University of Groningen

Evolutionary genetics and dynamics of transitions in sex determination

Schenkel, Martijn

DOI:
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.

Download date: 01-06-2023

CHAPTER V:

ESTABLISHMENT OF BABY-SEX CHROMOSOMES USING INTERPOPULATION CROSSES IN THE HOUSEFLY *MUSCA DOMESTICA*

Martijn A. Schenkel
Leo W. Beukeboom
Ido Pen

Abstract

The evolution of sex chromosomes has been studied extensively from a theoretical perspective, but empirical confirmation of the resulting models has been lacking. This is largely due to lack of suitable model systems. Empirical efforts have sought to exploit interspecific variation in sex chromosome systems and by seeking out younger sex chromosomes to study sex chromosome evolution prior to the onset of extensive differentiation between the minor (Y or W) and major (X or Z) chromosome. Here, I propose an complementary system for empirical research on sex chromosome evolution. The housefly (*Musca domestica*) features a multifactorial sex determination mechanism wherein ’standard’ populations have a XY system with a Y-chromosomal male-determining gene *Mdmd*. Other populations feature a ZW system wherein *Mdmd* resides on an autosome, and females carry a dominant feminizing gene *tra*^{D} that overrules *Mdmd*. Interpopulation crosses were used to introgress the autosomal *Mdmd* gene into a standard population. This creates a strain in which the *Mdmd*-carrying autosome turns into an incipient proto-Y-chromosome, and its complement into an incipient proto-X-chromosome. These 'baby-sex chromosomes' can be used to study the genomic changes of sex chromosomes in real time and in a repeatable manner. We discuss possible applications of baby-sex chromosomes for future studies such as testing the 'sexual antagonism' hypothesis of sex chromosome evolution.

Introduction

The evolution of sex determination (SD) mechanisms and of sex chromosomes have been thoroughly studied in a theoretical context (van Doorn, 2014; Abbott et al., 2017). With the advent of genomics and other genetic tools, empirical research on these topics has been steadily increasing, resulting in the identification of a wide variety of SD genes (Hall et al., 2015; Krzywinska et al., 2016; Sharma et al., 2017; Pan et al., 2019). They have also detected evolutionary strata, i.e. regions exhibiting different levels of differentiation, on the sex chromosomes of several species (Handley et al., 2004; Wright et al., 2014), suggesting sex chromosome differentiation takes place in distinct bouts. These studies have confirmed some of the predictions generated by the theoretical models, but have been unable to address others. Most of these studies have been made use of existing interspecific variation in SD systems and sex chromosomes to study the stages of sex chromosome evolution (e.g. (Zhou & Bachtrog, 2012a; b)). A complementary approach would be to develop model systems in which SD and sex chromosome evolution can be studied; this has two main benefits. First, many of the theoretical predictions on SD and sex chromosome evolution cannot be tested within a single comparative framework. This is due to a difference in approach in how these two topics are studied, as generally studies on sex chromosome evolution rely on analysing past patterns of evolution, whereas studies on SD transitions focus on tracking differentiation in real time (e.g. (Hamm & Scott, 2008)]). Second, the rigor with which theories are tested is insufficient as there is an intrinsic lack of repeatability, leading to a lack in statistical power. An analysis of a given naturally-occurring sex chromosome system provides an *N*=1 result of its realized evolution, and is therefore inadequate to make extensive inferences about the selective processes involved (Kawecki et al., 2012). Our understanding of sex chromosome evolution is further hindered by the degenerated nature of late-stage sex chromosomes. This erases signs of past adaptive evolution and impedes the use of genomic tools to study alterations in gene content and regulation, which are still often inapplicable to degenerated and repetitive regions which often make up a large proportion of late-stage Y-/W-chromosomes (but see Mahajan et al., 2018). Our understanding of the early phases of sex chromosome evolution therefore remains largely speculative.

The evolution of sex chromosomes is thought to be initiated by an autosome acquiring a sex-determining function, e.g. by the evolution of a new SD gene (Ohno, 1967). The sex-limited chromosome (Y in XY systems, W in ZW systems) which carries the new SD gene is termed the minor sex chromosome, and its non-sex-limited counterpart (X in XY, Z in ZW) the major sex chromosome. Y- and W-chromosomes are transmitted exclusively through one sex (males for Y, females for W), but evolve in a similar manner, and likewise X- and Z-chromosomes undergo similar evolutionary changes. The region linked to the SD gene becomes sex-linked and is transmitted through one sex only, e.g. males in case of a male-determining gene on a novel Y-chromosome. The "sexual antagonism" hypothesis (Rice, 1987a) has become the canonical model for sex chromosome evolution, but empirical support for it is ambiguous. This hypothesis postulates that the presence of a sex-determining allele on a chromosome causes the region linked to that allele to be exclusively transmitted through one sex, e.g. males for a male-determining allele on a Y chromosome. This transmission pattern promotes the spread of genetic variants that are beneficial to this sex, but harmful to the other, i.e. intralocus (IASC) genetic variants (see also Chapter 2). Extending the linked region by the evolution of recombination-suppressing mechanisms (e.g. inversions) allows for more of these genetic variants to be linked to the sex-determining allele. Stepwise evolution of recombination suppression along the Y-/W-chromosome has been inferred from the presence of distinct strata of genetic divergence between the sex chromosome pairs (X- *versus* Y-chromosome and Z- *versus* W-chromosome) in several species (Lahn & Page, 1999; Handley et al., 2004; Wright et al., 2014). However, it is unclear whether the establishment of these strata occurred in conjunction with the spread of specific IASC genetic variants and, if this were the case, whether the newly-non-recombining region was enriched for such variants prior to or after the evolution of suppressed recombination. This is due to the degenerative nature of the Y-chromosome, which as a non-recombining genomic entity is subject to accumulation mutation via Muller's ratchet and suboptimal adaptive evolution as a result of various types of Hill-Robertson interference (Bachtrog, 2013). These processes can result in erosion of past signs of adaptation and an accumulation of deleterious mutations that makes ultimate and proximate analyses of the gene content of late-stage Y-chromosomes a cumbersome or even impossible endeavour (Graves, 2006; Schenkel & Beukeboom, 2016).

To bypass the issue of Y-chromosome degeneration, studies on Y-chromosome evolution have made use of naturally-occurring, recently-evolved sex chromosomes such as in a variety of *Drosophila* species (Kaiser et al., 2011; Zhou & Bachtrog, 2012b; a). Although these approaches resolve the issue of Y-chromosomal degeneration, they are complicated by the fact that comparisons between Y-chromosomes of differing age or stage of evolution are interspecific. Similarly, despite having a clear benefit in terms of studying naturally-evolved Y-chromosomes, they do not permit repeatability because they only represent a single instance of past sex chromosome evolution, and no additional replicates may be generated. An alternative approach to study sex chromosome evolution is to generate new sex chromosomes by mimicking the evolution of a new SD gene on an autosome. This can be achieved by genetic modification, such as the insertion of existing SD genes on an autosome. This is, however, a technically challenging procedure requiring genetic engineering. A more feasible approach is to exploit existing variation in SD mechanisms in species with multifactorial SD, such as in the housefly *Musca domestica*, such that genes that are capable of acting as master SD genes, but do not do so owing to the presence of other dominant SD genes, are crossed into novel genetic backgrounds where they become the master SD gene. In this chapter, I briefly outline the different SD mechanisms that occur naturally in *M. domestica* (see also Chapter I) and explain how this variation can be exploited to establish novel sex chromosome systems using classical genetic crosses. I refer to these as baby-sex chromosomes, i.e. the chromosome carrying the male-determining factor as the baby-Y chromosome and its complimentary copy as the baby-X chromosome. Finally, I discuss how the baby-sex chromosome strains and associated controls can be used to test theories of sex chromosome evolution, such as the 'sexual antagonism' model.

Genetics of housefly sex determination

The housefly *Musca domestica* L. (Diptera: Muscidae) has a multifactorial sex determination system in which different genes can determine sex and different chromosomes can be sex chromosomes depending on the geographic origin of housefly populations (reviewed in (Feldmeyer et al., 2008; Hamm et al., 2015)). The genetic cascade directing sexual differentiation is similar to that of many other

insects in that *transformer* (*tra*, sometimes denoted *F*) activity determines whether an individual becomes male or female (Bopp, 2010; Hediger et al., 2010; Bopp et al., 2014). In female embryos, *tra* is active, whereas in male embryos, the activity of *tra* is shut down by a male-determining (M) factor such as the *Mdmd* gene (Sharma et al., 2017). In populations at high latitudes (northern hemisphere), males carry a Y-chromosome with two copies of *Mdmd* (Hediger et al., 1998), whereas females carry two X-chromosomes that lack this gene, resulting in an XY sex chromosome system. At lower latitudes, however, a dominant *tra* allele (*tra^D*) is found, which remains active even in the presence of *Mdmd* and/or other (yet unidentified) M-factors (Wagoner, 1969; McDonald et al., 1978; Hediger et al., 2010). As a result, all carriers of this allele are female. In populations harbouring the *tra^D* allele, M-factors are commonly found on autosomes, and individuals are often homozygous for an M-factor. M-factors can reside on all autosomes and both sex chromosomes (reviewed in (Hamm et al., 2015)), and in most cases correspond to duplicated copies of the *Mdmd* gene (Sharma et al., 2017; Li *et al.,* in prep) with the exception of the M-factor on autosome I, which has been shown to be an unknown but different gene. Effectively, in populations at low latitudes sex is determined by the presence or absence of the *tra^D* allele and the M-factor has become a co-factor for male development instead of a master SD gene, and therefore these populations have a ZW sex chromosome system.

In contrast to the Y-chromosomal *Mdmd* copy at high latitudes, the autosomal M-factors at low latitudes are not transmitted solely through males but also through *tra^D*-carrying females; the chromosomes carrying them have therefore not yet developed into mature Y-chromosomes. Despite the masculinizing function of an autosomal M-factor, the fact that it can be passed through *tra^D*-carrying females means the region linked to the M-factor is not exclusively transmitted through males. Consequently, sexually antagonistic mutations are less likely to become associated with the M-factor (Jordan & Charlesworth, 2012; Charlesworth et al., 2014). By extension, the evolution of recombination suppression is not expected to be favoured (Rice, 1987a). However, these M-carrying autosomes can develop into Y-chromosomes when they are introduced into a population without *tra^D* alleles (Figure 1A). The autosomal M-factor is turned into a dominant male-determining allele as carriers of the autosomal M-factor will now always develop as males. This procedure effectively mimics the evolution of a novel sex-determining gene and establishes a

male-limited region linked to the autosomal M-factor. This region can undergo enrichment for IASC loci and the evolution of recombination arrest as predicted by the "sexual antagonism" model of sex chromosome evolution. Autosomes carrying M-factors therefore have the potential to develop as Y-chromosomes, but are sheltered from this evolutionary fate by the presence of *tra^D*; in its absence, they can undergo sex chromosome evolution.

Interpopulation crosses can establish new baby-sex chromosomes

I applied the principles of introgressing autosomal M-factors from a population with *tra^D* (ITA3; males *tra/tra; X/X; M/M*; females *tra^D/tra; X/X; M/M*) into a background without it (GM, originating from GK3 through GK8; males *tra/tra; X/Y; +/+*; females *tra/tra; X/X; +/+*) as shown in Figure 1A; the introgression procedure is outlined below and shown in Figure 1B. Detailed methods are described under "Introgression of autosomal M-factors". In brief, I established a genetically variable strain GM (Gerkesklooster Mix) with XX females and XY^M males by mixing flies from strains GK3 through GK8 (collected in Gerkesklooster, The Netherlands; for further details see Supplementary Table 1). I introgressed an autosomal M-factor from the ITA3 strain into this GM background to dissociate the M-factor from *tra^D,* thereby turning it into a baby-Y-chromosome. This also increases the genetic variation within the strain, which may have become depleted in the original strain from which the autosomal M-factor has been derived. Five-hundred virgin GM females were paired with 500 ITA3 homozygous M/M males in a single cage to establish an F0 population of 1000 individuals with equal sex ratio. The resulting F1 offspring were sexed, 500 males were collected and paired with 500 new virgin GM females. This procedure was repeated until the F10 generation resulting in replacement of the ITA3 genetic background by the GM background for theoretically 99.9%. At this point cages were set up using both female and male offspring to establish the introgressed population featuring a baby-sex chromosome system. Using this approach, I established two experimental strains EA and EB with autosome III as a baby-sex chromosome pair; control strains CG and CI were established by repeatedly backcrossing GM males with GM females (CG) or ITA3 males with ITA3 females (CI) (see also Table 1). The GM strain was newly established every generation during introgression to prevent depleting the genetic variation in this strain over time.

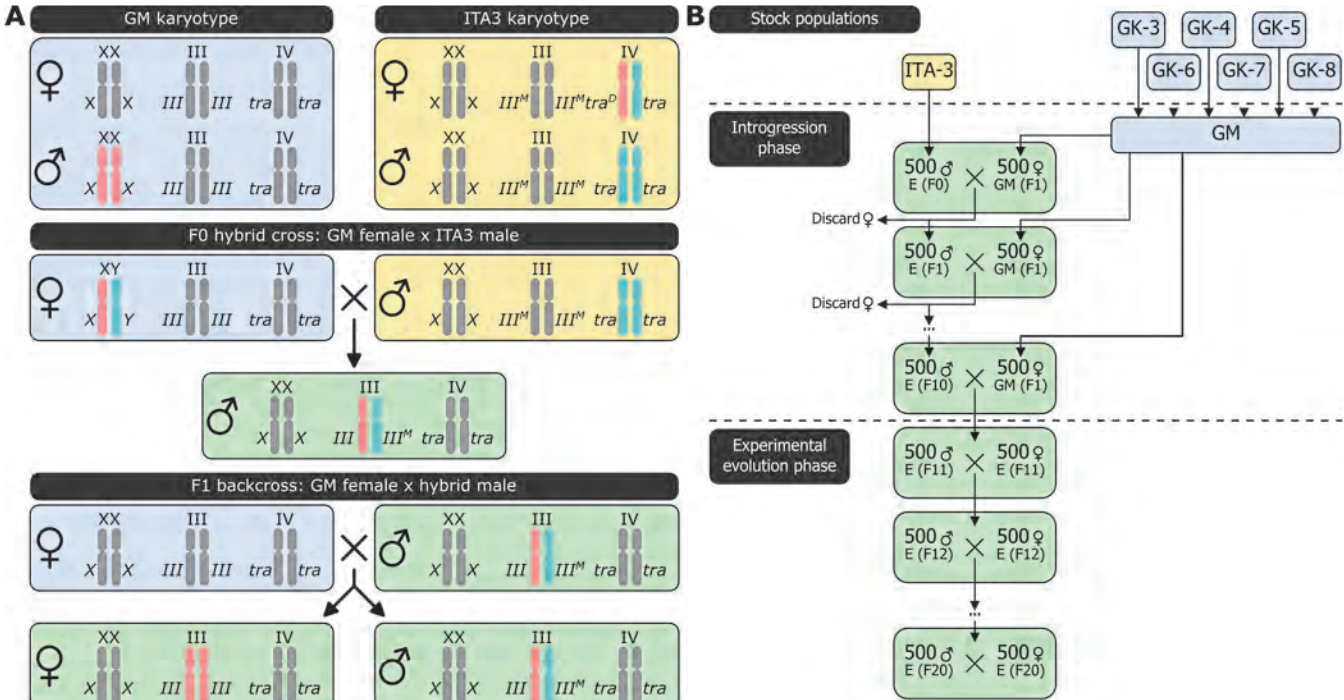


Figure 1: Establishment of baby-sex chromosomes in *Musca domestica*. (A) The M-factor on autosome III (III^M) from ITA3 functions as an utosomal co-factor for maleness, but can become a dominant male-determining gene when outcrossed to a GM background. III^M/III^M homozygosity in ITA3 results in all-male progeny in F1 offspring of GM × ITA3 crosses; backcrossing F1 males to GM females results in the formation of *III/III* females and III/III^M males. Coloured chromosomes indicate sex chromosomes, with pink chromosomes denoting female-biased sex chromosomes (*X* in *X/Y*, *III* in III/III^M, tra^D in tra^D/tra) and light blue chromosomes denoting male-biased chromosomes (*Y*, III^M, and *tra*). (B) Experimental procedure for introgression of ITA3-derived III^M into GM background. GM is newly established every generation by combining GK3 through GK8.

Table 1: Strains developed in introgression experiment.

	CG	**CI**	**EA & EB**
Male source population	GM	ITA-3	ITA-3
Female source population	GM	ITA-3	GM
Master sex determination gene	Y^M	tra^D	III^M
Sex chromosome system	XY	ZW	XY
Sex chromosome pair	XY	IV	III
Status of autosome III	Autosomal, no III^M present	Autosomal, fixed for III^M	Sex chromosome pair
XY genotype (M)	X/Y^M	*X/X*	*X/X*
XY genotype (F)	*X/X*	*X/X*	*X/X*
III genotype (M)	*+/+*	III^M/III^M	$III^M/+$
III genotype (F)	*+/+*	III^M/III^M	*+/+*
IV genotype (M)	*tra/tra*	*tra/tra*	*tra/tra*
IV genotype (F)	*tra/tra*	tra^D/tra	*tra/tra*

During introgression, offspring sex ratios were stably 50% female - 50% male except for the F1 and F2 generation in the EA and EB strains (Figure 2). For EA, the F1 generation had a strongly male-biased sex ratio (53/632 females/males, sex ratio 92.3% male) whereas for EB this male bias was less pronounced (281/471 females/males, sex ratio 62.6% male). A 100% male sex ratio is expected among these offspring because of M-factor homozygosity in the F0 ITA3 males, whereas none of the GM females carry tra^D; all offspring should therefore have a *tra/tra; X/X; M/+* genotype, which would result in maleness (see also Figure 1A). In the F2, the sex ratios for EA and EB reduced relative to the F1 sex ratios but remained slightly male-biased in both EA and EB (EA: 237/344 females/males, sex ratio 59.2% male; EB: 446/588 females/males, sex ratio 56.9% male). There are several - non-mutually exclusive - potential reasons as to why females were found in the F1 stage and similarly why sex ratios were still slightly male-biased among the F2 offspring. First, some tra^D-bearing ITA3 females may have been erroneously included during the F0 stage, resulting in some tra^D-bearing females being produced in both strains in the F1. Male-biased sex ratios in the F2 are then caused by the presence of *M/M* females carry tra^D, some males may be heterozygous for an M-factor and hence

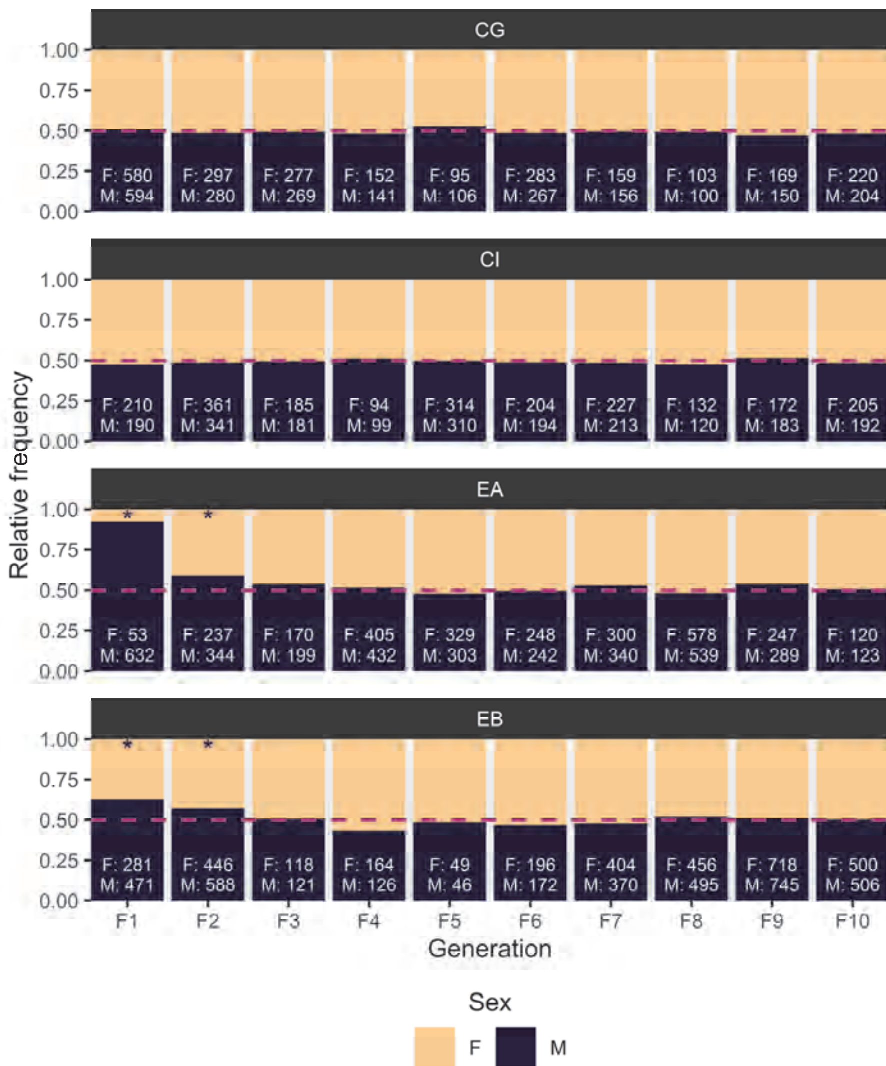


Figure 2: Sex ratios of developing offspring during introgression. Numbers indicate counts of females (F) and males (M) in each generation. Asterisks denote deviations from 50/50 sex ratios (binomial test, $P < 0.05$, Bonferonni corrected) which is depicted by the dashed pink line. Labels in facet strips indicate strain.

Biased sex ratios among F2 offspring may then be caused when the ancestral population harbours multiple unlinked M-factors, e.g. II^{M} and III^{M}. If one M-factor is fixed, additional M-factors are functionally neutral and may therefore persist. In the F1, this results in a typical male-biased sex ratio as observed here, and F1 males

that inherit a copy of the fixed M-factor and the unfixed M-factor produce F2 offspring with a 75% sex ratio. On a population level these effects could result in the slight male-biased sex ratio as observed among the F2 offspring.

Despite these discrepancies with regard to the sex ratio, the observed male-biased sex ratios among the EA F1 offspring in particular but also the EB F1 offspring suggest that the ancestral ITA3 population features a large fraction of *M/M* homozygous males or males with several unlinked M-factors. This indirectly would prove that this population also harbours tra^D-bearing females as well as that the autosomal M-factor is not (fully) sex-linked and hence not expected to undergo Y-chromosome evolution. Its baby-X complement is derived from the GM population and given that it is devoid of sex-determining variants the baby-X chromosome is guaranteed to not have evolved as an X-chromosome prior to the introgression stage. Aside from the male-biased sex ratio among F1 and F2 offspring in the EA and EB strains, no sex ratio biases were observed in the control strains CG and CI, and no sex ratio biases were observed in the EA and EB strains after this. It therefore seems unlikely that the autosomal M-factor exhibits meiotic drive or has any effect on individual fitness aside from its effect on sex determination.

Baby-sex chromosomes as a tool for testing the "sexual antagonism" model of sex chromosome evolution

Baby-sex chromosomes like those established in the EA and EB strains provide the opportunity to study sex chromosome evolution in a controlled and repeatable manner, and to do so from the very first stages after the sex chromosomes evolved a sex-determining function. This would allow us to study the role of IASC loci as predicted by the sexual antagonism hypothesis. If this hypothesis is correct and sufficient time has passed so that the baby-sex chromosome pair has been able to evolve, we expect that males experience a fitness benefit from carrying the baby-Y chromosome relative to males that lack the baby-Y, and inversely that females experience a fitness cost when carrying the baby-Y relative to females that do not. The baby-X chromosome can have a fitness benefit in females but a fitness cost in males depending on whether it is enriched for female-beneficial alleles that may have accumulated as a result of its female-biased transmission. To test these effects, fitness must be assessed of males and females bearing (1) the baby-X (*BX*) and (2)

baby-Y (*BY*) chromosomes from the EA/EB strains, (3) the M-factor-bearing autosome III (III^M) from the CI strain, and (4) the non-M-factor-bearing autosome III (*III*) from the CG strain. An additional control may be included involving (5) a non-M-factor-bearing autosome III (*+*) from a standardized tester strain (see details below). If the sexual antagonism is correct, the fitness relationships between the different genotype would be such that in males:

$$BY > \begin{matrix} III^M \\ III \\ + \end{matrix} \geq BX$$

In females, the relationship would be inverted so that:

$$BY < \begin{matrix} III^M \\ III \\ + \end{matrix} \leq BX$$

In males and females alike, the fitness effect of the baby-X chromosome is expected to be different from the various controls provided that it becomes enriched with female-beneficial alleles; otherwise, its effect is expected to equal that of the controls.

To test the fitness effects of the baby-X and baby-Y chromosomes, both need to be crossed into both male and female carriers. Baby-X females and baby-Y females may be easily obtained by a cross between a baby-sex chromosome-bearing male and a standard female (e.g. an EA or EB male with a GM female). However, baby-Y females and baby-X males are less easily generated. Baby-Y females must be generated by crossing the baby-Y chromosome into a genetic background with tra^D, e.g. by a cross between an EA male and a tra^D/tra ITA3 female. A baby-X male only arises when crossed into a background with another M-factor or a background without functional *tra*, so that a tra^{NULL}/tra^{NULL}; *BX/+* male is generated. The various control groups must also be acquired, where the procedure for III^M is similar to that for the baby-Y chromosome, whereas for III and + the procedure is similar for the baby-X chromosome. Generating the different treatment categories involving the baby-sex chromosomes would thus require a combination of different approaches. This may result in confounding effects on fitness estimates. A more straightforward

procedure may be to generate a single tester strain with *tra^{D}/tra^{NULL}; X/X; +/+* females and *tra^{NULL}/tra^{NULL}; X/X; +/+* males (where *tra^{NULL}* denotes a non-functional *tra* variant, *X* denotes a standard X-chromosome, and *+* denotes a wildtype non-M-factor-bearing autosome III). This strain can be used to cross the baby-X and baby-Y chromosomes, as well as the control autosomes *III^{M}* and *III* into both males and females in a comparable genetic background in just two generations. These crossing procedures are discussed in Box 1.

Box 1: Generating males and females carrying baby sex chromosomes using a tester strain.

To test the fitness effects of the baby-sex chromosomes and the autosomal variants that serve as controls, it is necessary to generate both males and females bearing either the baby-X, the baby-Y, the CG control autosome III (without an M-factor), or the CI control autosome III (with an M-factor). Ideally, all these types of individuals are generated within a uniform genetic background to prevent confounding effects that may occur when the baby-sex chromosomes or control autosomes are crossed into males and females using different strains.

To this end, it is useful to establish a "tester strain" with *tra^{D}/tra^{NULL}; X/X; +/+* females and *tra^{NULL}/tra^{NULL}; X/X; +/+* males, where *tra^{NULL}* refers to a loss-of-function mutant of *tra* that may be generated using e.g. CRISPR/Cas9; *X* indicates a 'standard' X chromosome, and + denotes a wildtype autosome III. M-factors can be absent as these serve no function in sex determination as maleness is achieved by *tra^{NULL}* homozygosity. The presence of *tra^{D}* is necessary to cross the baby-Y into females, whereas the *tra^{NULL}* allele is needed to ensure the baby-X can be crossed into males without use of alternative M-factors. This allows for the baby-sex chromosomes and the control chromosomes from the CG and CI strains to be crossed into males as well as females (Figure 3).

Crossing the baby-sex chromosomes into males and females is initiated by an F0 cross between a *tra^{D}/tra^{NULL}; +/+* tester female with a baby-X/baby-Y male (*tra/tra; BX/BY*) from the experimental strain resulting in four types of F1 offspring (3 female, 1 male). F1 offspring can then be crossed back to tester

females or males which results in different types of F2 offspring for each cross, but altogether includes males and females bearing a baby-X or baby-Y along with wildtype autosome III (*BX/+* or *BY/+*), or that are homozygous for the wildtype chromosome (*+/+*). An overview of the genotypes generated using this procedure is provided in Figure 3B. In this particular cross, control and BX males are however very rare (1/32 and 1/16 among F2 offspring); an additional cross may be carried out between a *tra/tra; BX/BX* female and a tra^{NULL}/tra^{NULL}; *+/+* tester male; the resulting all-female tra/tra^{NULL}; *BX/+* F1 offspring be backcrossed to tester males to obtain F2 offspring among whom *+/+* control and *BX/BX* males are expected to occur at a frequency of 1/4 each (for details, see Supplementary Figure 1).

For the CG and CI controls, the same tester strain can again be used to cross the autosome III with an M-factor (from CI, denoted III^{M}) or without it (from CG, denoted *III*) into males and females (for details, see Figure 3C and 3D). For the CI strain, only a single cross between a tester female and a CI male is necessary to acquire both males and females carrying the III^{M} chromosome. For the CG strain, an initial cross between a tester male and a CG female produces all-female F1 offspring with a tra/tra^{NULL}; *III/+* genotype. These can be backcrossed to a tester male to acquire F2 offspring males and females bearing the *III* chromosome.

These different crossing schemes together generate males and females bearing either (1) the baby-X chromosome, (2) the baby-Y chromosome, (3) the CG-derived autosome *III*, or (4) the CI-derived autosome *III* with an M-factor. There are however two caveats to this approach. First, the crossing procedure used here results in females that do carry the baby-X chromosome which lacks an M-factor and therefore tra^{D} is not strictly required for femaleness. This means two subtypes of baby-X-bearing females are generated within this crossing procedure, i.e. those with the tra^{D} allele and those without. Given its dominant feminizing effect, the tra^{D}-bearing chromosome functions as a W-chromosome and may therefore be enriched for female-beneficial alleles. This can distort the estimate of the overall fitness effect of the baby-X chromosome and must therefore be accounted for. Similarly, the CG-derived autosome *III* is not crossed to tra^{D}-bearing females in the current crossing scheme. These can however

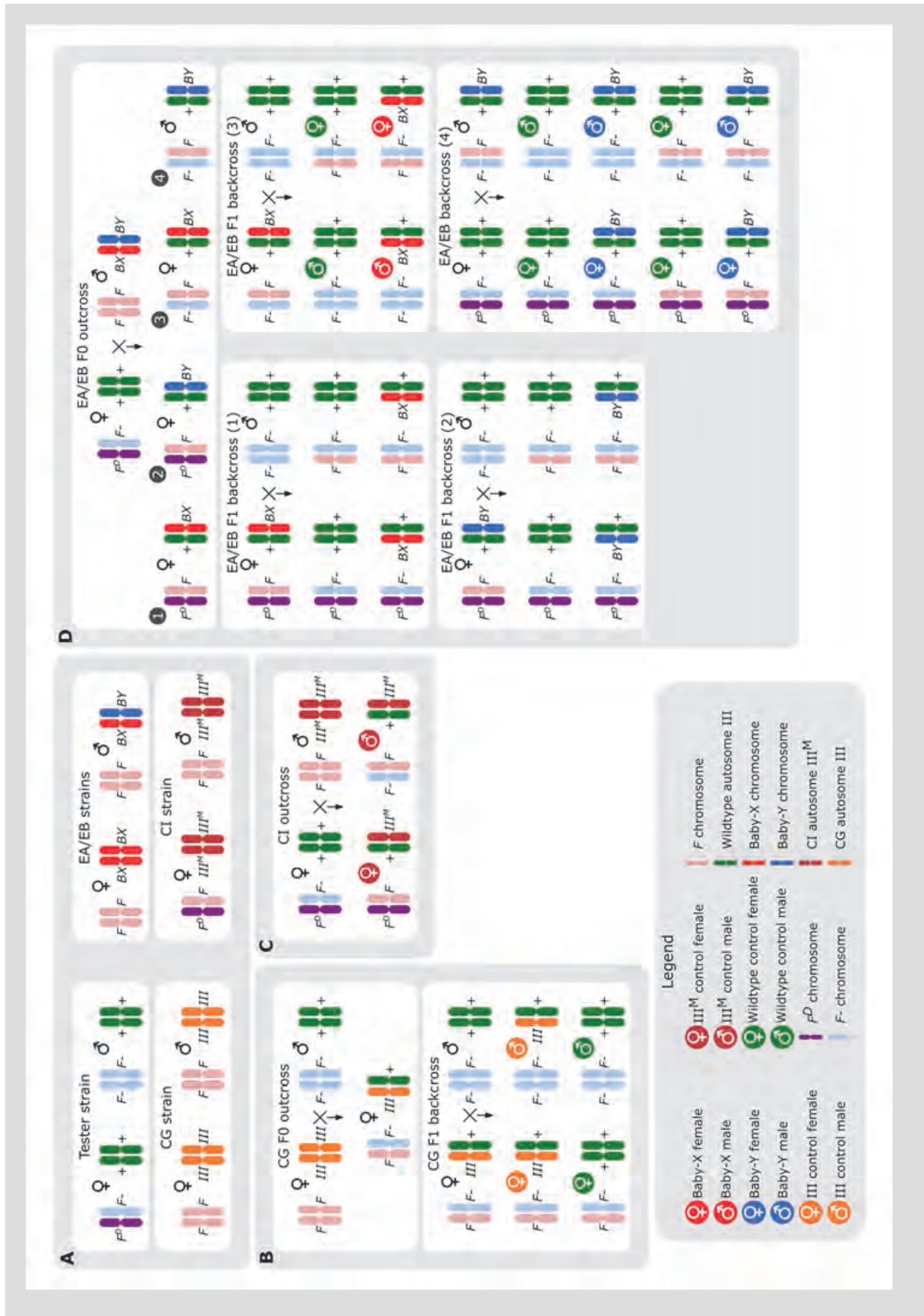

A
Tester strain
CG strain
EA/EB strains
CI strain
B
CG F0 outcross
CG F1 backcross
C
CI outcross
D
EA/EB F0 outcross
EA/EB F1 backcross (1)
EA/EB F1 backcross (2)
EA/EB F1 backcross (3)
EA/EB backcross (4)
Legend
Baby-X female
Baby-X male
Baby-Y female
Baby-Y male
III control female
III control male
III^M control female
III^M control male
Wildtype control female
Wildtype control male
F^D chromosome
F- chromosome
F chromosome
Wildtype autosome III
Baby-X chromosome
Baby-Y chromosome
CI autosome III^M
CG autosome III

Figure 3 (opposite page): Crossing schemes for generating all types of individuals required to assess the fitness effects of baby-X and baby-Y chromosomes in males and females. (A) Overview of strain karyotypes. For the CG strain, sex is determined by the XY chromosome pair (males XY, females XX; not depicted). Crossing schemes for generating (B) control males and females bearing the CG-derived *III* chromosome, (C) control males and females bearing the CI-derived III^M chromosome, and (D) males and females bearing the baby-X or baby-Y chromosome. Numbers in dark grey circles indicate the associated F1 backcross scheme. Alleles are depicted as follows: F^D = tra^D; *F* = *tra*; *F-* = tra^{NULL}; *BX* = baby-X; *BY* = baby-Y; + = wildtype (non-baby-sex chromosome); *III* = autosome III from CG strain; III^M = III^M chromosome from the CI strain. Sex determination can occur based on the *tra* (i.e. *F*) genotype or the baby-sex chromosome genotype, which are dominant in the following order: F^D > *BY* & III^M > *F* > *F-* > *BX* & + & *III*.

easily be generated by crossing F2 males with this chromosome to a tester female. The second issue is that in the crossing schemes, a sizeable fraction of the generated offspring bear none of the baby-sex chromosomes or associated controls, but instead are homozygous for the autosome *III* of the tester strain. This issue could be resolved by introducing a phenotypic marker into the tester strain, such as the *brown body* mutation that is located on autosome III (Hiroyoshi, 1977). Individuals that are homozygous for the tester strain-derived autosome III can then be identified by their brown colouration and be excluded from further experiments.

Discussion & Conclusion

The establishment of baby-sex chromosomes described here was carried out in the form of a large-scale introgression, yet the principles of this approach are in fact very simple. The introgression approach was necessary as the source strain for the autosomal M-factors had been maintained as a laboratory culture already for several years and hence was likely to have reduced genetic variation. To allow evolution to take place within the newly-established populations with baby-sex chromosomes, it was necessary to outcross this population to a genetically variable strain such as the GM strain used here. This approach may not be necessary if the ancestral strain used for the autosomal M (i.e. in which it is maintained as a co-factor for maleness behind tra^D which functions as the master sex-determining gene) harbours sufficient genetic variation; if this is the case, the two-generation crossing procedure in Figure 2 would be sufficient to generate a new baby-sex chromosome pair.

Aside from the exploitation of naturally-occurring variation, new sex chromosomes in the housefly may also be generated by genetic modification. The male-determining *Mdmd* gene can be inserted into new genomic locations using PiggyBac transformation (Wu, 2018, similar to the procedure used by Hediger et al. (2010) for *tra*), which could be used to generate new baby-Y chromosomes. This methodology may also be used to test theoretical models of SD transitions. For example, the Van Doorn & Kirkpatrick models (van Doorn & Kirkpatrick, 2007, 2010) posit that a new SD gene that evolves close to a locus experiencing IASC may result in a supergene complex that can outcompete an ancestral SD gene. By inserting *Mdmd* in the vicinity of a yet-to-be-identified IASC, such complexes can be generated and the validity of these models can be assessed. The same methodology may also be used to introduce tra^D in a novel genomic environment (Hediger et al., 2010). This opens up the possibility to not only study XY male heterogametic systems but also ZW systems in the same genomic background. Yet other genetic modification approaches, such as CRISPR/Cas9 knockout of *tra* or *Mdmd* may also be used to generate new baby-sex chromosomes of various kinds such as the tra/tra^{NULL} system. The availability of genetic modification combined with dominant male- and female-determining genetic variants in *M. domestica* therefore opens up a large variety of possibilities to generate new sex chromosome systems.

The *de novo* establishment of sex chromosomes, either by the introgression procedure as done here or by genetic modification, will be a powerful tool to study the early stages of sex chromosome evolution prior to decay of the Y-chromosome. This approach enables the study of sex chromosome evolution from the very initial stages, bypassing the issue of Y-chromosome degeneration, and to do so in a repeatable manner. The availability of different SD genes in the housefly *M. domestica* additionally enables these newly-established sex chromosomes to be crossed into both sexes, where their impacts on fitness can be evaluated in a sex-specific manner. The housefly therefore presents a unique model system wherein the "sexual antagonism" hypothesis of sex chromosome evolution can be tested with unprecedented rigour.

Methods

Fly strains & culturing procedures

Wildtype housefly stock strains were established from flies collected in Italy (ITA3) and The Netherlands (GK3 through GK8) (Figure 1); an overview of the stock strains is provided in Supplementary Table 1. Housefly stock strains were maintained at 25°C, 14:10 LD in 3250-ml bottle cages (Semadeni, Ostermundigen, Switzerland; 24 × 13.5 × 13.5 cm; L × W × H). All cages contained two 15-ml vials of sugar water (20% wt/vol) and two 15-ml vials of water, one 35-mm Petri dish with milk powder, and were shut off with tubular gauze. Sugar water and water was replaced three times per week. Egg-laying substrate was provided in cages when flies had become sexually mature (3-4 days after emergence, though mating may start earlier). Egg-laying substrate consisted of a mixture of wheat bran, flour, milk powder, and dry inactivated yeast (20: 3: 2.4: 1 ratios), of which approximately 200 g was mixed with a solution of Nipagin (Spruyt Hillen, Ijsselstein, The Netherlands; 5 ml 10% W/V in 99% ethanol, mixed with demi water to a total of 225 ml). Per cage, two 35-ml cups with egg-laying substrate were provided for 3-4 days, after which both cups were emptied into one 770-ml beaker. One hundred and fifty grams of egg-laying substrate was added to each beaker as larval feed. After this, beakers were kept under 25°C, 14:10 LD until adults emerged between 7 to 10 days later. In addition to standard culturing procedures, the ITA3 strain was cultured in a large Plexiglas cage (30 × 35 × 40 cm; L × W × H) to generate offspring for use in the introgression experiment. Each cage contained five 15-ml vials of sugar water and five 15-ml vials of water, as well as 3 Petri dishes with milk powder, and was set up using two beakers of newly-emerged adult flies. After sexual maturation, 6-12 egg-laying cups were provided which were further processed as per regular culturing conditions.

Establishment of a genetically variable strain with XY sex determination for use in introgression

The strains GK-3 to GK-8 were used to generate a genetically variable strain (GM) to be used for introgression in late 2017; although collected from the same field site,

the GK-1 and GK-2 strains were not used as these had been maintained under laboratory conditions for approximately 2 years and may have lost a substantial proportion of their genetic variation during this period. We set up cages containing newly-emerged flies for each strain as per normal culturing procedures. These beaker cages were placed within a setup consisting of two large plexiglass cages (30 × 35 × 40 cm; L × W × H) connected to each other by a circular opening (16 cm inner diameter), with lids consisting of plastic meshing for ventilation. Beaker cages were then opened to allow flies to exit into the Plexiglas cages. As per regular culturing procedures, vials containing sugar water and water were replaced three times per week, and egg-laying materials were provided upon sexual maturation. To ensure sufficient numbers of offspring could be cultured for use in the introgression experiment, we provided 24 egg-laying cups into the setup, which were further processed as per regular culturing conditions. The resulting F1 offspring were used in the introgression experiment. The GM strain was newly established every generation (rather than cultured from F1 and subsequent offspring) to prevent depleting the genetic variance in this population over the duration of the introgression phase.

Introgression of autosomal M-factors

To establish populations harbouring new proto-sex chromosomes for experimental evolution, I performed a population-scale introgression in which I crossed the autosomal M-factor of ITA-3 houseflies into a GM genetically variable background (Figure 2). To do so, I set up population cages containing 500 ITA-3 males and 500 GM virgin females. Based on the parental genotypes, this cross is expected to produce all-male F1 offspring, as all individuals should have a *tra/tra; X/X;* $III^{M}/+$ genotype. F1 males were backcrossed with GM females to produce equal numbers of *tra/tra*; *X/X;* $III^{M}/+$ F2 sons and *tra/tra*; *X/X; +/+* daughters. Daughters were discarded, and 500 randomly-collected sons were used to backcross to 500 GM females in a large Plexiglas cage (30 × 35 × 40 cm; L × W × H). This procedure was repeated until the F10 generation; F10 offspring were used to establish populations in a large Plexiglas cage (30 × 35 × 40 cm; L × W × H) which were maintained as per regular culturing procedures. Two replicate populations were established in this manner yielding experimental strains EA and EB; in addition, we

created control strains in which we either backcrossed ITA-3 males with ITA-3 females (strain CI), or GM males with GM females (CG) (see also Table 1). Every generation, I determined the sex ratio amongst the offspring for each strain by counting the number of males and females that emerged as adults from one randomly-selected beaker produced as per the culturing procedures described above.

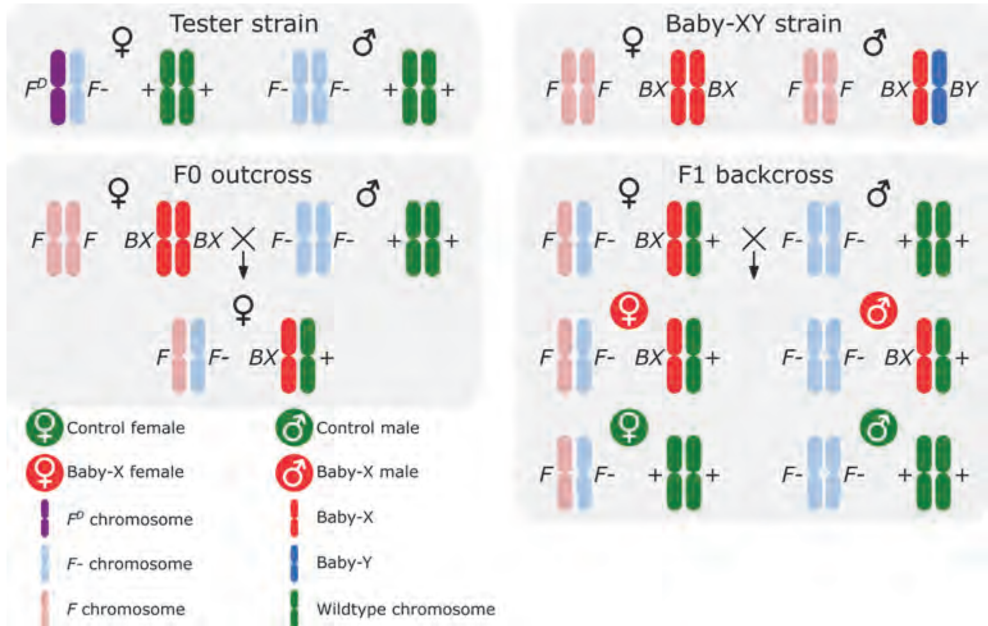
Statistical analysis

All data analysis was carried out in R (v.4.0.2, R Development Core Team, 2020) using RStudio (v.4.0.0, RStudio Team, 2020). Data wrangling and visualisation was carried out using the "cowplot" (Wilke, 2019), "tidyverse" (Wickham et al., 2019) and "viridis" packages (Garnier, 2018).

Acknowledgements
I thank Ljubinka Francuski for assistance in the introgression experiment, and the Evolutionary Genetics cluster for fruitful discussion.

Supplementary Material

Supplementary Figures



Supplementary Figure 1: Auxiliary crossing scheme for generating individuals for assessing the fitness of baby-X and baby-Y chromosomes. Alleles are depicted as follows: F^D = *tra*^{D}; *F* = *tra*; *F-* = *tra*^{NULL}; *BX* = baby-X; *BY* = baby-Y; *+* = wildtype (non-baby-sex chromosome). Sex determination can occur based on the *tra* genotype or the baby-sex chromosome genotype, which are dominant in the following order: F^D > *BY* > *F* > *F-* > *BX* & *+*. An initial cross between a baby-sex chromosome-bearing female (*F/F; BX/BX*) and a tester male (*F-/F-; +/+*) yields all-female F1 *F/F-; BX/+* offspring, which can be backcrossed to tester males to obtain *BX*-bearing females and males as well as control females and males. Contrary to the crossing procedure initiated by a cross between a tester female (F^D*/F-; +/+*) and a baby-sex chromosome-bearing male (*F/F; BX/BY*), wildtype control males and BX males each have a frequency of 1/4 among F2 offspring instead of 1/32 or 1/16.

Supplementary Tables

Supplementary Table 1: Overview of housefly stock strains used in the experiments.

Strain	**Origin**	**Coordinates**	**Collected**	**M location**	***tra^D***
GK3-GK8	Gerkesklooster, The Netherlands	53°24'N, 6°21'E	2018	XY	No
ITA3	Castellaneta Marina, Italy	40°47'N, 16°93'E	2013	III	Yes