

University of Groningen

Evolutionary genetics and dynamics of transitions in sex determination

Schenkel, Martijn

DOI:
[10.33612/diss.166344703](https://doi.org/10.33612/diss.166344703)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2021

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
Schenkel, M. (2021). *Evolutionary genetics and dynamics of transitions in sex determination*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen. <https://doi.org/10.33612/diss.166344703>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

CHAPTER IV:

DIVERGENT EVOLUTION OF GENETIC SEX DETERMINATION MECHANISMS ALONG ENVIRONMENTAL GRADIENTS

Martijn A. Schenkel
Jean-Christophe Billeter
Leo W. Beukeboom*
Ido Pen*

* These authors contributed equally to this work.

Abstract

Sex determination (SD) is an essential developmental process, but its molecular regulation is very diverse, both between and within species. SD mechanisms are traditionally categorized as either genetic (GSD) or environmental (ESD), depending on the type of cue that triggers sexual differentiation. However, mixed systems, with both genetic and environmental components, are more prevalent than previously thought. Here, we show theoretically that weak environmental effects on expression levels of genes within sex determination regulatory mechanisms can trigger within-species evolutionary divergence of SD mechanisms. This may lead to their stable coexistence and a polymorphic system along spatial environmental gradients. We apply our model to the SD system of the housefly, a species which has global latitudinal clines in the relative frequencies of XY and ZW systems. Our model predicts such patterns by assuming that specific genes in the housefly SD system have temperature-dependent expression levels. We conclude that environmental sensitivity of gene regulatory networks may play an important role in diversification of SD mechanisms.

Introduction

Sex determination (SD) is a crucial aspect of the development of sexually-reproducing organisms, yet the regulatory mechanisms directing SD are very diverse and prone to evolutionary change (Bachtrog et al., 2014; Beukeboom & Perrin, 2014). SD mechanisms are traditionally classified as either environmental sex determination (ESD) or genetic sex determination (GSD) depending on what type of cue initiates the determination of an individual's sex. In both ESD and GSD, the initial cue directs sexual differentiation by activating genes for male- or female-specific development and physiology (Ge et al., 2018). Under ESD, such cues include temperature, salinity, acidity, or the social environment during a sensitive phase in embryonic development (reviewed in Devlin & Nagahama, 2002; Janzen & Paukstis, 1991). In GSD systems, the causal agents are genetic differences between males and females, such as the male-determining *SRY* gene in mammals (Berta et al., 1990; Sinclair et al., 1990; Goodfellow & Lovell-Badge, 1993). Various chromosomal systems of SD are known, such as male heterogamety (females XX, males XY) and female heterogamety (females ZW, males ZZ), as well as systems without sex-specific chromosomes, such as haplodiploidy (Beukeboom & Perrin, 2014). Even in species that have similar chromosomal SD systems, *e.g.* male heterogamety, the genes that induce maleness can be very different. Turnover events, *i.e.* the evolution of a novel SD mechanism, are prevalent in some taxa but not others, indicating differential lability of SD systems. All eutherian mammals have male heterogamety whereas in fish and insects SD mechanisms vary between species and even between populations of some species. How this variety in SD systems has evolved is not yet well understood.

ESD and GSD could be considered as two extremes of a spectrum, rather than being mutually exclusive (Pen et al., 2010; Uller & Helanterä, 2011; Beukeboom & Perrin, 2014). Although pure ESD and GSD systems do exist, mixed systems occur in several organismal groups, such as amphibians, fish and insects (Doums et al., 1996; Devlin & Nagahama, 2002; Nigro et al., 2007; Pen et al., 2010; Hamm et al., 2015; Ma et al., 2016). In such systems, SD genes may bias the process of sex determination towards the male or female state, and environmental cues may counteract these genetic effects (Alho et al., 2010; Holleley et al., 2015). Furthermore, GSD has repeatedly evolved from ESD (Ezaz et al., 2009; Pen et al.,

2010; Muralidhar & Veller, 2018) and various evolutionary models have been proposed to explain such transitions (Ezaz et al., 2009; Pen et al., 2010; Muralidhar & Veller, 2018). For instance, changes in environmental conditions, like global warming, may bias population sex ratios and cause natural selection to favour genetic modifiers that re-establish balanced sex ratios (Fisher, 1930; Wilkins, 1995). Previous models have invoked selection caused by biased sex ratios, but have not considered the underlying molecular mechanisms of biasing development towards males or females. Even within GSD systems, genes involved in SD may be affected by environmental conditions, resulting in perturbations in the SD process. This may promote evolution of the GSD cascade by causing e.g. disturbed sex ratios, or via other yet unknown effects. Here, we model how environmental influences on the expression levels of genes involved in sex determination may affect the evolutionary dynamics of sex determination.

The process of sex determination involves a hierarchical regulation of various genes (Wilkins, 1995; Pomiankowski et al., 2004; Kopp, 2012). An initial cue targets a small number of regulatory genes that in turn regulate further downstream targets. Evolutionary transitions between GSD systems are thought to occur either by the displacement of the initial cue gene by another gene with a similar function or by the recruitment of a new regulatory gene on top of the ancestral SD cascade (Bull & Charnov, 1977; Wilkins, 1995; Pomiankowski et al., 2004). Downstream components may be more constrained in terms of evolutionary change due to their pleiotropic role in development but are not per se fully constrained (Herpin et al., 2013). The persistence of multifactorial GSD (where multiple genes may control sex) is still hard to explain as such systems are considered a transitory state towards monogenic SD (where a single gene controls sex) (Rice, 1986). Here, monogenic SD may be more robust as the possibility that individuals develop as intersexes is considered to be negligible, whereas under polygenic SD interference between different SD genes may result in failure to commit development to either the male or female state. A new evolutionary framework is therefore necessary which integrates the emerging views that (1) SD is not solely environmentally or genetically determined but instead can be affected by both types of cues; and (2) change in SD cascades does not occur only at the top but may also occur via changes in the underlying genetic network.

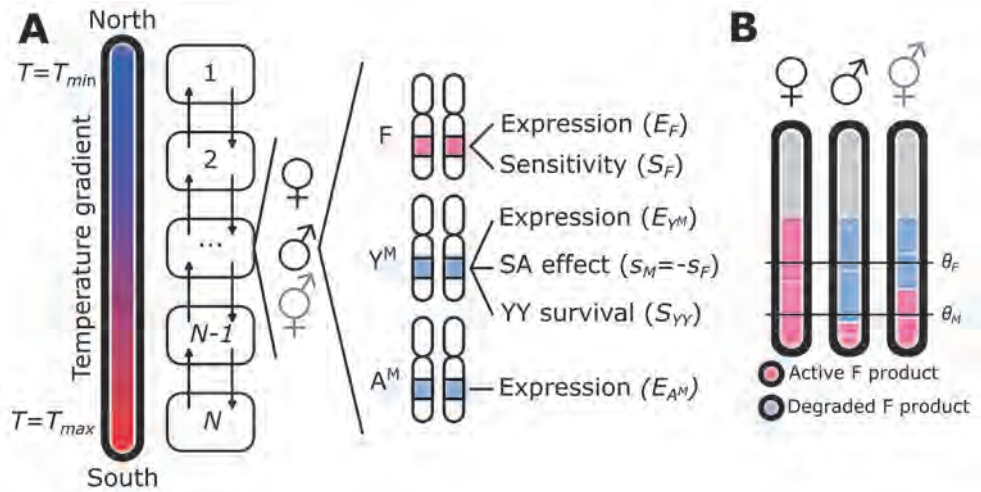


Figure 1: Model overview and hypothesis. (A) Demes 1 through N are arranged along a linear gradient where T increases from T_{min} to T_{max} . Each deme contains a variable number of females, males, and intersexes. Reproduction occurs by mating between males and females within the same deme; intersexes do not reproduce. Dispersal occurs at a rate $d/2$ to each neighbouring deme (indicated by arrows between demes). Each individual carries three loci F , Y^M , and A^M that together determine sex. (B) Sex is determined based on the net total active F product. Active F product is derived from the F locus, and is degraded by M derived from Y^M and/or A^M in individuals where these genes are expressed. Temperature positively affects the expression level of F . If the net total active F product exceeds θ_F , individuals become females, whereas if it falls below θ_M individuals become males; otherwise, individuals develop into infertile intersexes.

Here, we develop a theoretical model to study the evolution of polymorphic SD systems with spatial variation in environmental conditions affecting expression of sex determination genes. Our model is inspired by the polymorphic SD system of the housefly *Musca domestica* and assumes temperature as the environmental cue as variation in SD systems was found to be correlated with temperature in natural populations, and has been found to affect SD in some SD mutant strains (see also Box 1). Nonetheless the model has broader relevance to other species and environmental cues. Sex is determined by the activity of a temperature-sensitive feminizing gene F , which can be inhibited by one or more M gene(s) to induce maleness (Figure 1B). We investigate here (1) which variants of F can evolve under which conditions and how these variants yield stable SD systems (in combination with M), (2) when and how transitions from an XY to a ZW system via evolution of a dominant female-determining F variant which is insensitive to M can occur and (3)

how fitness effects associated with an M-factor affects the spread of this dominant female-determining F variant as well as other M-factors. We use our model to show how the multifactorial SD system of houseflies can evolve (Box 1).

Results & Discussion

Model overview

We modelled a metapopulation consisting of a linear array of subpopulations (demes) arranged along a temperature gradient, i.e. from north to south, with dispersal between neighbouring demes (Figure 1A). Sex is determined by the activity of a gene F , which induces feminization when it exceeds a threshold θ_F , whereas maleness is achieved by inhibition of F activity below a second and lower threshold θ_M . When F activity falls between these thresholds, individuals develop into infertile intersexes (Figure 1B). Inhibition of F is achieved by M, which is generated either by the Y-chromosomal locus Y^M and/or the autosomal A^M locus. All three loci F , Y^M , and A^M are unlinked to each other. We incorporate two equivalent M-producing genes because M-factor polymorphism is one of the key sources of variation in *M. domestica* SD (discussed in Box 1), and to allow us to study when different types of SD transitions (XY-XY as well as XY-ZW transitions) can occur.

Each F allele has a certain genetically determined expression level E_F and a sensitivity to inhibition by M S_F . Expression of F is also positively affected by temperature. Thus, the activity A of a single F allele is given by:

$$A = \max(0, (E_F(1 + \beta_T T) + \varepsilon)(1 - S_F M)) \quad (1)$$

Here, E_F and S_F denote the genetic expression level and sensitivity of the F allele; $1 + \beta_T T$ denotes the rate at which temperature increases F expression, with T indicating normalized local temperature (0 in first or most northern deme, 1 in last or most southern deme); ε indicates normally-distributed noise in F expression reflecting environmental and developmental noise; and finally M represents the amount of M product generated by Y^M and/or A^M . The term $E_F(1 + \beta_T T) + \varepsilon$ denotes gross F expression, whereas $1 - S_F M$ indicates what proportion of F product is retained after breakdown by M. The combined activity of both F alleles determines

sex as described above. All traits, that is, F expression levels and sensitivity as well as Y^M and A^M expression levels, can undergo mutations. Mutations that affect F expression do not affect sensitivity and vice versa. A trait value of 0 corresponds to an unexpressed state or full insensitivity to M; we exclude traits with a value of 0 from undergoing mutation so that evolution to such a state reflects a loss-of-function without back mutation. An exception applies to A^M ; we assume that no A^M alleles are initially present, but that A^M may evolve *de novo* at a rate μ_D . We assume a standard XY male heterogametic sex determination mechanism as the starting point, where females carry two X-chromosomes lacking Y^M alleles, and males carry an X-chromosome without and a Y-chromosome with an active Y^M allele. We incorporate fitness effects in the form of sexually antagonistic genetic variation (beneficial to males with additive fitness effect s_M , detrimental to females with additive fitness effect $s_F = -s_M$), which affects fitness via mating success and/or a reduced survival s_{YY} of YY homozygotes. When $\mu_D = 0$, no A^M expression can evolve and Y^M is always maintained as the male-determining gene. When $\mu_D > 0$, A^M displaces Y^M as the male-determining gene in absence of fitness effects associated with Y^M , which can inhibit SD transitions (Bull & Charnov, 1977; van Doorn, 2014), in which case Y^M may be retained either locally or globally (see details below).

F activity relative to sex determination thresholds determines which SD systems are evolutionary stable

The evolvability of F expression and sensitivity enables the evolution of different types of F alleles. In this section, we first verbally explore the model described above to identify the different systems that are able to produce an equal sex ratio; we do so by first considering how loss of expression and/or sensitivity affects the realized F activity in presence and absence of M. We then determine how different combinations of F and M genotypes (where M indicates either Y^M or A^M) form recurrent pairs, i.e. genotype combinations that result in the same genotypes when crossed with each other (Bull & Charnov, 1977), that result in a female-specific and a male-specific genotype. Loss of expression can result in non-functional F alleles, and may enable a transition to a system where presence/absence of one or more expressed F alleles determines sex, rather presence/absence of M (e.g. females F/O ; $+/+$, males O/O ; $+/+$, where F indicates expressed F and O unexpressed F , and $+$ indicates the absence of M). Loss of sensitivity allows for F to become a dominant

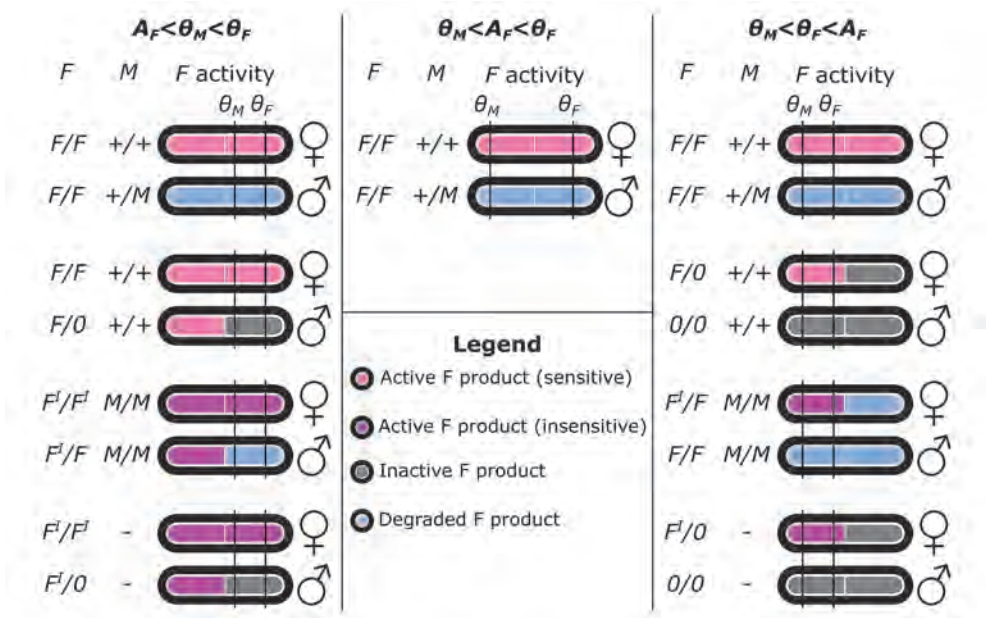


Figure 2: Potential sex determination systems depending on threshold values of F activity. Each system is defined by a recurrent pair of female and male genotypes that can only generate those two genotypes (conform Bull & Charnov (1977)). Here we define three alleles for locus F : a regular F that is expressed and sensitive to M ; a variant F^I that is insensitive to M and expressed; and a variant 0 that is unexpressed. For M , we distinguish between active (M) and inactive ($+$) variants. We use A to refer to the activity of a single F allele. In systems with both F^I and 0 , M has no function and may be present or absent; this is indicated by a dash (-).

female-determining allele F^I and the evolution of an SD system where its presence always leads to feminization and its absence is required for masculinization. How F alleles may evolve, and how this can lead to changes in the SD mechanisms is largely determined by the activity of a single F allele relative to the two SD thresholds θ_F and θ_M (Figure 2). Evolution of F also affects the dynamics of M ; when F comes to be unexpressed, M is no longer required to induce masculinization and may be lost. Alternatively, when F evolves to be insensitive to M , M alleles may accumulate as it can now be transmitted through males and females alike. The possibility for F to evolve thus enables a wide variety of different SD systems to arise from a rudimentary SD cascade.

The assumed initial SD system where a Y-chromosomal Y^M acts as a dominant male-determining gene over F , results in equal numbers of males and females

regardless of the relationship between F activity and the SD thresholds θ_F and θ_M . Moreover, male heterogamety with Y^M or A^M as a dominant male-determiner is the only system which can yield equal numbers of males and females when $\theta_M < A < \theta_F$. Temperature-dependent expression of F in our model can result in a change of magnitude of A relative to the thresholds θ_M and θ_F , i.e. this relationship may change from $A < \theta_M < \theta_F$ to $\theta_M < A < \theta_F$, or to $\theta_M < \theta_F < A$. In both cases, in some parts of the population the relationship $\theta_M < A < \theta_F$ applies so that some form of male heterogamety involving Y^M or A^M is expected to occur in these parts. Under other conditions however, stable SD systems can be formed with other variants of F . When the relationship between F activity A and the SD thresholds θ_M and θ_F changes, this can trigger a transition in SD. This could then result in local transitions in SD, so that SD mechanisms become differentiated between populations. It should be noted that a shift in the relationship between F activity and the SD thresholds does not necessarily cause SD transitions, unless such a shift has a direct impact on the sex ratio. However, following such shifts transitions may occur as a result of other factors capable of driving SD transitions now that alternative SD mechanisms are mechanistically feasible. For example, the evolution of an insensitive F variant is disfavoured when $\theta_M < A < \theta_F$, but is made possible when environmental conditions change so that $\theta_M < \theta_F < A$. Even if this change does not affect the sex ratio, other mechanisms of SD turnover (e.g. sexually antagonistic selection on a loci linked to F ; van Doorn & Kirkpatrick (2010)) may nonetheless apply to drive the invasion of an insensitive F variant.

Following evolution of an insensitive variant of F (denoted F') and fixation of M (further discussed below), two possible further evolutionary changes become possible (Supplementary Figure 1). First, under these conditions, the product of the regular F allele is always broken down and hence these alleles serve no further function in the SD process. This renders them selectively neutral and thereby prone to accumulation of loss-of-function mutations that result in pseudogenisation. Once expressed, sensitive F alleles are purged and individuals either carry F' , whose product is insensitive to M , or unexpressed F alleles that do not generate any F product. Male development occurs in the latter case and is due to a lack of F product rather than by F breakdown mediated by M . This in turn enables the second change: because M is no longer required for sex determination, M may degrade via similar mechanisms as initially occurred for F . Sex is then determined solely by the genotype

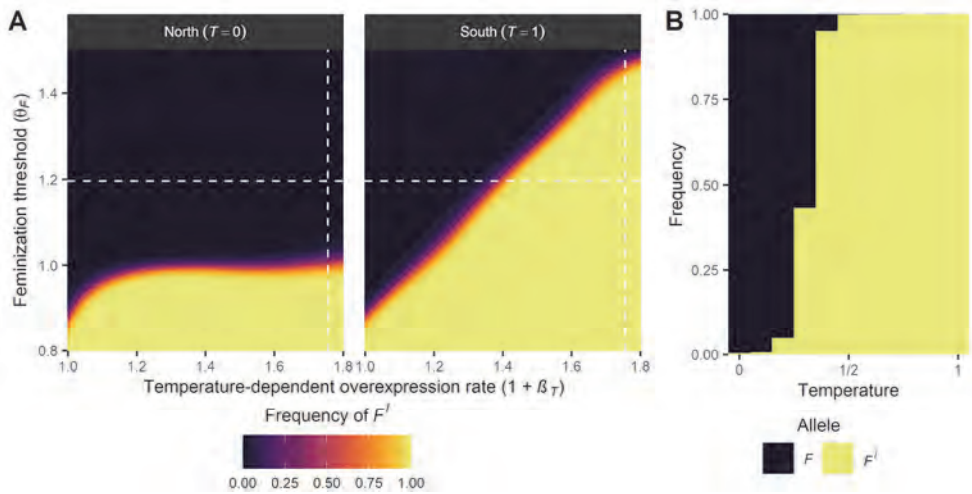


Figure 3: Conditions for spread of a dominant female-determining gene F^I . (A) Equilibrium frequency of F^I at the northern and southern end (first/last deme) of the population range; local temperatures are indicated in brackets (parameter values: $\theta_M = 0.2$; $\mu_D = 0.001$). In the northern deme, temperature-dependent overexpression of F is absent whereas overexpression is maximal in the southern deme. (B) Between these two extremes, the equilibrium frequency of F^I increases along the temperature gradient (shown here for $\theta_F = 1.20$; $\theta_M = 0.2$; $1 + \beta_T T = 1.76$; $\mu_D = 0.014$). Depicted in A and B are the frequency of the F^I allele at the maternally-inherited locus in females. White dashed lines in A indicate the parameter values for the simulation results depicted in B (vertical line: $1 + \beta_T$; horizontal line: θ_F).

at the F locus, though which genotypes induce femaleness or maleness depends on the relationship between potential activity of an F allele A and the feminization threshold θ_F (see Figure 2 for details). Note however that though F and M may through such changes become selectively neutral with regard to sex determination in our model, in natural systems they may still perform other functions and therefore be under selection through pleiotropic effects, preventing mutational decay. Alternatively, expression of F and/or M may be costly to maintain, in which case selection will actively favour genetic decay.

Evolution of F insensitivity and establishment of XY-ZW gradients

The evolution of F insensitivity can also enable a transition from male heterogamety to female heterogamety. This can occur via the evolution of a dominant feminizing allele of F and subsequent fixation of Y^M or A^M in both sexes. Dominance of F is achieved by becoming insensitive to M , though only when $A > \theta_F$ (conform Figure

2). By increasing the value of A , temperature-dependent overexpression of F could then enable a (local) transition to female heterogamety. To test this hypothesis, we used simulations in which we varied the rate at which temperature affects F expression ($1 + \beta_T$) and the value for feminization θ_F . We found that F^I could spread in the entire population when the feminization threshold $\theta_F < 1$ (Figure 3A) regardless of temperature as it was detected in the northern-most ($T = 0$, left panel) as well as the southern-most deme ($T = 1$, right panel). When $\theta_F > 1$, F^I was unable to spread to colder demes as it would result in intersexual development (Supplementary Figure 2), but could be detected in warmer demes depending on how strongly temperature affected F expression (i.e. the value of $1 + \beta_T$). Under these conditions, we find that F^I follows a frequency gradient where it is absent in northern i.e. colder demes, but increases in frequency as temperatures increase, i.e. towards the south (Figure 3B).

These results underline that the distribution of F^I is constrained by F potential expression level, and show that temperature-dependent effects on expression can establish gradients when temperature varies. Due to its feminizing effect, an F^I allele is transmitted as if it were a W-chromosome, and occurs only in F^I/F females or intersexes. In presence of M , the F^I product is not broken down but the product of regular sensitive F alleles is degraded, so that regular F alleles do not contribute to the total F product. This scenario becomes increasingly probable as either Y^M and/or A^M frequency increases. Therefore, feminization of developing embryos under these conditions must be achieved solely by the activity of the F^I allele. Because F^I is insensitive to M , the total F activity is determined by its gross expression, i.e. its genetic expression rate (E_F) multiplied by the temperature-dependent overexpression rate ($1 + \beta_T T$; see Equation 1). When no temperature-dependent overexpression occurs, feminization is only achieved when the genetic expression level exceeds the feminization threshold, i.e. $E_F > \theta_F$, but in presence of temperature effects is achieved when $E_F(1 + \beta_T T) > \theta_F$. This explains why when $\theta_F < 1$, F^I can spread everywhere independent of temperature (Figure 3A, left panel). but when $\theta_F > 1$, F^I frequencies depend on the temperature-dependent overexpression rate (Figure 3A, right panel) and the temperature (Figure 3B).

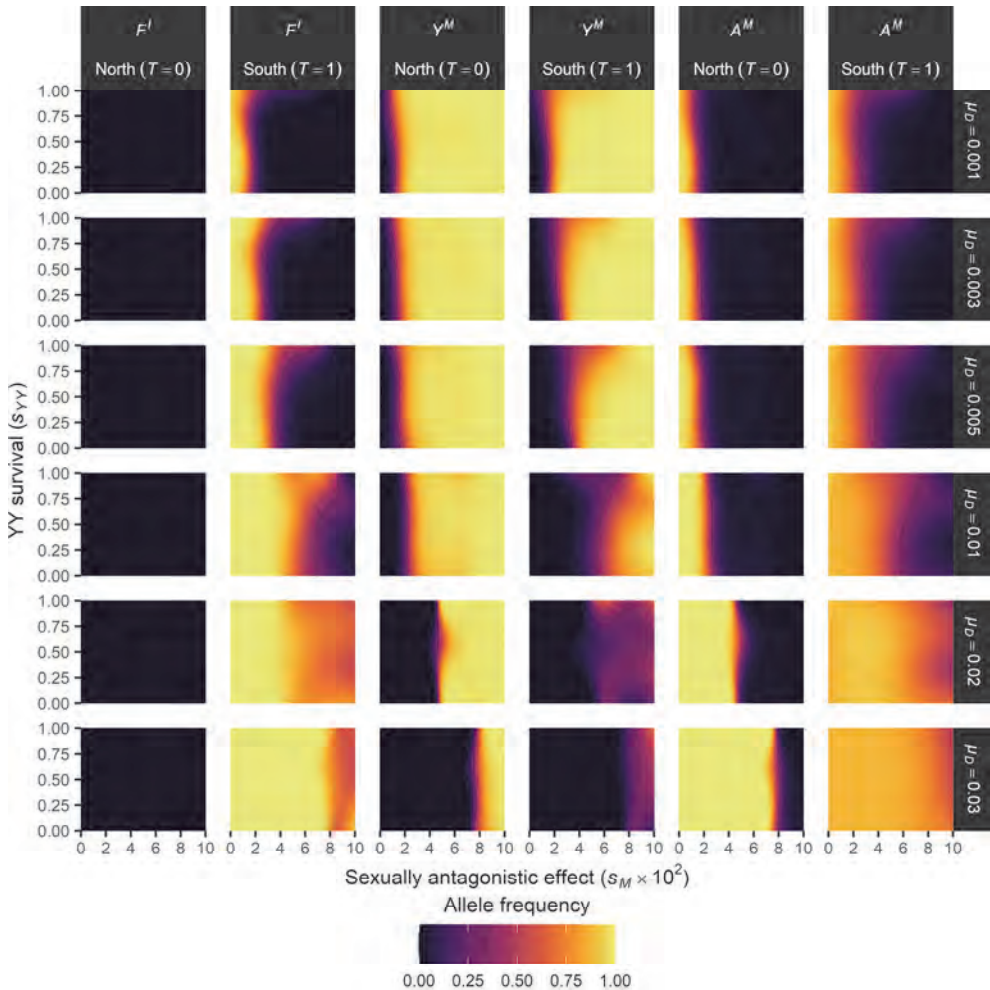


Figure 4: Y-chromosomal fitness effects alter the scope for SD transitions. Shown are the predicted equilibrium frequencies of F^I in females (maternally-inherited alleles), Y^M and A^M in males (paternally-inherited alleles). Horizontal strips indicate locus and temperature, vertical strips the A^M activation rate μ_D . Further parameter values used in simulations: $\theta_F = 1.2$; $\theta_M = 0.3$; $\beta_T = 0.5$. Predicted allele frequencies were smoothed with generalized additive models.

Y-chromosomal fitness effects modulate the conditions under which SD turnover can occur and can enable complex SD polymorphisms

In the simulations above, we have assumed that Y^M is not associated with any fitness effects, thereby not affecting the spread of F^I . Under these conditions, we find that Y^M is lost and replaced by A^M owing to novel A^M alleles being formed through

mutation. Y^M instead is lost as it is not favoured by selection nor recurrently formed through mutation. In the presence of F^I , A^M is found at high frequencies on both paternally- and maternally-inherited chromosomes and functions as a co-factor for maleness as it is necessary to break down the product derived from the regularly-sensitive and expressed F allele in non- F^I -bearing individuals to induce maleness. In the absence of F^I , A^M functions as a dominant male-determining gene and is thus found in a heterozygous state in males but is absent in females. However, Y^M being the ancestral SD gene may have induced the chromosome on which it is located to undergo Y-chromosome evolution (reviewed in Bachtrög, 2013; Schenkel & Beukeboom, 2016). If so, the Y^M -bearing chromosome is expected to become enriched for sexually antagonistic genetic variants as well as recessive deleterious mutations. In effect, this would cause Y^M to be associated with higher fitness in heterozygous X/Y^M males, but have a fitness cost to Y^M -bearing females (who also carry F) as well as all homozygous Y^M/Y^M individuals. Both sexually antagonistic genetic variation and cost of homozygosity can in theory strongly affect evolutionary transitions in SD (reviewed in van Doorn, 2014). We therefore performed additional simulations where we considered a sexually antagonistic fitness effect of the Y-chromosome (see Supplementary Table 1), causing the Y-chromosome to have a positive fitness effect in males but a negative fitness effect in females; these effects are effectuated during mating, where an individual's probability of mating is proportional to its fitness relative to that of other same-sex individuals. In addition to sexually antagonistic fitness effects which are mediated during the mating phase, we include costs of Y^M/Y^M homozygosity as a form of viability selection during embryogenesis.

When Y^M is associated with SA genetic variation, we find that selection acts to maintain it over A^M (Figure 4). SA effects can also inhibit invasion of F^I provided that the strength of SA selection is sufficiently high. This occurs as negative fitness effects of Y^M on female carriers reduces the fitness of F^I/F females relative to non- Y^M -carrying F/F females. However, when the rate of introduction of novel A^M alleles is sufficiently high, F^I can always invade even if the selective effects associated with Y^M are strong. Under these conditions, A^M evolves more readily and its overrepresentation in the gene pool is more pronounced (similar to Y-chromosomal meiotic drive; Jaenike (2001); Kozielska et al. (2010)). This causes a more strongly male-biased sex ratio among the offspring and hence the selective benefit for F^I as

a female-determining factor increases. When the sex ratio is sufficiently male-biased, F^I -bearing individuals (females) benefit derived from sex ratio selection outweighs the costs of bearing the female-detrimental Y-chromosome.

The fitness effect of the Y-chromosome can enable maintenance of both Y^M and A^M in the population, albeit in different locations. In case such polymorphism persists, Y^M tends to be disfavoured in the presence of F and instead A^M is found in both sexes. In the absence of F , Y^M is restricted to males and maintained in heterozygous state so that only the male-beneficial effect through sexually antagonistic selection is exposed to selection, thereby promoting its maintenance over A^M as the male-determining gene. The Y^M/A^M polymorphism is highly similar to the distribution of Y-chromosomal versus autosomal M-factors in the housefly *M. domestica*. In this species, autosomal M-factors are more prevalent at lower latitudes and coincide with a dominant feminizing allele tra^D , resulting in three latitudinal gradients in Y-chromosomal M-factors, autosomal M-factors, and the tra^D allele. We find that Y-chromosomal fitness effects enable the evolution of this complex system in our model (see Box 1). This provides an adaptive explanation for the evolution of this system in contrast to existing models of SD evolution, which have been unable to predict when stable polymorphisms for *tra* versus tra^D along with Y-chromosomal M-factors versus autosomal M-factors may occur in general, and in particular when these coincide along environmental clines as observed under natural conditions.

Box 1: Evolution of the polymorphic housefly system

Our model has been inspired by the common housefly *Musca domestica*, in which wild populations harbour different SD systems (reviewed in (Hamm et al., 2015)). Here, we discuss how our model may explain the evolution of this system. In houseflies, like many insects, sex is determined by a linear cascade of genes, two of which vary between populations. First, *transformer* (*tra*) induces female development when active in developing embryos (Hediger et al., 2010); activity depends on pre-mRNA splicing, a process that is sensitive to temperature as well as other stressors (Palusa et al., 2007). Male development is achieved by inhibition of the *tra* loop by switching the splicing of *tra* pre-mRNA to a male-specific variant, which encodes a truncated and presumably non-functional

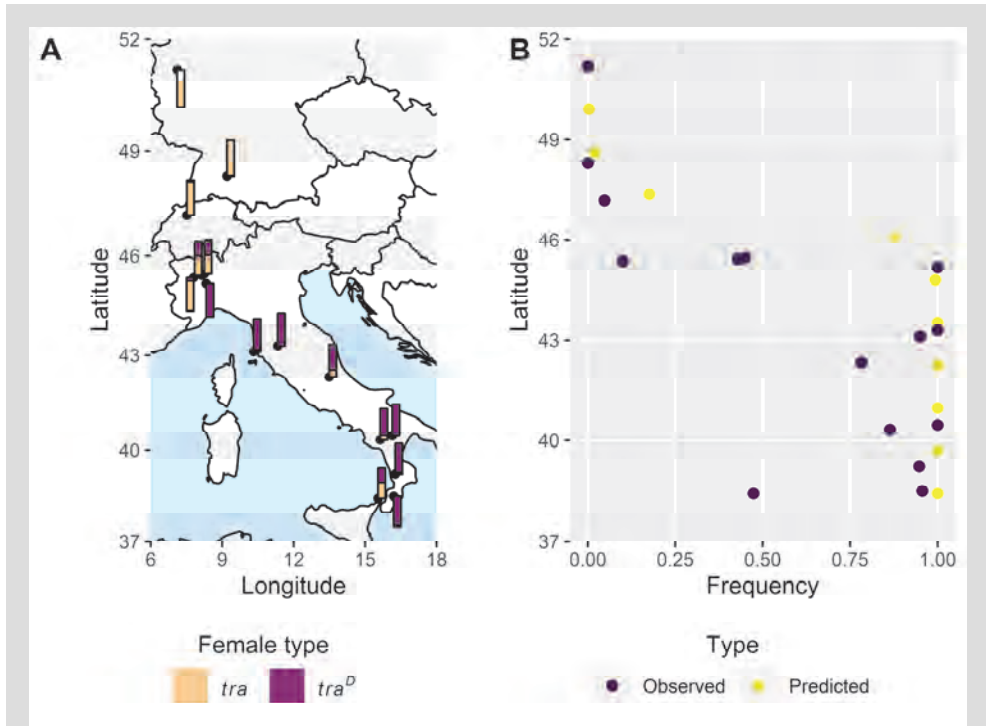


Figure 5: Model predictions vs. the observed latitudinal frequency gradient of a female-determining gene in the housefly *Musca domestica*. (A) The frequency of *tra/tra* (light orange) and *tra^D/tra* (purple) females in Europe. Adapted from (Kozielska et al., 2008). (B) Observed (purple) and predicted (yellow) frequencies of *tra^D*-bearing females (deme position scaled to match observed latitude range). Parameter values used: $\beta_F = 1.5$; $\theta_F = 1.15$; $\mu_D = 0.005$; $s_M = -s_F = 0.05$; $S_{YY} = 0.9$; $d = 0.1$.

protein; this inhibition is achieved by a masculinizing factor (M-factor) such as *Mdmd* (Sharma et al., 2017). A variant of *tra*, *tra^D*, has been found to be insensitive to inhibition by M-factors and instead induces female development in all carriers regardless of whether they carry any M-factors. Second, M-factors in *M. domestica* are found on different chromosomes in different populations. Most of these M-factors are homologous to *Mdmd* and likely represent duplicated copies (Sharma et al., 2017), but the M-factor on autosome I remains unidentified. In our model, F represents *tra* and likewise *tra^D* corresponds to the dominant F^D as discussed previously; Y^M and A^M of our model both represent M-factors.

High-latitude *M. domestica* populations have a standard male heterogamety (XY) system in which the Y-chromosomal *Mdmd* gene induces maleness, and all individuals carry two regularly-sensitive *tra* alleles. At lower latitudes, however, females carry the insensitive *tra^D* allele, and both sexes can be homozygous for an autosomal copy of *Mdmd*, so that these populations have a female heterogamety (ZW) system (Figure 5; (Kozielska et al., 2008; Hamm et al., 2015)). The transition between these two SD systems is gradual, so that gradients exist in the frequency of Y-chromosomal *Mdmd* (decreasing with lower latitude), autosomal *Mdmd*, and *tra^D* (both increasing with lower latitude). Temperature likely plays a causal role in maintaining these gradients by affecting the SD process, though the exact mechanism for this is still undetermined (Feldmeyer et al., 2008; Feldmeyer, 2009). However, temperature effects on housefly SD have been reported in the form of biased sex ratios produced in wildtype crosses (Feldmeyer, 2009) as well as in females carrying the *masculinizer* (*man*) mutation, another variant of *tra* (Schmidt, Hediger, Nöthiger, et al., 1997; Hediger et al., 2010). The *man* mutation represents a maternal-effect male-determining gene, where *man*-carrying females may produce male progeny even if the progeny do not carry an M-factor. Although these females can produce all-male broods, the effect is incomplete and temperature-sensitive (Schmidt, Hediger, Nöthiger, et al., 1997), with offspring sex ratios more male-biased at higher temperatures. Altogether, temperature seems to have an important influence on SD in *M. domestica*, but the underlying mechanisms are not yet fully understood.

The housefly system with its different SD systems is represented in our model by gradients in the frequency of Y^M (decreasing with temperature), A^M , and F (both increasing with temperature). Presumably, *tra^D* is limited to warmer localities for similar reasons as F in our model, i.e. because a single *tra^D* allele may not be sufficient to induce feminization at low temperatures. Y-chromosomal *Mdmd* and autosomal *Mdmd* may likewise follow similar dynamics as Y^M and A^M . As described in the main text, Y-chromosomal fitness effects can yield gradients in Y^M and A^M , but may also prevent the spread of F , particularly at low rates at which novel A^M alleles enter the population. The evolution of a

polymorphic SD system like that of houseflies therefore depends on a balance between the Y-chromosomal fitness effects and the rate at which new autosomal

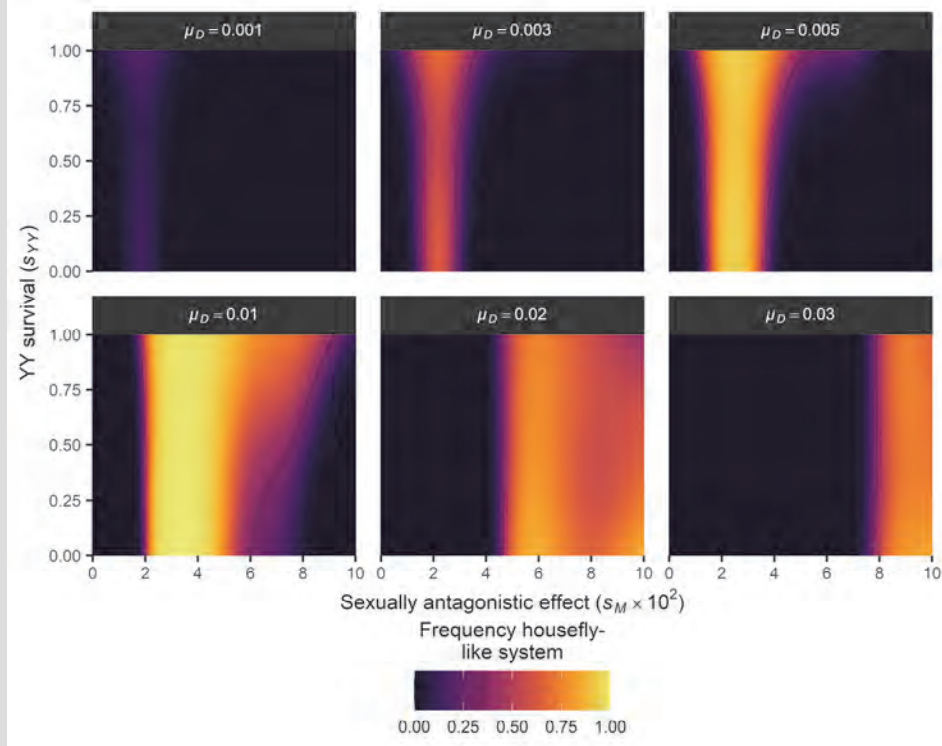


Figure 6: Evolution of a housefly-like system. A housefly-like system is defined by Y^M being the major allele at $T = 0$ but the minor allele at $T = 1$ (in males, paternally-inherited allele), and vice versa for A^M (in males, paternally-inherited alleles) and F (in females, maternally-inherited alleles). Frequency here denotes the expected frequency of observing a housefly-like system at equilibrium in the model. Parameter values: $\theta_F = 1.2$; $\theta_M = 0.3$; $1 + \beta_T = 1.5$. Values in the strips above each graph denote the value of the A^M activation rate μ_D . Simulations were scored as exhibiting a housefly-like system as described above (total $N = 10000$ simulations). To obtain these predicted scores, we fitted a binomially-distributed generalized additive model.

Mdmd arises. In our model, we find that a housefly-like system can evolve under various rates of A^M *de novo* evolution (Figure 6). Higher rates of A^M evolution require stronger SA effects for Y^M to be maintained in lower temperature demes. Costs of YY homozygosity do not appear essential for Y^M to be lost in presence of F , though they may increase the likelihood of Y^M being lost in favour of A^M in

the presence of F^I by reducing the fitness of YY homozygotes. Possibly, the negative effect Y^M has in females due to sexually antagonistic selection may be sufficient to promote the loss of Y^M .

Our model can explain the evolution of a complex variable SD system like that of the housefly *M. domestica*. Based on these results, we propose a novel hypothesis for the evolution of the housefly system in which the sex determination cascade is shaped by a combination of environmental influences on *tra*, recurrent evolution of autosomal *Mdmd*, and fitness effects associated with the Y-chromosomal copy of *Mdmd* (Figure 7). This hypothesis generates a number of testable predictions. First, in houseflies we expect that the feminizing function of *tra* may be elevated at higher temperatures, allowing for a single functional *tra* allele (such as *tra^D*) to be sufficient for feminization. The inverse is true at lower temperatures, meaning that both *tra* alleles must be functional for feminization. A *tra^D/tra; M/M* genotype then should yield a female phenotype at high temperatures, but a male or intersexual phenotype at low temperatures. The same is true for a *tra/tra⁰; +/+* genotype, where *tra⁰* is a loss-of-function variant such as *tra^{man}* (Schmidt, Hediger, Nöthiger, et al., 1997; Hediger et al., 2010) or an artificially-generated knockout and *+* denotes the absence of active M-factors. Temperature-dependent effects on *tra* can then be elucidated by rearing individuals with such genotypes at different temperatures. Although it is still unclear how exactly *tra* maintains its own activity on a molecular level, the influence of temperature-dependent effects must be effectuated during sex determination which occurs during early embryonic development (Hediger et al., 2010; Sharma et al., 2017). Further studies on molecular *tra* function and the effect of temperature during this stage on individual sex determination may be mutually informative. The second prediction is that *Mdmd* must be transmitted (or at some point have been transmitted) at a super-Mendelian rate so that a male-biased sex ratio is established. In houseflies, evolution of autosomal *Mdmd* is likely to occur or have occurred via transposition of *Mdmd* (Green, 1980; Sharma et al., 2017). Relative to *de novo* evolution, this mechanism may result in much higher evolutionary rates and consequently stronger male-biased sex ratios. Other mechanisms of *Mdmd* evolution such as meiotic drive should in

principle have the same effect (Kozielska et al., 2010), but do not appear to occur in *M. domestica*. Third, Y-chromosome bearing males must have a fitness

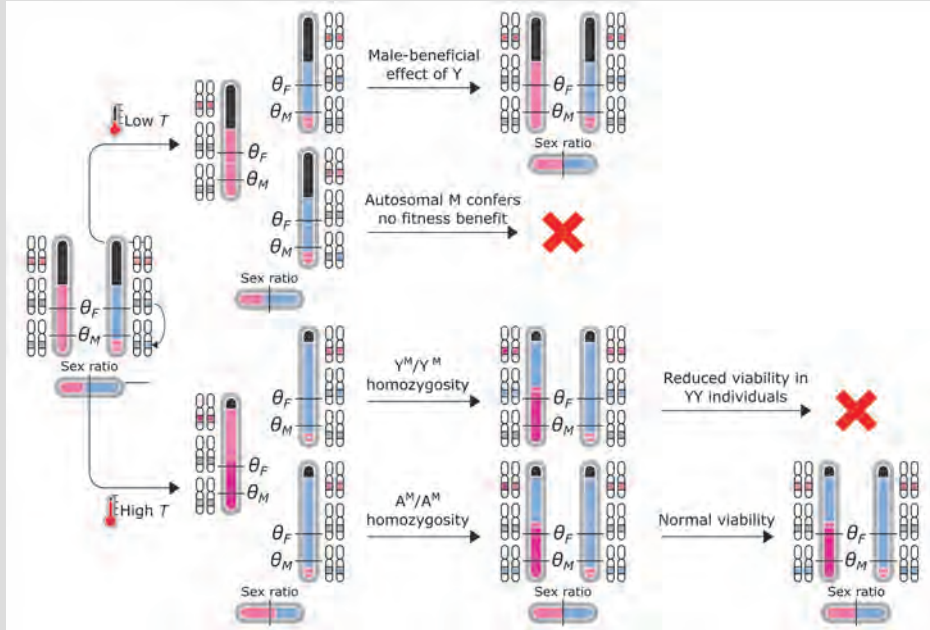


Figure 7: A novel hypothesis for the evolution of the housefly polymorphic sex determination mechanism. Evolution of A^M (here represented by transposition of Y^M) results in an excess of M factors in the population and a male-biased sex ratio. At low temperatures, the M-insensitive P^D allele cannot evolve, and both Y^M and A^M persist in a heterozygous state. Because Y^M is associated with a fitness benefit in males, Y^M -bearing males outcompete A^M -bearing males, resulting in a return to the ancestral state with Y^M as the sole male-determining allele in a XY system. In contrast, at high temperatures, the P^D allele can evolve, and has a fitness benefit as a result of sex ratio selection. As P^D spreads, M alleles can be transmitted by females resulting in the formation of homozygous Y^M/Y^M and A^M/A^M individuals. As Y^M homozygosity is associated with a viability cost, these individuals have lower fitness than A^M/A^M individuals. This results in a loss of the Y^M allele and fixation of the A^M allele in its place as a co-factor for male development. In effect, a transition has occurred from XY male heterogamety to ZW female heterogamety as the sex-determining role has been taken over by P^D .

benefit over XX males with autosomal M-factors in absence of F^D , but must have a fitness cost in presence of F^D . We assume here that the fitness benefit is conferred by sexually antagonistic genetic variants on the Y-chromosome and that the fitness cost is due to presence of recessive deleterious mutations. However, as discussed above, other scenarios may also be possible, provided

that they yield these fitness effects in absence and presence of F . Recent work found no evidence to suggest that temperature affected the expression of genes linked to either a Y-chromosomal M-factor or an autosomal M-factor in adult males (Adhikari et al. in prep), suggesting that these chromosomes are not simply exposed to temperature-dependent selective pressures; instead, selection may affect these M-factors through other mechanisms. Finally, for a housefly-like system to evolve, a balance must exist between these processes. Rather than being particularly dependent on a single parameter, whether or not the housefly system may have evolved via adaptive evolution as described under our hypothesis hinges on the manner in which these different processes interact in this species.

Evolution of polymorphic genetic sex determination can be driven by environmental heterogeneity

Here, we have presented a model for the evolution of SD systems in a context where sex is determined by genetic factors in combination with environmental effects. In our model, female development is induced when the activity of a feminizing gene F exceeds a certain threshold, whereas male development is achieved when F activity is below a different and lower threshold. This can be caused by inhibition of F activity by the maleness-promoting gene(s) Y^M and/or A^M , or by loss of F expression. We incorporate a positive effect of temperature on the expression of a feminizing locus F . We find that several different SD systems can be formed depending on the activity of an F allele relative to the threshold values for masculinization and feminization. Temperature-dependent effects on F may alter the relationships between F activity and SD thresholds, thereby enabling the evolution of different genetic SD systems. A particular prediction is the transition from male heterogamety to female heterogamety, which occurs in our model via the evolution of an insensitive dominant feminizing variant (F) that induces femaleness even in the presence of Y^M and/or A^M , which normally inhibit F , resulting in maleness. F can spread when activity of a single F or F allele is sufficient to induce feminization; when activity is modulated by temperature, this can lead to local invasions of F and subsequently

differentiation between populations along temperature gradients. Differentiation can also occur for Y^M and A^M , with Y^M being favoured in absence of F and A^M in presence of F . This occurs when Y^M is associated with certain fitness effects such due to linkage with sexually antagonistic variants or recessive deleterious mutations. Altogether, we show that this can lead to multiple gradients in SD genes as found in *M. domestica*.

Conclusion

Our model can easily be amended to other systems provided that they have a basic GSD framework which is influenced by an environmental effect. Environmental effects on genetic sex determination systems are being reported in an increasing number of species. Although temperature-dependent effects are well-documented, other environmental effects may also influence SD in certain systems such as hormonal imbalance in fish due to pollution (Devlin & Nagahama, 2002). Likewise, other components and assumptions of the model that are based on the housefly system may be represented differently in other species but with similar effects. For example, the impact of A^M evolution is not due to a specific mechanism of mutation, but more generally by causing a male-biased sex ratio, thereby promoting the invasion of F . Male-biased sex ratios occur due to A^M overrepresentation in the gene pool via its *de novo* evolution, but the same effect can be achieved by e.g. translocation of a Y-chromosomal male-determining gene or via association with meiotic drivers (Green, 1980; Kozielska et al., 2010). Additionally, we see that in absence of an association between Y^M and certain fitness effects, A^M replaces Y^M altogether, yet still drives the invasion of F , showing that our model does not strictly require a third locus. Inversely, it is likely that a more complex genetic basis generates similar evolutionary patterns, e.g. when various genes can evolve into a male-determining gene, possibly via different mechanisms (Bopp, 2010). In this scenario, many genes having a small chance to evolve into a male-determining gene may have the same net effect on sex ratio as a gene that is prone to evolving a masculinizing function. In this light, it will be interesting to test whether genes that have acquired a sex-determining function in one species are prone to evolve a similar function in a related species, where it has no sex-determining function.

Previous work has shed light on the evolution of ESD and GSD systems, and when transitions between these two may arise (Pen et al., 2010; Muralidhar & Veller, 2018; Schwanz et al., 2020). However, our understanding of the evolution of

polymorphic SD systems and the potential for environmental heterogeneity to influence this process remains limited. Our results highlight environmental effects on GSD systems, and under which conditions this can lead to within-species polymorphism in GSD systems. Moreover, even in systems that appear to be fully GSD, the role of environmental influences on the SD processes must not be ignored as these may have played an important role in their evolution. In extension of this, changes in environment, e.g. due to global warming, may impose yet unforeseen selective pressures on species when GSD systems evolved under different conditions.

Acknowledgements

We thank the Center for Information Technology of the University of Groningen for providing access to the Peregrine high-performance computing cluster; Pina Brinker, Fangying Chen, and Peter Hoitinga for feedback on an earlier version of the manuscript; and Babak Arani and the Evolutionary Genetics cluster for fruitful discussion.

Methods

An overview of all parameters and their standard values are provided in Supplementary Table 2.

Life cycle

We simulate a population consisting of N demes arranged along a linear gradient (Figure 1A); we vary the environmental cue T normalized from 0 in the first deme to 1 in the last deme. T positively affects the expression of a feminizing locus F and may thereby increase the probability of an individual developing as a female (for details see section "Sex determination"). In each deme, we generate a fixed number of K individuals upon initiation. Individuals have a diploid genome consisting of three linkage groups. One linkage group carries the F locus, from which the F product that induces feminization is derived. The two remaining linkage groups can carry a male-determining M locus, whose protein M degrades the F product. We designate one of

these linkage groups as being the original sex chromosome pair and accordingly refer to these as XY chromosome pair; we refer to the M-locus on this chromosome pair as Y^M . The other linkage group is referred to as the autosomal chromosome pair, and we refer to its M-locus as A^M . Sex is determined by the total amount of F product available after interacting with M. We assume a male heterogametic system in which females carry two X-chromosomes that lack an active M allele, and males carry one such X-chromosome and a Y-chromosome that harbours an active M allele. All individuals lack active M alleles on their autosomes upon initiation. Reproduction occurs by random mating between males and females within each deme, during which all linkage groups follow Mendelian segregation. During reproduction, a total of K juveniles are created within each deme. Each allele can undergo mutation with a certain probability. Juveniles then disperse at a rate d to a random neighbouring deme. After dispersal, all adults die and are replaced by the juveniles. For all simulations, a “burn-in” period of 20,000 generations is applied during which all demes have a $T = 0$. After this, we increase T in each deme to its final value (as determined by their position in the gradient) during a warmup period of 10,000 generations. We use a burn-in and warmup phase to ensure that the system is able to evolve to a stable state prior to incorporating environmental effects and to ensure selection due to environmental effects does not change abruptly. After increasing the warmup phase, we keep conditions stable for 200,000 generations to allow the system to evolve to a new equilibrium. At this point, we analyse the genotypes of all individuals to determine which SD systems have evolved in which demes.

Sex determination

An individual's sex is determined by the net activity F_{tot} of the two alleles at the F locus, which is determined based on the initial expression level of each F allele minus the amount of F product that is degraded by M (see Figure 1B). Initially, each allele produces an amount E_F of F product which has a sensitivity S_F to breakdown by M. Environmental effects on SD are included solely with regard to F expression, so that the total expression A (prior to eventual breakdown by M) of an F allele is given by:

$$A = E_F (1 + \beta_T T) \quad (2)$$

Here, β_T refers to the rate at which F expression increases between $T = 0$ and $T = 1$. To model within-deme heterogeneity in T and developmental noise, we add some noise to the expression of F by adding a Gaussian amount ϵ with $\mu = 0$ and $\sigma = \sigma_F$. The cumulative amount M_{tot} of M protein is determined by summing up the expression level E_M of all active M alleles. Breakdown of F protein by M is determined per F allele by the expression level and the sensitivity of the expressed product as follows, so that the net activity A' of an F allele is given by:

$$A' = \max(0, (E_F(1 + \beta_T T) + \epsilon)(1 - S_F \times M_{tot})) \quad (3)$$

We sum up the net activity of both alleles of F to obtain the net activity of the F locus, F_{tot} , based on which sex is determined:

$$F_{tot} = \sum_{i=1}^2 A'_i \quad (4)$$

in which i is used to indicate the maternal ($i = 1$) and paternal ($i = 2$ allele). Individuals develop into males if $F_{tot} < \theta_M$ or into females if $F_{tot} > \theta_F$. If $\theta_M \leq F_{tot} \leq \theta_F$ individuals develop into infertile intersexes. The expression and sensitivity of F , as well as the expression of M , can vary from 0 to 1.

Reproduction and mutation

Reproduction occurs by mating between females and males residing in the same; an individual's probability of being sampled as a mate is proportional to its fitness relative to other same-sex individuals and depends solely on sex chromosome genotype (see "Fitness effects of the Y-chromosome"); in absence of Y-chromosomal fitness effects mating effectively occurs between randomly-sampled males and females. Mutations can occur within each gamete, and occur independently for F sensitivity, F expression, and M expression each with a trait-dependent probability μ ; mutations occur prior to sex determination in developing offspring so that this is functionally equivalent to mutations occurring during gametogenesis. Mutations may either result in a certain shift in trait value (regular mutations) or in a loss-of-function type mutation where the trait value is set to zero (null mutation). When mutations occur, they have a trait-dependent probability μ_{null} of being a null mutation. Regular

mutations result in a change in a trait value by summing the current trait value with a value sampled from normal distributions with mean zero and standard deviations σ_{F_E} , σ_{F_S} , and σ_{M_E} for F sensitivity, F expression, and M expression. Once a trait value equals zero through either regular or null mutations, we consider this to be a loss of function and prevent the trait from undergoing further mutation so that there is no gain of function. The exception to this is the *de novo* evolution of the A^M locus. New (expressed) A^M alleles may arise *de novo* with a frequency μ_D . The expression level of newly-evolved A^M locus is sampled from a normal distribution with mean μ_E and standard deviation σ_E .

Fitness effects of the Y-chromosome

We incorporate two possible fitness effects that are associated with Y^M . First, we incorporate a sexually antagonistic fitness effect of the Y-chromosome with additive fitness effects s_F and s_M in females and males (see Supplementary Table 1). This causes the Y-chromosome to have a positive fitness effect in males but a negative fitness effect in females, and resembles a scenario in which the Y-chromosomal M-locus is tightly linked to one or more sexually antagonistic alleles. These fitness effects are effectuated during the mating phase of our model; rather than sampling one male and one female at random, each of these are sampled with a likelihood that is proportional to their fitness score relative to other same-sex individuals in their deme. Second, we incorporate a fully recessive viability cost of Y^M/Y^M homozygosity, whereby developing YY individuals survive to the adult stage with a probability $s_{YY} \leq 1$. Survival of X/X and X/Y^M individuals is not affected. This resembles the effect of mutation accumulation on the sex-limited chromosome. For both fitness effects, we assume that Y^M is fully linked to the loci under selection so that no recombination occurs.

Data analysis

We categorize all F alleles based on their sensitivity level and their expression level, where an F allele is considered insensitive if its sensitivity is equal to 0 (and sensitive if it is larger than 0), and is unexpressed when its expression is less than $\theta_M/2$. We classify all alleles that are insensitive and expressed as F alleles, and all others as F

alleles. We categorize all M alleles based on their expression level; we consider an M allele unexpressed when its expression is lower than $(2 - \theta_F)/2$, and expressed if its expression equals or exceeds $(2 - \theta_F)/2$. The threshold for F expression is based on the assumption that an expression value below $\theta_M/2$ is insufficient to prevent maleness when an individual is homozygous. Similarly, an M allele may break down a small amount of F product without preventing female development. Assuming F is fully expressed (and hence a total of 2 F product is generated), breakdown by M can at most be $2 - \theta_F$. When both F alleles are fully sensitive, this means that an individual can be homozygous for an M that breaks down at most $(2 - \theta_F)/2$ and still develop into a female.

Following categorization, for each deme we count the number of F and F' as well as the number of unexpressed and expressed M alleles on the maternally-inherited and paternally-inherited alleles. Of this, we take two separate subsets for the allele counts at the two environmental extremes $T = 0$ and $T = 1$. We fit generalized additive models (GAMs) with binomial distributions on the allele frequencies at both extremes for each locus. For the F locus, we did so on the allele frequency of the F^I allele on the maternally-inherited copy and for both Y^M and A^M we fit GAMs on the frequency of expressed alleles on the paternally-inherited copies. In each GAM, we included a full tensor product smooth between the parameters of interest. Thin plate regression splines were used as the base functions for each parameter.

To assess whether a polymorphic system resembling that found in natural housefly populations had evolved, we determined for each simulation whether (1) F' was the minor allele at $T = 0$ and the major allele at $T = 1$; (2) Y^M was the major allele at $T = 0$ and the minor allele at $T = 1$; and (3) A^M was the minor allele at $T = 0$ and the major allele at $T = 1$. Minor (major) alleles are defined as having a frequency below (over) $1/2$. Simulations that met all three criteria were considered to have evolved a housefly-like SD system. The resulting scores were used to fit a binomially-distributed GAM with a full tensor product smooth between A^M *de novo* rate (μ_D), sexually-antagonistic fitness effect (s_M ; all simulations assume $s_F = -s_M$), and YY survival rates (s_{YY}).

All data analysis was carried out in R (v.4.0.0; R Development Core Team (2020)) and RStudio (v.1.2.5033; RStudio Team, 2020), using the 'cowplot' (Wilke, 2019), "maps" (Becker et al., 2018), "mapsproj" (McIlroy et al., 2020), 'mgcv'

(Wood, 2017), 'tidyverse' (Wickham et al., 2019), and 'viridis' (Garnier, 2018) packages.

Supplementary Material*Supplementary Tables*

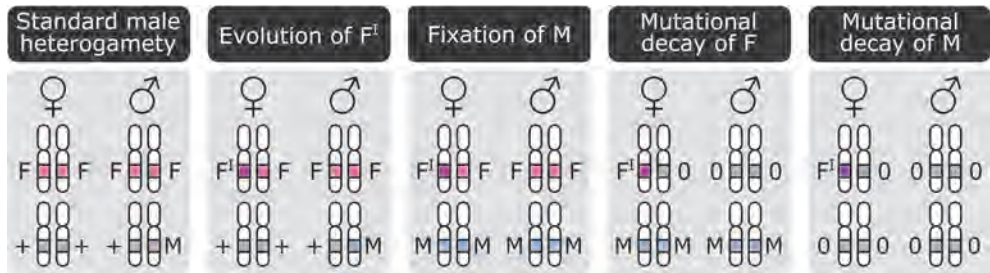
Supplementary Table 1: Sexually antagonistic fitness effects of the Y-chromosome. s_F and s_M denote the fitness effect of the Y-chromosome in females and males, whereas h_F and h_M denote the dominance of these fitness effects in XY heterozygotes. We assume the Y-chromosome to be deleterious to females ($s_F < 0$) but beneficial to males ($s_M > 0$).

	XX	XY	YY
Female	1	$1 + h_F s_F$	$1 + s_F$
Male	1	$1 + h_M s_M$	$1 + s_M$

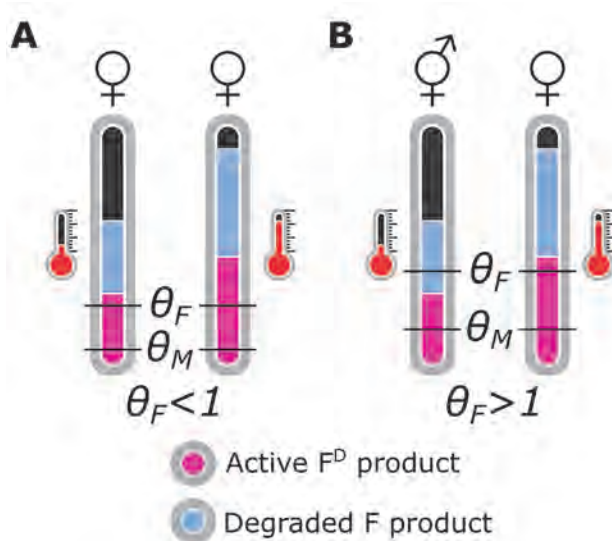
Supplementary Table 2: Parameters in the model and their standard values.

Variable	Description	Standard value
N	Number of demes	11
K	Population size per deme	10,000
d	Rate at which dispersal occurs	0.1
θ_F	Lower threshold for female development	1.2
θ_M	Upper threshold for male development	0.3
s_{YY}	Survival rate of YY individuals	1
h_F	Dominance of Y-chromosomal fitness effect in females	0.5
h_M	Dominance of Y-chromosomal fitness effect in males	0.5
s_F	Selective effect of Y-chromosome in females	0
s_M	Selective effect of Y-chromosome in males	0
μ_{null}	Proportion of mutations that are null mutations	0.01
μ_E	Mutation rate for F expression level	0.1
σ_{FE}	SD for mutation effect for F expression level	0.1
$\mu_{E_{null}}$	Proportion of mutations in F expression that are null mutations	μ_{null}
μ_S	Mutation rate for F sensitivity	0.1
σ_{FS}	SD for mutation effect for F sensitivity	0.1
$\mu_{S_{null}}$	Proportion of mutations in F sensitivity that are null mutations	μ_{null}
μ_M	Mutation rate for M expression level	0.1
σ_{ME}	SD for mutation effect for M expression	0.1
$\mu_{M_{null}}$	Proportion of mutations in M expression that are null mutations	0
μ_D	Mutation rate for <i>de novo</i> A^M evolution	0
β_E	Correlation coefficient between temperature and F expression	0.5
μ_E	Mean expression level for A^M when evolving <i>de novo</i>	0.9
σ_E	Standard deviation for A^M expression level when evolving <i>de novo</i>	0.05
σ_F	Standard deviation for noise in F expression level	0.05

Supplementary Figures



Supplementary Figure 1: Genetic decay of F and M following the evolution of an insensitive feminizing F^I allele. Starting from a standard male heterogametic system, a feminizing F^I allele can invade resulting in a transition to a female heterogamety system. For this to occur, an M allele must be fixed in both sexes. Fixation of M ensures the regularly-sensitive F product is always degraded and hence these alleles confer no function leaving them susceptible to genetic decay by mutation accumulation. This can eventually lead to loss of functional F in favour of non-expressed F alleles (denoted O). Loss of F obviates the need for M to be maintained as no regular F product is generated that must be broken down to ensure maleness in non- F^I -bearing individuals (genetic males). Similar to F previously, M functionality is no longer necessary and mutations may accumulate by which M becomes unexpressed (also denoted O). Note that F^I is depicted here as a dominant feminizing allele (females F^I/F , males F^I/F) but similar scenarios apply for a recessive feminizing F^I allele (females F^I/F^I , males F^I/F). The only difference is that the F allele that is susceptible to decay is now only found in males in a heterozygous state rather than in heterozygous females and homozygous males; decay of F and M can occur according to the same principles as when F^I is a dominant feminizing allele.



Supplementary Figure 2: Constraints on the spread of a dominant feminizing F allele. (A) When $\theta_F < 1$, F can invade in the entire population as a single F allele generates sufficient product to induce female development. (B) When $\theta_F > 1$, F cannot spread in absence of temperature-dependent overexpression as it does not generate sufficient product to induce femaleness, and instead intersexual development is induced in F -bearing individuals. At high temperatures, F overexpression enables a single F allele to be sufficient for feminization, allowing for F to spread. In all cases, we assume a $F/F; M/M$ genotype so that regular F product is degraded.

