theta((\prod_i Male(b,a) \times \prod_j Female(a,a) \times Food) \lceil \{\surd\})$$
$$\Theta((\prod_i Pair(b,a,a,a) \times Food) \lceil \{\surd\})$$
$$\Theta((\prod_i Young(b,a,a,a) \times Food) \lceil \{\surd\})$$
$$\Theta((\prod_i Animal(a,a) \times \prod_j Animal(b,a) \times Food) \lceil \{\surd\})$$
$$\Theta((\prod_i Ani(a,a) \times \prod_j Ani(b,a) \times Food) \lceil \{\surd\}) \qquad \}$$

symmetrically for $a+b \leq Low$ and $High \leq 2b$.

3. With $2a \leq Low < a+b < High \leq 2b$:

$$\{ \quad \Theta((\prod_i Female(a,a) \times \prod_j Female(a,b) \times \prod_k Male(a,b) \times \prod_l Female(b,a)$$
$$\times \prod_m Male(b,a) \times \prod_n Male(b,b) \times Food) \lceil \{\surd\})$$
$$\Theta((\prod_i Pair(b,b,a,a) \times \prod_j Pair(b,b,a,b) \times \prod_k Pair(b,a,a,a) \times \prod_l Pair(b,a,a,b)$$
$$\times \prod_m Pair(a,b,a,a) \times \prod_n Pair(a,b,b,a) \times Food) \lceil \{\surd\})$$
$$\Theta((\prod_i Young(b,b,a,a) \times \prod_j Young(b,b,a,b) \times \prod_k Young(b,a,a,a) \times \prod_l Young(b,a,a,b)$$
$$\times \prod_m Young(a,b,a,a) \times \prod_m Young(a,b,b,a) \times Food) \lceil \{\surd\})$$
$$\Theta((\prod_i Ani(a,a) \times \prod_j Ani(a,b) \times \prod_k Ani(b,a) \times \prod_l Ani(b,b) \times Food) \lceil \{\surd\}) \qquad \}$$

Proof: For the case 1 above it is straightforward to demonstrate closure from the definition of the processes. For the second and third cases there is a problem as to why they do not become a system of type 1. It is certainly possible for the systems to reach at least one of the following states:

$$\Theta((\prod_i Ani(a,a) \times Food) \lceil \{\surd\})$$
$$\Theta((\prod_i Ani(b,b) \times Food) \lceil \{\surd\})$$

In which case the system is now apparently of type 1. It should be noted that if either of these cases occur then there will be no more males (females) to pair with and the population will go extinct. The probability of this event is comparable to that for spontaneous extinction[10] as a result of all offspring being one sex (a possibility for all of our three systems). Thus it is not the case that the system can 'lose' either of the polygenes without the population going extinct, the process becoming equivalent to our *Dead* process. Hence we can see that the second and third types are stable in the same sense as the first, as long as spontaneous extinction as a result of monosexuality is avoided. It is this observation that helps in resolving the issues of reachability.

To demonstrate the inevitability of reaching one of our three stable states we observe from the proof above that particular gene counts will only be retained if they are necessary

[10] In some senses we are making a distinction between very long lived transients and short lived ones. There is an ultimate closed set of the singleton *Dead*, but the time scale to reach it is many orders of magnitude greater than evolutionary time. Since it results from all surviving offspring in a generation 'guessing' their sex incorrectly, it is a very low probability event. Thus we distinguish between stable sets which are persistant on this time scale from others.

for the maintenance of the presence of *both* sexes. Hence, in a complex system where many gene counts are present there is a non-zero probability that any particular count can be lost, either through the parent failing to mate or its offspring having the count failing to survive, this process will continue until one of the three states above is reached. It is sufficient to demonstrate the effect of adding a single alternative count into the above three systems and consider its consequences for the presences of both males and females to observe the non-existence of any other stable solutions.

Proposition 2.2 *The addition of animals with a new factor count to any of the stable states is not a stable state, and the system subsequently evolves to a stable state.*

Proof: We consider each of the three types in turn.

1. Consider a type 1 with $Low < 2a < High$ and an extra factor c then we have animals of type $2a$, $a + c$ and $2c$, there are two cases to consider:
   (a) $c < a$: then since $a + c$ and $2c$ fall into either the female only or the 50/50 regions we can lose the c gene without affecting sex ratio. if $2c$ is in the random region then we can lose the a gene, but will still have a type 1 solution but with c in place of a. Finally if both $a + c$ and $2c$ lie below Low then we *cannot* lose the a gene as it is the only source of males, but the system can lose the c without such problems.
   (b) $c > a$: as above if $2c \leq High$ then c can replace a otherwise c will inevitably be lost.

   Hence the system will always stay as type 1, it should be noted that the probability of losing a gene from the system in competition is many orders of magnitude greater than that of spontaneous extinction.
2. Consider an XY system with $2a \leq Low$ and $High \leq a + b$, and a competitor gene c, then there are three basic cases:
   (a) $c < a$: in this case $2c$ animals will always be female, if $b + c \geq High$ then the a or c genes are interchangable and the system will become a, b or c, b XY; failing that c can be evolved out, but a and b cannot so the system becomes a, b XY.
   (b) $a < c < b$: if $2c \leq Low$ then as above, similarly $2c \geq High$ see (c) below. So if $Low < 2c < High$ then the system can lose both a, b to become a type 1 system or lose the c to become type 2. In this case there is no evolutionary advantage and the timescale will be that of symmetry breaking.
   (c) $b < c$: by an identical argument to that for $c < a$ and results in either an a, c or a, b type 2 system, and there is no need to retain all three genes counts.
3. Type 3 with $2a \leq Low < a + b < High \leq 2b$, and introducing a new value c we have three cases

(a) $c < a$: if $Low \leq c + b \leq High$ then we can get c, b type 3 otherwise the system *must* evolve to a, b Type 3.

(b) $a < c < b$: if $2c \leq Low$ and $Low < c + b < High$ then it can evolve to either a, b or c, b type 3; alternatively if $c + b \geq High$ then it can become either a, b type 3 or c, b XY. Similarly if $2c \geq High$ and $Low < a + c < High$ then it can evolve to a, c or a, b type 3; or if $a + c \leq Low$ it can become a, c ZW or a, b type 3. If $Low < 2c < High$ then we have a competition between c, c type 1 or a, b type 3, but the system will eventually reach one of them.

(c) $b < c$: if $Low < a + c < High$ then it can evolve to a, c or a, b type 3, otherwise it will become a, b type 3.

The basic method of system evolution is to lose particular gene counts until a point is reached at which selection for the rarer sex forces the sex ratio to equality. Those genes that do not aid in the maintenance of equal sex ratio are evolved out, otherwise there is a competition between essentially equivalent genes which will be resolved by symmetry breaking. That is once one of the competitors gets an advantage then it is likely to increase that advantage in the next generation until it reaches fixation. This process is very fast in comparison with that of spontaneous extinction.

In a particular competitive circumstance the actual resolution will depend upon the initial frequencies of the various gene counts. This is of particular relevance when considering the effect of mutation or recombination upon the stable states of the system. In particular if mutants arise infrequently, or on becoming dominant would lead to a non-Fisherian sex ratio, then they are highly unlikely to become common within a stable population.

Note that in the above, type 3 solutions in competition will often be changed into type 1 or 2 and thus we could expect to find relatively few of these solutions.

2.4 Genotype Frequencies in Type 3

Earlier we stated that the distribution of genotypes in a type 3 system was not the standard 1:2:1 ratio. In this section we calculate the stable genotype frequencies, assuming that animals have to occur in male/female pairs to breed. From the type 3 system we have that females are (a, a) or (a, b) and males are (a, b) or (b, b). We can assume from symmetry that the proportions of (a, a) and (b, b) animals must be the same. Hence, we proceed by assuming that a proportion β of *males* are (a, b) and that α of *males* are (b, b) and for *females* β are (a, b) and that α of *females* are (a, a); tracing the breeding ratios (we label the axes with the frequency of males or females of that type). Hence the gene frequencies of the offspring are as follows:

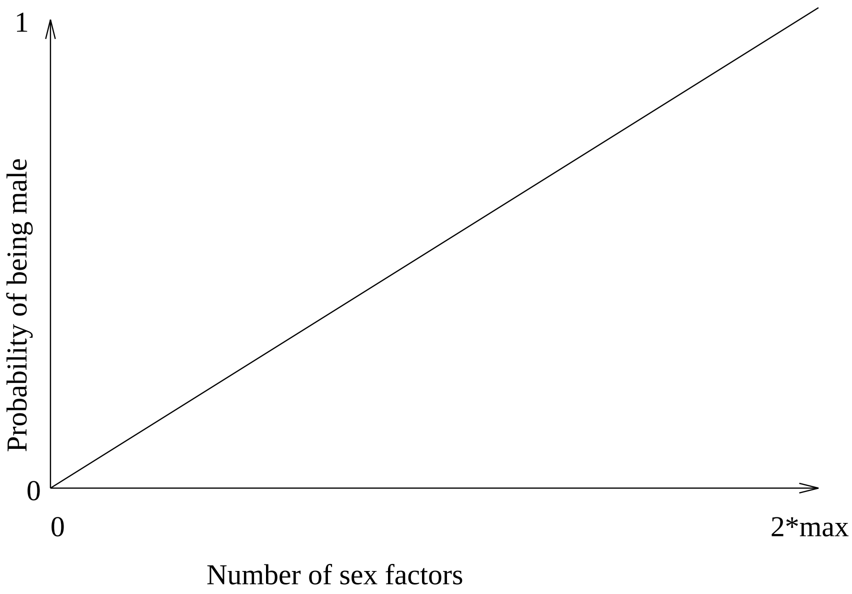

Figure 6: The result of breeding two polygenic shrimp.

$females$	$(a,a)\alpha$	$(a,b)\beta$
$males$ $(a,b)\beta$	$(a,a)\frac{1}{2}\alpha\beta$ $(a,b)\frac{1}{2}\alpha\beta$	$(a,a)\frac{1}{4}\beta^2$ $(a,b)\frac{1}{2}\beta^2$ $(b,b)\frac{1}{4}\beta^2$
$(b,b)\alpha$	$(a,b)\alpha^2$	$(a,b)\frac{1}{2}\alpha\beta$ $(b,b)\frac{1}{2}\alpha\beta$

The above system is stable when the resultant number of a genotype remains the same in subsequent generations. Solving for genotype (a,b) we obtain the equation $\beta = \frac{1}{2}\beta^2 + \alpha^2 + \alpha\beta$ which simplified with the eqaution $\alpha + \beta = 1$ gives us the quadratic equation $\beta^2 - 4\beta + 2 = 0$ which has a solution in $[0,1]$ of $2 - \sqrt{2}$ (or approximately 0.59). In the above eqautions setting $\alpha = \frac{1}{2}$ and $\beta = \frac{1}{2}$ does not give a stable solution.

So the ratio of genotypes in type 3 is 0.205: 0.59 : 0.205 or approximately 1:3:1. Exactly $\frac{\sqrt{2}-1}{2} : 2 - \sqrt{2} : \frac{\sqrt{2}-1}{2}$.

3 Allocation with Probabilistic Determination

In this case we assume that the probability of being **male** is a linear function of the total number of sex factors carried (see Figure 6).

Formally this is represented by changing the definition of the state $Animal(x,y)$ to the following:

$$Animal(x,y) \stackrel{def}{=} (x+y).\surd : Male(x,y) + (2max - (x+y)).\surd : Female(x,y)$$

Whilst leaving the definitions of the other states unchanged.

In this case we lose the Type 2 solutions of the deterministic function and only retain the Type 1 and 3 solutions. Initially one might assume that all solutions should be of Type 1, but initially the sex bias selection pressure is very low, so the genes necessary for a Type 1 solution (for example with $max = 20$ then 10, 11 or 9) could have been lost by probabilistic drift. In this case the best solution the system can adopt is a Type 3 with a sex ratio as close to Fisher as possible. In the latter case, as above, we do not generally observe population ratios of 1:2:1.

3.1 Determining population makeup.

In the case of the Type 3 solution we give some gene flow equations to calculate the stable population. Take the following collection of variables:

- x the proportion of (a, a) animals in the population;
- y the proportion of (a, b) animals in the population;
- z the proportion of (b, b) animals in the population;
- f_x the probability of an (a, a) animal being female;
- f_y the probability of an (a, b) animal being female;
- f_z the probability of an (b, b) animal being female.

We define $m_x = 1 - f_x$ and m_y, m_z respectively.

Then define the proportion of females f in the population to be $f = xf_x + yf_y + zf_z$, and the proportion of males m in the population to be $m = xm_x + ym_y + zm_z$.

Consider the matings that lead to an (a, a) animal. This can be the result of $(a, a), (a, a)$ or $(a, a), (a, b)$ or $(a, b), (a, b)$ matings. So the proportion of (a, a) animals in the next generation is the probability of an $(a, a), (a, a)$ mating times the probability it results in an (a, a) animal, plus the probability of an $(a, a), (a, b)$ mating times the probability it results in an (a, a) animal plus the probability of an $(a, b), (a, b)$ mating times the probability it results in an (a, a) animal.

So a similar chart to our earlier one for type 3 solutions in the determined systems would be:

$females$	$(a,a)\frac{xf_x}{f}$	$(a,b)\frac{yf_y}{f}$	$(b,b)\frac{zf_z}{f}$
$males$ $(a,a)\frac{xm_x}{m}$	$(a,a)\frac{xm_x}{m}\frac{xf_x}{f}$	$(a,a)\frac{1}{2}\frac{xm_x}{m}\frac{yf_y}{f}$ $(a,b)\frac{1}{2}\frac{xm_x}{m}\frac{yf_y}{f}$	$(a,b)\frac{xm_x}{m}\frac{zf_z}{f}$
$(a,b)\frac{ym_y}{m}$	$(a,a)\frac{1}{2}\frac{ym_y}{m}\frac{xf_x}{f}$ $(a,b)\frac{1}{2}\frac{ym_y}{m}\frac{xf_x}{f}$	$(a,a)\frac{1}{4}\frac{ym_y}{m}\frac{yf_y}{f}$ $(a,b)\frac{1}{2}\frac{ym_y}{m}\frac{yf_y}{f}$ $(b,b)\frac{1}{4}\frac{ym_y}{m}\frac{yf_y}{f}$	$(a,b)\frac{1}{2}\frac{ym_y}{m}\frac{zf_z}{f}$ $(b,b)\frac{1}{2}\frac{ym_y}{m}\frac{zf_z}{f}$
$(b,b)\frac{zm_z}{m}$	$(a,b)\frac{zm_z}{m}\frac{xf_x}{f}$	$(a,b)\frac{1}{2}\frac{zm_z}{m}\frac{yf_y}{f}$ $(b,b)\frac{1}{2}\frac{zm_z}{m}\frac{yf_y}{f}$	$(b,b)\frac{zm_z}{m}\frac{zf_z}{f}$

We can generate three equations in our variables x, y and z, which will hold when the genotype proportions are stable, from the above matrix:

$$\begin{aligned} x &= \frac{xm_x}{m}\frac{xf_x}{f} + \frac{1}{2}\frac{xm_x}{m}\frac{yf_y}{f} + \frac{1}{2}\frac{ym_y}{m}\frac{xf_x}{f} + \frac{1}{4}\frac{ym_y}{m}\frac{yf_y}{f} \\ y &= \frac{xm_x}{m}\frac{zf_z}{f} + \frac{1}{2}\frac{xm_x}{m}\frac{yf_y}{f} + \frac{1}{2}\frac{ym_y}{m}\frac{zf_z}{f} + \frac{1}{2}\frac{ym_y}{m}\frac{xf_x}{f} \\ &\quad + \frac{1}{2}\frac{ym_y}{m}\frac{yf_y}{f} + \frac{zm_z}{m}\frac{xf_x}{f} + \frac{1}{2}\frac{zm_z}{m}\frac{yf_y}{f} \\ z &= \frac{1}{2}\frac{ym_y}{m}\frac{zf_z}{f} + \frac{1}{4}\frac{ym_y}{m}\frac{yf_y}{f} + \frac{zm_z}{m}\frac{zf_z}{f} + \frac{1}{2}\frac{zm_z}{m}\frac{yf_y}{f} \end{aligned}$$

Unfortunately the above equations are too complicated to solve in general. For illustration we provide some numerical solutions to these equations for particular values of f_x, f_y and f_z:

f_x	f_y	f_z	x	y	z
1	1/2	0	0.2071	0.5857	0.2071
3/4	1/2	1/4	0.236	0.527	0.236
17/20	1/2	3/20	0.225	0.549	0.225
13/20	9/20	5/20	0.36	0.49	0.14
13/20	8/20	2/20	0.44	0.46	0.09

From the above table it is clear that populations with equilibriums far away from that of Hardy-Weinberg are possible.

4 The Effect of Recombination

For our purposes we take recombination as follows. Given an animal with genotype (a,b) then produce an animal with genotype (c,d) with c linearly distributed in the range 0 to $a+b$ and $d=(a+b)-c$. We assume a variable rate at which this process occurs in animals. Since sex is determined by the total number of sex factors, which is left invariant by the above, we can impose recombination at any stage in our system as it will be independent of sex choice and gene selection.

Formally the presence of recombination in our evolving system can be represented by the following processes:

$$
\begin{array}{rcll}
RMale(x,y) & \stackrel{def}{=} & \omega.\overline{mate}(x,y) : Dead + 1.\surd : Dead & \\
RFemale(z,w) & \stackrel{def}{=} & \omega.mate(x,y) : RPair(x,y,z,w) + 1.\surd : Dead & \\
RPair(x,y,z,w) & \stackrel{def}{=} & \sum_{i \in Offspring} w_i.\surd : RYoung(x,y,z,w)^i & \\
RYoung(x,y,z,w) & \stackrel{def}{=} & 1.\surd : RAnimal(x,z) & \\
 & & +1.\surd : RAnimal(x,w) & \\
 & & +1.\surd : RAnimal(y,z) & \\
 & & +1.\surd : RAnimal(y,w) & \\
RAnimal(x,y) & \stackrel{def}{=} & \omega.eat : RAniR(x,y) + 1.\surd : Dead & \\
RAni(x,y) & \stackrel{def}{=} & \left\{ \begin{array}{ll} \surd : RMale(x,y) & \text{if } x+y \geq High \\ \surd : RFemale(x,y) & \text{if } x+y \leq Low \\ 1.\surd : RFemale(x,y) + 1.\surd : RMale(x,y) & \text{otherwise} \end{array} \right. &
\end{array}
$$

In the above the process $RaniR$ has yet to be defined and we do so below:

$$
\begin{array}{rcl}
RAniR(x,y) & \stackrel{def}{=} & norec.\surd : \surd : RAni(x,y) + rec.\surd : RecAni(x,y) \\
RecAni(x,y) & \stackrel{def}{=} & \sum_{i \in [0,x+y]} 1.\surd : RAni(i,(x+y)-i)
\end{array}
$$

In the above the variables $norec$ and rec express the probability that recombination will occur in this animal. The process $RecAni$ performs the recombination, it is at this point that differences emerge between our polyfactorial system and one dependent on multi-factors[11]. If the animal doesn't recombine then we need to wait an extra tick to allow the system to resynchronise.

Recombination allows gene counts that have been lost to be reintroduced, it is a mechanism for avoiding the consequences of 'bad' choices (Muller's ratchet [Mul64, Fel74]). Unfortunately for an analysis of the stable solutions of our system recombination is a major challenge. The behaviour of the system is given by a dynamic competition between drift removing rare genotypes and recombination reintroducing them. These rates are very difficult to estimate and are highly dependent on the parameter values. Essentially the rate of drift is a function of effective population size, and the recombination rate is dependent on the degree of linkage of the genes. Traditional analyses of these complex systems assume that linkage is low [Bar86], however other authors [Wri68] suggest that linkage between related genes may be very high.

However there are some observations that we can make dependent on the relative speed of the underlying rates in the system:

- Recombination very much slower than drift. In this case our system will behave as before. It will reach an essentially stable solution, which will be characterised by brief departures from it as a result of a recombination event. It is extremely unlikely that the system's nature will change.

[11] In a multi-factor system an appropriate recombination rule would be:
$RecAni(x,y) \stackrel{def}{=} 1.\surd : Rani(x,y) + 1.\surd : Rani(y,x)$
effectively swapping the genetic material from the mother and father. This form of recombination would have **no** effect on our conclusions.

- Recombination and drift rates are comparable. In this case the system will tend to be constantly in flux between different possible stable solutions, occasionally residing in one for a period before moving off to another.
- Recombination faster than drift. The system never stabilises but the observed distribution of polyfactors is **highly** non-Guassian [TB90]. The effect of sex selection forbids certain states and the system will tend to avoid them.

Essentially recombination has an effect as structure starts to emerge in the system. If it is 'strong enough' to break down the structure then it will probably never emerge, otherwise the structures we see can emerge despite the presence of the recombination. In simulo [HT94b] rates of recombination of much greater than 1/10,000 in a population of 2000 animals caused the system to start to oscillate, and recombinant rates of the order of 1/100 cause a complete loss of any structure. Note, the system does not go to a normal distribution, it becomes highly variable, exhibiting almost no structural properties, in a sense it is completely noise dominated.

5 Multiple Chromosomes

We start be analysing the system when the polygenic factors are distributed over two chromosomes. This extension of our earlier system requires that we track 4 factor counts rather than 2, representing the counts on each of the 4 chromosomes.

A male process with genotype ma_1, ma_2, mb_1, mb_2 tries to export its genes to a female process and then dies:

$$2Male(ma_1, ma_2, mb_1, mb_2) \overset{def}{=} \omega.\overline{mate}(ma_1, ma_2, mb_1, mb_2) : Dead + 1.\surd : Dead$$

We abbreviate the formal parameter list $ma_1, ma_2, mb_1, mb_2, fa_1, fa_2, fb_1, fb_2$ by $\tilde{p}$. A female process takes in the genes specified by the male to produce a process with their combined genetic information:

$$2Female(fa_1, fa_2, fb_1, fb_2) \overset{def}{=} \omega.mate(ma_1, ma_2, mb_1, mb_2) : 2Pair(\tilde{p}) + 1.\surd : Dead$$

The pair now produces some number of offspring which will acquire their genetic type in the correct proportions from their parents:

$$2Pair(\tilde{p}) \overset{def}{=} \sum_{i \in Offspring} off_i.\surd : 2Young(\tilde{p})^i$$

In the above off_i represents the probability of having i offspring, and $Offspring$ the set or permissible values of number of young.

The young now acquire their genotype in the correct proportions:

$$
\begin{aligned}
2Young(\tilde{p}) \stackrel{def}{=} \; & 1.\surd : 2Animal(ma_1, fa_1, mb_1, fb_1) \\
& +1.\surd : 2Animal(ma_1, fa_1, mb_1, fb_2) \\
& +1.\surd : 2Animal(ma_1, fa_1, mb_2, fb_1) \\
& +1.\surd : 2Animal(ma_1, fa_1, mb_2, fb_2) \\
& +1.\surd : 2Animal(ma_1, fa_2, mb_1, fb_1) \\
& +1.\surd : 2Animal(ma_1, fa_2, mb_1, fb_2) \\
& +1.\surd : 2Animal(ma_1, fa_2, mb_2, fb_1) \\
& +1.\surd : 2Animal(ma_1, fa_2, mb_2, fb_2) \\
& +1.\surd : 2Animal(ma_2, fa_1, mb_1, fb_1) \\
& +1.\surd : 2Animal(ma_2, fa_1, mb_1, fb_2) \\
& +1.\surd : 2Animal(ma_2, fa_1, mb_2, fb_1) \\
& +1.\surd : 2Animal(ma_2, fa_1, mb_2, fb_2) \\
& +1.\surd : 2Animal(ma_2, fa_2, mb_1, fb_1) \\
& +1.\surd : 2Animal(ma_2, fa_2, mb_1, fb_2) \\
& +1.\surd : 2Animal(ma_2, fa_2, mb_2, fb_1) \\
& +1.\surd : 2Animal(ma_2, fa_2, mb_2, fb_2)
\end{aligned}
$$

This is the standard pattern of inheritance in a diploid system.

Now the animals have to eat in order to survive, again we abbreviate the formal parameter list aa_1, aa_2, ab_1, ab_2 by $\tilde{a}$:

$$2Animal(aa_1, aa_2, ab_1, ab_2) \stackrel{def}{=} \omega.eat : 2Ani(\tilde{a}) + 1.\surd : Dead$$

Finally, the animals sex is determined, we abbreviate $aa_1 + aa_2 + ab_1 + ab_2$ as σ_a:

$$2Ani(\tilde{a}) \stackrel{def}{=} \begin{cases} \surd : 2Male(\tilde{a}) & \text{if } \sigma_a \geq High \\ \surd : 2Female(\tilde{a}) & \text{if } \sigma_a \leq Low \\ 1.\surd : 2Female(\tilde{a}) + 1.\surd : 2Male(\tilde{a}) & \text{otherwise} \end{cases}$$

Letting $xa1_i, xa2_i, xb1_i, xb2_i, ya1_j, ya2_j, yb1_j, yb2_j$, with $i \in I$ and $j \in J$ arbitrary indexing sets, range over some initial sex factor distribution, the complete system is given by:

$$2Poly \stackrel{def}{=} \Theta((\prod_i 2Male(xa1_i, xa2_i, xb1_i, xb2_i) \times \prod_j 2Female(ya1_i, ya2_i, yb1_i, yb2_i) \times Food)\lceil\{\surd\})$$

The system $2Poly$ is the same as the system $Poly$ in that the sex decision process and all other evolutions are identical excepting that the polygenic factors are now distributed over 2 chromosome pairs.

5.1 Stable solutions with 2 Pairs

Observing how the stable solutions arose in our earlier system we can envisage that with 2 chromosomes present the system will evolve by losing factors until a point is reached at which the factors have to be retained in order that the system does not go extinct. The first observation is that the two chromosome pairs evolve *independently*. There is no communication between them. Hence, there can be no selection for preservation of structure between the chromosomes. Therefore any stable solution, for both chromosomes, must be stable with respect to each pair separately as no exchange of factors between them is possible. That is to say chromosomes belonging to *different* pairs do not recombine together.

So as candidate stable solutions we can consider the systems we observed earlier, noting that we can perform the proof of stability on each of the chromosomes separately as we did before.

This gives rise to 6 possible combinations of pairs (since order is unimportant)

1. Type 1 $\times$ Type 1: in this case we have a_1 and a_2 on the two chromosomes and given $Low \leq 2(a_1 + a_2) \leq High$ the system will be stable.

2. Type 1 $\times$ Type 2: in this case we have a_1 and a_2, b_2 so if $2(a_1 + a_2) < Low$ and $2a_1 + a_2 + b_2 > High$ the system[12] is stable.

3. Type 1 $\times$ Type 3: in this case we have a_1 and a_2, b_2 and as long as $2(a_1 + a_2) < Low \leq 2a_1 + a_2 + b_2 \leq High < 2(a_1 + b_2)$ the system will be stable.

4. Type 2 $\times$ Type 2 in this case we have a_1, b_1 and a_2, b_2 factors present in the system. Let $s_i = 2a_i$ and $l_i = a_i + b_i$ $i = 1, 2$ then (WLOG) there are two potentially stable cases:

   (a) $s_1 + s_2$ and $s_1 + l_2$ less than Low, as well as $l_1 + s_2$ and $l_1 + l_2$ greater than $High$: if b_2 becomes a_2 then we still get both type below and above the limits respectively, thus maintaining the sex ratio. So in this case the pair that does not change the position of the other values with respect to the limits can be evolved out to a Type 1 $\times$ Type 2.

   (b) $s_1 + s_2 < Low \leq s_1 + l_2,\ l_1 + s_2 \leq High < l_1 + l_2$: now if either of the pairs loses one of the gene counts then one of the determined regions will be lost. So given that none of $2(a_1 + a_2)$, $2(a_1 + b_2)$, $2(b_1 + a_2)$ and $2(b_1 + b_2)$ can form a type 1 solution then the above will be stable. This is the only **new** form of solution we observe, and gives rise to up to 9 genotypes, we leave an analysis of the population genetics of this system for later, as the system is highly unlikely to evolve in practice.

5. Type 2 $\times$ Type 3: as above we have l_1 and s_1 modifying the small, medium and large values in the type 3 solution ($s_2 = 2a_2$, $m_2 = a_2 + b_2$ and $l_2 = 2b_2$). In this case there are two arrangements, either there is no-overlap between the two versions of the type three and they lie outside the boundaries in which case the system will become a Type 1 $\times$ Type 2 of either a_2, a_2 and a_1, b_1 or b_2, b_2 and a_1, b_1. Alternatively the two sets overlap; then if $s_1 + s_2,\ l_1 + s_2 < Low \leq s_1 + m_2,\ l_1 + m_2 \leq High < s_1 + l_2,\ l_1 + l_2$ in which case either of two Type 1 $\times$ Type 3 will evolve. Otherwise the system will evolve to Type 1 $\times$ Type 1.

[12] It may seem possible to lose the a_1 if $2a_2 < Low$ and $a_2 + b_2 > High$, but there is no other alternative gene to go on chromosome pair one so a_1 will be maintained. In the presence of mutation or recombination, however, the information on pair one could be lost.

6. Type 3 $\times$ Type 3: the system will evolve to Type 1 $\times$ Type 1 unless the following pattern occurs: defining $s_i = 2a_i$, $m_i = a_i + b_i$ and $l_i = 2b_i$ then if $s_1 + s_2$, $s_1 + m_2$, $s_1 + l_2 < Low \leq m_1 + s_2$, $m_1 + m_2$, $m_1 + l_2 \leq High < l_1 + s_2$, $l_1 + m_2$, $l_1 + l_2$ in which case the differences in a_2 and b_2 are irrelevant and they can be evolved out, leading to a Type 1 $\times$ Type 3. There is a symmetric case when pair 2 dominates.

At this stage we have identified essentially four types of solutions: our previous three solution types, and a new version of the mixed type in which there are four defining factors on two chromosome pairs.

By the same arguments used to demonstrate that our previous three solutions were the only ones possible in a single chromosome system we could demonstrate that the above are the only stable solutions for 2 chromosome pairs.

5.2 More than 2 Chromosomes

As we in Section 5.1, introducing an extra chromosome pair into a system generally leads to the information on one of the pairs being lost. If we consider crossing a further chromosome pair with the solutions above then we obtain no new stable solutions, and this continues as we consider crossing over as many extra chromosomes as we wish to add.

Essentially all but one (occasionally two) of the chromosome pairs can be treated as an additive constant which move the limits appropriately for the 'lead' sex determining chromosome pair. Thus if the polygenes are distributed over many chromosomes we would expect to see one of the four following outcomes:

- All but two of the chromosome pairs have identical counts of the polygenes and the last two are arranged so that the totals over all the pairs are as 4b;
- All but one of the chromosome pairs have identical numbers and one pair is arranged as type 1 2 or 3 as for a single chromosome pair.

6 Conclusions

We have presented a polygenic system which can evolve, in the absence of selection, into one of three basic types when present on a single chromosome pair. The first two of these three solutions represent well known forms of sex determination. The first is where sex is determined with respect to environmental cues alone (ESD; [CB77]). The second is where sex is purely genetically determined; the system evolving to one of two states where sex is determined according to the genotype based on one chromosome pair (sex chromosomes; [Ohn67]). These states (reviewed in [Bul83]) are male heterogamety (male is XY, female is XX; as in mammals) and female heterogamety (female is ZW, male is ZZ; *e.g.*, birds). The third solution is essentially a combination of the previous two. Its existence in natural populations has not been explicitly demonstrated, although there is considerable evidence of mixed ESD/genotypic systems which may conform to the structure [BVB82, NAG88a, LC93, CVE92, Jan92, WA94].

Whilst it is known [Ric86] that polygenic sex determination is unstable to invasion by mutants, we have demonstrated that in finite populations polygenic systems of sex determination are *intrinsically* unstable. A polygenic sex determining system can evolve towards a conventional system of genetic sex determination without the introduction of any sex determining mutants. The outcome is purely the result of two processes that affect factor frequencies in the gene pool. Random drift reduces genetic variability as a consequence of sampling error in finite populations [Wri68, Wri69]. Fisherian selection [Fis30] maintains only those combinations of factors that yield an unbiased population sex ratio; this latter process occurs as the rarer sex necessarily makes a relatively greater genetic contribution to the following generation.

In the presence of selective advantage for ESD we might expect to see a preference for type 1 or 3 solutions but we leave that analysis for a later date. It is clear from our results that the assumption that the stable distribution for polygenic sex determining factors is Guasssian is false (as in [TB90]) if there is a high degree of linkage, even if the factors are distributed over many chromosome pairs.

Of particular interest is the observation that generally if the factors are distributed over many chromosomes then one (occasionally two) pairs will end up with absolute influence on sex determination. Hence we see the automatic evolution of sex chromosomes, with the gene count on the autosomes (non sex chromosomes) all being identical between members of each pair. These latter observations concur with some recent genetic observations where DNA sequences homologous to known sex factors on sex chromosomes have been found widely distributed throughout the autosomes [Sin88]. These factors appear to be present in wide range of species[DF88, BHS88, Pag87]. Indeed [Pag87] suggest that a possible model consistent with their findings would be one of additive sex factors.

From the perspective of the original problem of distinct ESD and GSD *Gammarus* populations, the outcome is very satisfactory. We can account for the observation of differing sex determination methods within the same species from the result that one of the populations is of solution type 1 and the other is of type 3, both having diverged from a single ancestral polygenic state. We may use the same observation to interpret a similar cross-species pattern found in the fish *Menidia menidia* [CVE92, LC93].

From the perspective of experimental design, most of the solutions give a small number of possible outcomes for offspring sex ratio, dependent on the particular female/male cross. The populations with the most complex structure are those that arise as a probabilistic type 3. In that case there can be a considerable variation in familial sex ratio.

There are many questions of rate left to study. Often when two viable solutions are competing the system can remain stable in this configuration for a very long time. Whilst the process of symmetry breaking is fast when compared with that of spontaneous extinction, it can be very slow (in large populations) in comparison with the generational time scale. When the effective population is many thousands it can take thousands of generations for a symmetry break to occur. Equally there are questions of the probability of particular solutions being reached both in the presence and absence of selective advantage. We hope to address these questions in a later paper.

Taking a polygenic (multifactorial) system as the genetic basis of sex determination we can observe evolution to many of the known forms of sex determination. More interestingly we see that in a multichromosome system generally one of the pairs will take the lead in determining the animal's gender. Finally it is possible within the same basic genetic system to have both environmental and genetic sex determination.

In a general sense the process algebraic approach to the description of evolutionary systems has great potential. The question of the effect of point of gene expression on evolutionary dynamics [HT94a] may be very precisely addressed in this context. Within our process algebraic description is a very general evolutionary system which could be exploited to investigate problems of evolution other than sex. The effects of linkage and finite population size in evolution are known to be very hard to understand [Wri68, Wri69]. The process algebraic approach may give new insights in this area.

7 Acknowledgements

Chris Tofts is supported by an EPSRC Advanced Fellowship (B/93/AF/1556). Melanie Hatcher is supported by an NERC Research Grant (GST/02/688). The authors thank Alison Dunn and Jonathan Adams for many useful discussions of this work. The final version of this paper is much improved as a result of diligent reading of an early draft by Alan Williams.

References

[AGN87] Adams, J., Greenwood, P. & Naylor, C. Evolutionary aspects of environmental sex determination. Int. J. Inv. Rep. Dev. 11, 123-135, 1987.

[Aid36] Aida, T., Genetics 21, 136, 1936.

[Bar86] Barton, N.H., The Maintenance of Polygenic Variation Through a Balance Between Mutation and Stabilising Selection, Genet. Res. 47(3):209-216, 1986.

[BlC89] Blackmore, M.S. & E.L. Charnov. Adaptive variation in environmental sex determination in a nematode. Am. Nat., 134, 817-823, 1989.

[Bul80] Bull, J.J. Sex determination in reptiles. Q. Rev. Biol., 55, 3-21, 1980.

[BC88] Bull, J.J. & Charnov, E.L. How fundamental are Fisherian sex ratios? *Oxf. Surv. Evol. Biol.* **5**, 96-135, 1988.

[Bul83] Bull, J. Evolution of Sex Determining Mechanisms, Benjamin/Cummings, California, 1983.

[BVB82] Bull, J.J., R.C. Vogt & M.G. Bulmer Heritability of sex ratio in turtles with environmental sex determination. Evolution, 36, 333-341, 1982.

[BHS88] Bull, J., D. Hillis and S. O'Steen, Mamalian ZFY Sequences Exist in Reptiles Regardless of Sex-Determining Mechanism, Science 24: 567-568.

[BB82] Bulmer, M.G. and J.J. Bull, Models of Polygenic Sex Determination and Sex Ratio Evolution, Evolution 36: 13-26, 1982.

[BB91] Belew, R., and L.B., Brooker, **Proceedings of the Fourth (4th) International Conference on Genetic Algorithms University of California**, Morgan Kaufman, 1991.

[BW88] Birky, C. W., and J.B. Walsh, Effects of Linkage on Rates of Molecular Evolution, Proc. Natl. Acad. Sci. USA 85: 6414-6418, 1988.

[BW90] Baeten, J.C.M. and W.P. Weijland, Process Algebra, Cambridge Tracts in Theoretical Computer Science 18, Cambridge University Press 1990.

[Cha82] Charnov, E.L. The Theory of Sex Allocation. Princeton: Princeton University Press, 1982.

[CB77] Charnov, E.L. & J.J. Bull. When is sex environmentally determined? Nature 266, 828-830, 1977.

[CHJA81] Charnov, E.L., R.L. Los-den Hartogh, W.T. Jones & J. van den Assem. Sex ratio evolution in a variable environment. Nature 289, 27-33, 1981.

[CVE92] Conover, D.O., D.A. Vanvoorhees & M. Ethisham. Sex ratio selection and the evolution of environmental sex determination in laboratory populations of *Menidia menidia*. Evolution, 46, 1722-1730, 1992.

[CK81] Conover, D.O. & B.E. Kynard. Environmental sex determination - Interaction of temperature and genotype in a fish. Science, 213, 577-579, 1981.

[CC84] Cox, T.F. and M.A.A. Cox, The Maximum Non-unit Eigenvalue for the Wright-Fisher Model of Random Drift Incorporating Selection, Theor. Popul. Biol. 25:29-40, 1984.

[CK70] Crow, J.F. and M. Kimura, An introduction to Population Genetics Theory, Harper and Row, New York, 1970.

[Dav87] Davis, L., **Genetic Algorithms and Simulated Annealing**, Research Notes in Artificial Intelligence, Pitman, 1987.

[Dav91] Davis, L., **Handbook of Genetic Algorithms**, Van Nostrand Reinhold, 1991.

[DAS93] Dunn, A.M., Adams, J. & Smith, J.E., Transovarial transmission and sex ratio distortion by a microsporidian parasite in a shrimp. *J. Invert. Pathol.* **61**, 248-252, 1993.

[DF88] Deeming, D. C., and M.W.J. Ferguson, Environmental regulation of sex determination in reptiles, Phil. Trans. R. Soc. Lond. B 322:19-39, 1988

[Esh75] Eshel, I. Selection on sex ratio and the evolution of sex-determination. *Heredity* **34**, 351-361, 1975

[Fel74] Felsenstein, J., The Evolutionary Advantage of Recombination, Genetics 78:737-756, 1974.

[Fis30] Fisher, R.A. **The Genetical Theory of Natural Selection**. Oxford: Clarendon Press, 1930.

[Fre76] Fredga, K., A. Groop, H. Winking and F. Frank, Fertile XX and XY type females in the wood lemming *Myopus schisticolor*, Nature 261: 225-227, 1976.

[FS88] Frank, S.A. & Swingland, J.R., Sex ratio under conditional sex expression. *J. theor. Biol.* **135**, 415-418, 1988.

[Gol89] Goldberg, D. E., **Genetic Algorithms in Search, Optimisation and Machine Learning**, Addison-Wesley, 1989.

[Gor54] Gordon, M., Two Opposing Sex Determing Mechanisms, one XX-XY and the other WY-YY in Different Natural Populations of the Platyfish, *Xiphophorus maculatus*, Proc 9th Int. Cong. Gen. 2:960-964, 1954.

[GW88] Gabriel, W., and G. P. Wagner, Parthenogenetic Populations Can Remain Stable in Spite of High Mutation Rate and Random Drift, Naturwissenschaften 75:204-205, 1988.

[Ham67] Hamilton, W.D. Extraordinary sex ratios. *Science* **156**, 477-488, 1967.

[Ham72] Hamilton, W.D. Altruism and related phenomena, mainly in social insects. *Ann. Rev. Ecol. Syst.* **3**, 193-232, 1972.

[Hen88] M. Hennessey, **Algebraic theory of Processes**, M.I.T. Press 1988.

[Hoa85] Hoare, C. A. R., **Communicating Sequential Processes**, Prentice-Hall 1985.

[How42] Howard, H.W. The genetics of *Armadillidium vulgare* Latrielle. II. Studies on the inheritance of monogeny and amphogeny. *J. Genetics* **44**, 143-159, 1942.

[HT94a] Hatcher, M.J. and C. Tofts, The Effect of Point of Expression on ESS Sex-ratio, in press J. Theor. Biol.

[HT94b] Hatcher, M.J. and C. Tofts, The Evolutionary Dynamics of a Polygenic Sex Determination System, in preperation.

[Jan92] Janzen, F.J., Heritable variation for sex ratio under environmental sex determination in the common snapping turtle (*Chelydra serpentina*). Genetics, 131, 155-161, 1992.

[Jon83] Jones, K.W., Evolution of Sex Chromosomes, in Development in Mammals, Vol 5, ed. M.H. Johnston, 297-320, Elsevier New York, 1983.

[Kim83] Kimura, M., The Neutral Theory of Molecular Evolution, Cambridge University Press, Cambridge.

[KL86] Karlin, S. and S. Lessard, Theoretical Studies on Sex Ratio Evolution, Princeton University Press, 1986.

[Kos64] Kosswig, C., Polygenic Sex Determination, Expermintia 20:190-199, 1964.

[LC93] Lagomarisino, I.V. & D.O. Conover. Variation in environmental and genotypic sex determining mechanisms across a latitudinal gradient in the fish, *Menidia menidia*. Evolution, 47, 487-494, 1993.

[LB84] Lloyd, D.G. & K.S. Bawa. Modification of the gender of seed plants in varying conditions. Evol. Biol., 16, 255-338, 1984.

[Lan85] Lande, R., Expected time for random genetic drift of a population between stable phenotypic states, Proc. Natl. Acad. Sci. USA 82:7641-7645, 1985.

[Mil80] R. Milner, **Calculus of Communicating Systems**, LNCS 92. Springer Verlag, 1980.

[Mil83] R. Milner, Calculi for Synchrony and Asynchrony, Theoretical Computer Science 25(3), pp 267-310, 1983.

[Mil90] R. Milner, **Communication and Concurrency**, Prentice Hall, 1990.

[MMS93] Moody, M.E., L.D. Mueller and D.E. Soltis, Genetic Variation and Random Drift in Autotetraploid Populations, Genetics 134:649-657, 1993.

[Mul64] Muller, H.J., The Relation of Recombination to Mutation Advantage, Mutation Research 1:2-9, 1964.

[Nag90] Nagylaki, T., Models and Approximations for Random Drift, Theor. Popul. Biol. 57:192-212, 1990.

[Nag93] Nagylaki, T., The Evolution of Multilocus Systems Under Weak Selection, Genetics 134: 627-647, 1993.

[NAG88a] Naylor, C., J. Adams & P. Greenwood. Population dynamics and adaptive sexual strategies in a brackish water crustacean, *Gammarus duebeni*. J. Anim. Ecol., 57, 493-507, 1988.

[NAG88b] Naylor, C., J. Adams & P. Greenwood. Variation in sex determination in natural populations of a shrimp. J. evol. Biol., 1, 355-368, 1988.

[NT80] Nei, M. and F. Tajima, Genetic Drift and Estimates of Effective Population Size, Genetics 98: 625-640, 1981.

[Nun93] Nunney, L., The Infleunce of Mating System and Overlapping Generations on Effective Population Size, Evolution 47(5):1329-1341, 1993.

[Nur74] Nur, U., The expected changes in the frequency of alleles affecting the sex ratio. Theor. Popul. Biol. 12, 49-85, 1974.

[Ohn67] Ohno, S., **Sex Chromosomes and Sex Linked Genes**, Springer Verlag Berline, 1967.

[Ohn79] Ohno, S., **Major Sex Determining Genes**, Berlin: Springer-Verlag, 1979

[Pag87] Page, D.C., R. Mosher, E. Simpson, E. Fisher, G. Mardon, J. Pollack, B. McGillivray, A de la Chapelle and L. Brown, The Sex-Determining Region of the Human Y Chromosome Encodes a Finger Protein, Cell 51:1091-1104, 1987.

[Poe88] Poethke, H.K., Sex Ratio Polymorphism: The Impact of Mutation and Drift on Evolution, Acta Biotheretica, 37:121-147, 1988.

[Ric86] Rice, W.R., On the Instability of Polygenic Sex Determination: the Effect of Sex Specific Selection, Evolution 40:633-9, 1986.

[RJ93] Rigaud, T. & Juchault, P. Conflict between feminizing sex ratio distorters and an autosomal masculinizing gene in the terrestrial isopod *Armadillidium vulgare* Latr. Genetics 133, 247-252, 1993.

[Sin88] Sinclair, A.H., J. Foster, J. Spencer, D. Page, M. Palmer, P. Goodfellow and J. Marshall Graves, Sequences Homologous to ZFY, a candidate human sex-determining gene, are autosomal in marsupials, Nature: 336:780-781.

[TW73] Trivers, R.L. & D.E. Willard. Natural selection of parental ability to vary the sex ratio of offspring. Science, 179, 90-92, 1973.

[TB90] Turelli, M., and N.H. Barton, Dynamics of Polygenic Characters Under Selection, Theor. Popul. Biol 38:1-57, 1990.

[THF92] Tofts,C., M.J. Hatcher, N. Franks, Autosynchronisation in *Leptothorax acervorum*; Theory, Testability and Experiment, Journal of Theoretical Biology 157: 71-82.

[Tof90a] Tofts, C., A Synchronous Calculus of Relative Frequency, CONCUR '90, Springer Verlag, LNCS 458, 1990.

[Tof90b] C. Tofts, The Autosynchronisation of *Leptothorax acervorum* (Fabricius) Described in WSCCS, LFCS-Report Number 128, Dept. Of Computer Science Edinburgh, 1991.

[Tof91] Tofts, C., Task Allocation in Monomorphic Ant Species, LFCS-Report Number 144, Dept. Of Computer Science Edinburgh, 1991.

[Tof93a] Tofts, C., Using Process Algebra to Describe Social Insect Behaviour, Transactions on Simulation, 1993.

[Tof93b] Tofts, C., A Study of Task Alocation in Monomorphic Ant Species: Theory, Bull. Math. Bio. 55(5):891-918, 1993.

[Tof94] Tofts, C., Processes with Probabilities, Priority and Time, Formal Aspects of Computer Science, 6(5):536-565, 1994.

[Tur84] Turelli, M. Heritable Genetic Variation via Mutation-Selection Balance: Lerch's Zeta Meets the Abdominal Bristle, Theor. Popul. Theo. 25:138-193, 1984.

[UB79a] Uyenoyama, M.K. & B.O. Bengtsson Towards a genetic theory for the evolution of the sex ratio. Genetics 93, 737-754, 1979

[UB79b] Uyenoyama, M.K. & B.O. Bengtsson Towards a genetic theory for the evolution of the sex ratio II. Haplodiploid and diploid models with sibling and parental control of the brood sex ratio and brood size. Theor. Popul. Bio. 20, 57-79, 1979.

[Ver65] Verner, J., Selection for Sex-ratio, Am. Nat. 99: 419, 1965.

[WA94] Watt, P.J. & J. Adams. Adaptive variation in sex determination in a crustacean, *Gammarus duebeni*. J. Zool., 232, 109-116, 1994.

[WC78] Werren, J.H. & E.L. Charnov, Facultative sex ratios and population dynamics. Nature 272, 349-350, 1978.

[Wer87] Werren, J.H. The coevolution of autosomal and cytoplasmic sex ratio factors. J. theor. Biol. 124, 313-334, 1987.

[WG90] Wagner, G. P., and W. Gabriel, Quantitative variation in Finite Parthenogentic Populations: What Stops Muller's Ratchet in the Absence of Recombination? Evolution 44(3):715-731

[Win34] Winge, O. The experimental alteration of sex chromosomes and vice versa as illustrated by *Lebistes*, Compt. Rend. Lab., Carlsburg. Ser. Physiol, 21:1-49, 1934.

[Wri39] Wright, S., The Distribution of Self-Fertility Alleles in Populations, Genetics 24:538-542, 1939.

[Wri68] Wright, S., **Evolution and The Genetics of Populations I**, Genetic and Biometric Foundations, The University of Chicago Press, 1968.

[Wri69] Wright, S., **Evolution and The Genetics of Populations II**, The Theory of Gene Frequency, The University of Chicago Press, 1969.

[Yam55] Yamamoto, T., Genetics 40:406, 1955.

[Zim91] Zimmerman, J.K. Ecological correlates of labile sex expression in the orchid *Catasetum viridiflavum*. Ecology, 72, 597-608, 1991.