Supplementary Methods

Derivation of analytical approximations for the invasion fitness of a W allele

1. INTRODUCTION

In this supplementary material we derive population genetic recursions for the change of allele frequencies and genetical associations under the influence of sex determination, sexually antagonistic selection, Mendelian segregation and mutation. The exact recursions are simplified by means of a Quasi-Linkage Equilibrium (QLE) approximation, following Barton & Turelli (1991) and Kirkpatrick et al. (2002), to estimate the fitness of a dominant feminizing W allele that invades into a population with XY sex determination. Since the indirect fitness effects of sexually antagonistic loci on the sex determining alleles are approximately additive if the individual sex-antagonistic loci have multiplicative effects on viability (this assumes the absence of epistasis between sex-antagonistic loci), we analyze a model consisting of three loci:

- Locus $s_1$ is the original sex-determination locus. This locus carries two alleles, $Y$, a masculinizing factor, and $X$. The frequency of the allele $Y$ is denoted as $y$
- Locus $s_2$ carries a rare novel sex-determination modifier allele, $W$. The frequency of this allele is denoted as $w$
- Locus $a$ is a locus with sexually antagonistic alleles 0 and 1. The frequency of allele 1 is denoted as $p$

To keep track of the genetic state of the population we derive equations that describe the change of genetical associations from one generation to the next. Genetical associations are defined as moments of the distribution of genotype frequencies. Details and results for the change of associations under selection, transmission and mutation can be found in Kirkpatrick et al. (2002). The following notation is used to denote the associations at different stages in the life cycle:

- **zygotes**
  - sex determination
  - juveniles
  - selection
  - adults
  - random mating

- **mating pairs**

- **Mendelian segregation**

- **zygotes**
  - change of reference values
  - zygotes
  - mutation

- **zygotes**
  - next generation

Outline of the analysis

The next section of this document (Section 2, ‘The Toolbox’) contains definitions for procedures that implement the change of associations in different stages of the life cycle. We also define procedures to simplify and linearize expressions, and we calculate the selection coefficients that quantify the effect of sex-determination and viability selection on individual alleles and statistical associations between alleles. After the preparatory work in the Section 2, we proceed to calculate the invasion fitness of the mutant sex-determination allele for the case that linkage between the sex-antagonistic locus and the sex-determination loci is weak. Section 4 then concentrates on a number of alternative scenarios. We analyze the effects of tight linkage between the sex-antagonistic locus and the novel sex-determination factor, tight linkage between the sex-antagonistic locus and the ancestral sex-determination factor, and we consider a homologous sex-determination mutation. The final section considers a population that has nearly fixated for ZW sex determination and derives the fitness of the $X$ allele when it is rare to examine the scope for bistability or a protected polymorphism of sex determining genes.

This document has been generated using Mathematica 5.2 (Wolfram 2003). The original file is available from the authors on request and can be used to verify the calculations.
2. THE TOOLBOX

Preliminaries

This is to control Mathematica's message output.

\[ \text{Off[General::"spell1"]} \]

Positions and reference values

The set \( S \) contains the positions influencing sex determination

\[ S = \{ s_{1f}, s_{1m}, s_{2f}, s_{2m} \} \]

The set \( A \) contains the positions that affect survival from the juvenile to the adult stage

\[ A = \{ a_{f}, a_{m} \} \]

The set \( W \) contains all the positions

\[ W = \text{Join}[S, A] \]

The reference values \( \phi_i \) are set equal to the allele frequencies at the zygote stage, such that first-order associations at the zygote stage vanish. Distinct reference values are used for positions with male or female sex-of-origin.

\[ \text{subsReferenceValues} = \{ \text{DeleteCases}[S, a_{f}], \text{DeleteCases}[S, a_{m}] \} \]

General procedures

Genetical associations are defined as expectations over the distribution of genotypes of the moments \( z \), where

\[ z = \phi \cdot H - \phi \cdot L \]

The following procedures are used to calculate \( z \):

\[ \text{GetIndex[position_, L_List]} := \text{First[Flatten[position, L]]} \]

\[ \text{GetMoment[K_List, Genotype_List, L_List]} := \]

\[ \text{Module[]} \]

\[ \text{Do} \]

\[ \text{Return}[\phi] \]

Associations with repeated positions are simplified by the procedure \( \text{RemoveRepeatedPositions} \)

\[ \text{RemoveRepeatedPositions[D_1[K1_List], D_2[K2_List]} := \]

\[ \text{Module[]} \]

\[ \text{If} \]

\[ \text{For example, the association} \]

\[ \text{If mating is random, then associations involving positions with different sex-of-origin can be decomposed into products of lower-order associations that involve only positions with a single sex-of-origin. The following procedure implements this decomposition.} \]
ExpandIntersexualAssociations =

\[ D_i[[K_{list}]] \Rightarrow Module[[K_f = Sort[DeleteCases[K, _a]],
K_m = Sort[DeleteCases[K, _e]]], D_i[K_f] D_i[K_m]]; \]

For example:

\[ D_i[[\{s_{1f}, s_{2a}\}]] / ExpandIntersexualAssociations \]

\[ D_i[[\{s_{1f}\}]] D_i[\{s_{2a}\}] \]

As long as the modifier allele is rare, its frequency changes at a geometric rate that is approximately constant (i.e., independent of the frequency of the modifier). The invasion of the modifier can thus be studied by linearizing the associations at the zygote stage with respect to the frequency of the modifier. The frequency of the modifier is written as \( \epsilon_w \) and it changes at a geometric rate that is approximately constant (i.e., independent of \( \epsilon_w \)). Therefore, we expand \( D_i[K] \) as

\[ D_i[K] = D_i[K] + \epsilon_w D_i[K] + \epsilon_w D_i[K] + \epsilon_w \epsilon_m D_i[K] + (\epsilon_w^2 + \epsilon_w^2) \epsilon_{k}. \]

For some associations, the lowest-order term \( D_0 \) vanishes. For other associations, specific higher-order terms can be omitted since they do not appear in the equations derived below. The procedure \texttt{Linearize} generates expansions exactly up to the order needed. The higher-order coefficients \( \epsilon_k \) are included as a means to check the results for errors.

\[ \text{Linearize} = \{ D_i[[K_{list}]] \Rightarrow Module[[n = Count[K, s_{2f}], k = Count[K, s_{1f}] + Count[K, s_{2a}]],
If[And[n == 0, k == 0],
D_0[K] + \epsilon_w D_0[K] + \epsilon_w (1 + \epsilon_k) \epsilon_k,
If[n == 1, \epsilon_w \omega (D_0[K] + \epsilon_w D_{m,w}[K]) + \epsilon_w^2 (1 + \epsilon_k) \epsilon_k, \epsilon_w^2 (1 + \epsilon_k) \epsilon_k]
], y_f \rightarrow \epsilon_w \omega_{y_f}, y_m \rightarrow \frac{1}{2} + \epsilon_w \omega_{y_m}, w_f \rightarrow 2 \epsilon_w \omega, w_m \rightarrow 0 \}; \]

In this analysis, we focus on the invasion of a dominant feminizing allele at the novel sex-determination locus. Hence, the \( W \) allele will appear only in females. At invasion, the allele frequencies at the ancestral sex determination locus are close to 0 in females and close to \( \frac{1}{2} \) in males. Accordingly, \texttt{Linearize} also substitutes \( y_f = w_m = 0 \) and \( y_m = \frac{1}{2} \).

Sex-determination procedures

The change of associations due to sex determination is determined by sex-determination coefficients \( \sigma_k \). The values of these coefficients are related to the fraction females among juveniles with genotype \( G \). We use the following procedure to specify the mapping from genotype to phenotype, which then implicitly defines the sex-determination coefficients \( \sigma_k \).

\[ \text{GetEquations}[\text{Genotype}_{list}, \text{fractionFemale}_{_}_] :=
Module[[K = Subsets[Q]], lhs = 1, i],
Do[lhs ++ \sigma_k[i] \{GetMoment[K[[i]], Genotype, S] - D_i[K[[i]]] \},
{i, 1, Length[K]] ];
(lhs /. ExpandIntersexualAssociations /. subsReferenceValues /. Linearize /. \{\epsilon_w \rightarrow 0, \epsilon_m \rightarrow 0\}) = \text{fractionFemale} / Q] \]

For example, the following equation accounts for the fact that male individuals develop as males:

\[ \text{GetEquations}[[0, 1, 0, 0], 0] \]

\[ 1 + \frac{\sigma_{(s_{1a})}}{2} = 0 \]

Similar equations for the other genotypes constrain the values of the sex-determination coefficients and the sex ratio \( Q \). The solutions for the \( \sigma_k \) that involve only ancestral sex-determination positions are calculated as
never created. As a result, some sex-determination coefficients remain unresolved (e.g.,
we do not need to make assumptions on the sex of individuals that are homozygous for
the mutant allele are solved as follows:

\[
\text{subsResCoeffs} = \text{Solve[}
\{\text{GetEquationS}[[0, 0, 0, 0], 1], \\
\text{GetEquationS}[[1, 0, 0, 0], 0], \\
\text{GetEquationS}[[0, 1, 0, 0], 0], \\
\text{GetEquationS}[[1, 1, 0, 0], 0], \\
\text{Join}[\text{DeleteCases}[	ext{Table}[[\text{Subsets}[$Q$]], \{x, 2, \text{Length}[\text{Subsets}[$Q$]]\}], \\
\sigma_{s_{}} /; \text{Count}[[K, s_{2}]] \geq 1], \{Q\}] / \text{First} / \text{ColumnForm}
\}
\sigma_{(a_{1f})} \rightarrow -1
\sigma_{(a_{1m}, a_{2m})} \rightarrow 2
\sigma_{(a_{1m})} \rightarrow -2
Q \rightarrow \frac{1}{2}
\]

The sex-determination coefficients for the mutant allele are solved as follows:

\[
\text{subsMutCoeffs} = \text{Solve[}
\{\text{GetEquationS}[[0, 0, 1, 0], 1], \\
\text{GetEquationS}[[1, 0, 1, 0], 1], \\
\text{GetEquationS}[[0, 1, 1, 0], 1], \\
\text{GetEquationS}[[1, 1, 1, 0], 1], \\
\text{GetEquationS}[[0, 0, 0, 1], 1], \\
\text{GetEquationS}[[1, 0, 0, 1], 1], \\
\text{GetEquationS}[[0, 1, 0, 1], 1], \\
\text{GetEquations}[[1, 1, 0, 1], 1]] / . \text{subsResCoeffs}, \\
\text{Cases}[	ext{Table}[[\text{Subsets}[$Q$]], \{x, 2, \text{Length}[\text{Subsets}[$Q$]]\}], \\
\sigma_{s_{}} /; \text{Count}[[K, s_{2}]] == 1], \{Q\}] / \text{First} / \text{ColumnForm}
\sigma_{(a_{2f})} \rightarrow 1
\sigma_{(a_{2m})} \rightarrow 1
\sigma_{(a_{1f}, a_{2f})} \rightarrow 1
\sigma_{(a_{1f}, a_{2m})} \rightarrow 1
\sigma_{(a_{1m}, a_{2f})} \rightarrow 2
\sigma_{(a_{1m}, a_{2m})} \rightarrow 2
\sigma_{(a_{1f}, a_{1m}, a_{2f})} \rightarrow -2
\sigma_{(a_{1f}, a_{1m}, a_{2m})} \rightarrow -2
\]

We do not need to make assumptions on the sex of individuals that are homozygous for $W$; if $W$ is dominant, such individuals are never created. As a result, some sex-determination coefficients remain unresolved (e.g., $\sigma_{(a_{1f}, a_{2m})}$), but these do not appear in our final results.

The procedure \text{SexDetermination} expresses associations in male and female juveniles after sex-determination in terms of associations at the zygote stage.

\text{SexDetermination} = \{
\text{Df}_{2}[[K, \text{List}]] \rightarrow \\
\text{Module}[[\{L = \text{Subsets}[$Q$], \text{sum} = D_{1}[K], i\}], \\
\text{Do}[[\text{sum} = \sigma_{l} \rightarrow \sigma_{l} \rightarrow \text{RemoveRepeatedPositions}[[D_{1}[K], D_{1}[L[[i]]]]] - \\
D_{1}[K] D_{1}[L[[i]]], \{i, 1, \text{Length}[L]\}], \\
\text{sum} /. \text{ExpandIntersexualAssociations} / . \text{subsReferenceValues} / . \text{subsResCoeffs} / . \text{subsMutCoeffs}],
\text{Dm}_{2}[[K, \text{List}]] \rightarrow \\
\text{Module}[[\{L = \text{Subsets}[$Q$], \text{sum} = D_{1}[K], i\}], \\
\text{Do}[[\text{sum} = \sigma_{l} \rightarrow \sigma_{l} \rightarrow \text{RemoveRepeatedPositions}[[D_{1}[K], D_{1}[L[[i]]]]] - \\
D_{1}[K] D_{1}[L[[i]]], \{i, 1, \text{Length}[L]\}], \\
\text{sum} /. \text{ExpandIntersexualAssociations} / . \text{subsReferenceValues} / . \text{subsResCoeffs} / . \text{subsMutCoeffs}]]
\]

The resulting expressions are already quite complicated. The following example shows the change of the frequency of allele $Y$ in males due to sex determination:
\[ Dm_{2}([{s}_{1n}]) / . \text{SexDetermination} \]

\[ 2(1 - y_n) y_n + 2(1 - y_n) y_n D_1([{s}_{1f}, {s}_{2f}]) - D_1([{s}_{1n}, {s}_{2n}]) - \]

\[ 2(1 - 2 y_n) D_1([{s}_{1n}, {s}_{2n}]) - \sigma([{s}_{1f}, {s}_{2f}, {s}_{2n}]) D_1([{s}_{1f}, {s}_{2f}]) D_1([{s}_{1n}, {s}_{2n}]) - \]

\[ (1 - 2 y_n) \sigma([{s}_{1f}, {s}_{2f}, {s}_{2n}, {s}_{2n}]) D_1([{s}_{1f}, {s}_{2f}]) D_1([{s}_{1n}, {s}_{2n}]) \]

The result simplifies after linearization, and we find the expected result of a value that is close to \( \frac{1}{2} \) (in the absence of the \( W \) allele, the frequency of \( Y \) at the sex chromosome that is inherited from the father changes by \( \frac{1}{2} \), from \( \frac{1}{2} \) in zygotes before sex determination to \( 1 \) in male juveniles):

\[ \text{Series}[Dm_{2}([{s}_{1n}]) / . \text{SexDetermination} / . \text{Linearize} / . \{\epsilon_{a} \rightarrow 0\}, \{\epsilon_{w}, 0, 1\}] \]

\[ \frac{1}{2} + \frac{1}{2} w D_{w}([{s}_{1f}, {s}_{2f}]) \epsilon_{w} + O[\epsilon_{w}]^{2} \]

**Viability selection procedures**

These two procedures relate the viability of a male or female juvenile to its genotype. The relationship is expressed in terms of viability-selection coefficients \( \alpha_{K} \), which measure the strength of selection on the positions in \( K \). Throughout, we assume that selection is weak, allowing us to truncate these expressions at first order in \( \epsilon_{a} \).

\[ \text{GetEquationAfemale[GenotypeList, selectionCoeff_] :=} \]

\[ \text{Module}[\{ K = \text{Subsets}[\mathbb{A}], \text{lhs} = 1, i, \}
\]

\[ \text{Do[} \text{lhs} += \epsilon_{a} \alpha_{[K[[i]]]}/.[\text{GenotypeList} \rightarrow \text{selectionCoeff}]\text{,} i\}
\]

\[ \text{GetEquationAmale[GenotypeList, selectionCoeff_] :=} \]

\[ \text{Module}[\{ K = \text{Subsets}[\mathbb{A}], \text{lhs} = 1, i, \}
\]

\[ \text{Do[} \text{lhs} += \epsilon_{a} \alpha_{[K[[i]]]}/.[\text{GenotypeList} \rightarrow \text{selectionCoeff}]\text{,} i\}

For example, to specify that heterozygote females have fitness \( 1 + h_{2}v_{r} \), we write

\[ \text{GetEquationAfemale[\{0, 1\}, h_{2} v_{r}] \}
\]

\[ 0 = \epsilon_{a} ( - h_{2} v_{r} - p_{t} \alpha_{[(s, t), (s, t)]} + \alpha_{[(s, m), (s, m)]} - p_{t} \alpha_{[(m, t), (m, t)]} + p_{t} p_{a} \alpha_{[(t, a), (t, a)]} + \alpha_{[(t, a), (t, a)]} - p_{t} p_{a} \alpha_{[(t, a), (t, a)]} + \alpha_{[(t, a), (t, a)]} ) \]

\[ \text{GetEquationAmale[\{0, 1\}, h_{2} v_{r}] \}
\]

\[ 0 = \epsilon_{a} ( - h_{2} v_{r} - p_{t} \alpha_{[(s, t), (s, t)]} + \alpha_{[(s, m), (s, m)]} - p_{t} \alpha_{[(m, t), (m, t)]} + p_{t} p_{a} \alpha_{[(t, a), (t, a)]} + \alpha_{[(t, a), (t, a)]} - p_{t} \alpha_{[(t, a), (t, a)]} + \alpha_{[(t, a), (t, a)]} ) \]

The values of the selection coefficients \( \alpha_{K} \) and the mean fitness in males and females are defined by additional equations for the other genotypes. The solution of the resulting system of linear equations is given by
The procedure \textsf{ViabilitySelection} expresses associations in adults after viability selection in terms of associations at the juvenile stage.

\textbf{ViabilitySelection} = \{
    \textbf{Df}_3[\textit{K}_\text{List}] \rightarrow \\
    \text{Module}[\textit{L} = \text{Subsets}[\textit{A}], \text{sum} = \text{Df}_2[\textit{K}], i], \\
    \text{Do}[\text{sum} += \varepsilon_0 \alpha_{\{i\}}/(1-\text{ss} \text{ko})(\text{RemoveRepeatedPositions}[\text{Df}_2[\textit{K}]], \\
        \text{Df}_2[[\text{L}[[i]]]] - \text{Df}_2[\textit{K}] \text{Df}_2[[\text{L}[[i]]]]), \{i, 1, \text{Length}[\text{L}]\}]; \\
    \text{sum} /\!\!/ \text{subsReferenceValues}], \\
    \text{Dm}_3[\textit{K}_\text{List}] \rightarrow \\
    \text{Module}[\textit{L} = \text{Subsets}[\textit{A}], \text{sum} = \text{Dm}_2[\textit{K}], i], \\
    \text{Do}[\text{sum} += \varepsilon_0 \alpha_{\{i\}}/(1-\text{ss} \text{ko})(\text{RemoveRepeatedPositions}[ \\
        \text{Dm}_2[\textit{K}], \text{Dm}_2[[\text{L}[[i]]]] - \text{Dm}_2[\textit{K}] \text{Dm}_2[[\text{L}[[i]]]]); \\
        \{i, 1, \text{Length}[\text{L}]\}]; \\
    \text{sum} /\!\!/ \text{subsReferenceValues}];
\}

\textbf{Transmission and the change of reference values}

The procedure \textsf{getTransmissionCoefficient} calculates the probability that a set of positions K ends up in a single gamete. These probabilities are dependent on the order of the genes on the chromosome, and below, we will consider different cases:

\textbf{getTransmissionCoefficient}[\textit{K}_\text{List}] := \text{Module}[\textit{i} = 1, j, t = 1/2, r = 0, \text{L}], \\
    \text{If}[\text{Length}[\textit{K}] = 0, \text{Return}[1]]; \\
    \text{While}[! \text{MemberQ}[\textit{K}, \text{geneOrder}[[\textit{i}]]], \text{i}++]; \\
    \text{j} = \text{i} + 1; \\
    \text{While}[\text{j} \leq \text{Length}[\text{geneOrder}], \\
        \text{r} = (1-n) \text{recombinationRate}[[\text{j} - 1]] + r(1 - \text{recombinationRate}[[\text{j} - 1]]); \\
        \text{L} = \text{Cases}[\textit{K}, \text{geneOrder}[[\text{i}]] \mid \text{geneOrder}[[\text{j}]]]; \\
        \text{If}[\text{Length}[\text{L}] = 2, \\
            \text{t} = \text{If}[\text{Count}[\text{L}, _x] = 1, \text{r}, 1 - \text{r}]; \\
            \text{i} = \text{j}; \\
            \text{r} = 0;]; \\
            \text{j}++]; \\
    \text{Return}[\text{Simplify}[\text{t}]];\}
Since the reference values are not equal for all positions at a single locus, the associations have to be adjusted to the reference values after the transmission event. This is done by the procedure \texttt{getAdjustedAssociation}.

\begin{verbatim}
getAdjustedAssociation[K_List, isFemale_] := Module[
  {L = Subsets[K], M, Mc, i, j, sum = 0, product},
  Do[
    M = L[[i]];  
    Mc = Sort[Complement[K, M]];  
    product = 1;
    Do[product *= (\[Phi][K[[j]]] - \[Phi][isFemale, M[[j]]]/\[Phi][M[[j]]]/\[Phi][M[[j]]] - \[phi][M[[j]]])' ,
      {j, 1, Length[M]}]; 
    sum += product If[isFemale, Df[j][Mc], Dm3[Mc]],
    {i, 1, Length[L]}];
  sum /. subsReferenceValues]
\end{verbatim}

The procedure \texttt{Transmission} expresses the association in the new generation of zygotes after transmission in terms of the associations in the adults of the parental generation.

\begin{verbatim}
Transmission = D4[K_List] := Module[{Kf = DeleteCases[K, _n],
  Km = DeleteCases[K, _l], Lf, Lm, sumFemale = 0, sumMale = 0},
Lf = Tuples[Kf /. {locus_ -> {locus, locus}}];
Lm = Tuples[Km /. {locus_ -> {locus, locus}}];
Do[sumFemale += getTransmissionCoefficient[Lf[[i]]]
    getAdjustedAssociation[Lf[[i]], True], {i, 1, Length[Lf]}];
Do[sumMale += getTransmissionCoefficient[Lm[[i]]]
    getAdjustedAssociation[Lm[[i]], False], {i, 1, Length[Lm]}];
sumFemale + sumMale];
\end{verbatim}

The reference values have to be updated at the end of the generation. A change of reference values changes the associations. The procedure \texttt{ChangeReferenceValues} expresses the associations in zygotes after the change of reference values in terms of associations before the change.

\begin{verbatim}
ChangeReferenceValues = D5[K_List] := Module[{L = Subsets[K], M, Mc, i, j, sum = 0, product},
  Do[
    M = L[[i]];  
    Mc = Sort[Complement[K, M]];  
    product = 1;
    Do[product *= -D4[{M[[j]]}], {j, 1, Length[M]}]; 
    sum += product D4[Mc],
    {i, 1, Length[L]}];
  sum /. subsReferenceValues];
\end{verbatim}

**Complete lifecycle including mutation**

By concatenating the procedures for transmission, viability selection and sex determination, we map the associations at the zygote stage from one generation to the next.

\begin{verbatim}
LifeCycle = 
    SexDetermination);
\end{verbatim}

In the final step of our preparatory work, we include mutation to the processes that affect the change of allele frequencies and associations. Mutation occurs only at the sex-antagonistic locus. The rate of mutation is scaled by $\epsilon_n \epsilon_m$, where $\epsilon_n$ is again an order-of-magnitude parameter that will be used later on to linearize the results for small mutation rates.
GetFrequencyChange[k_, a, orderA_List, orderW_List] :=
Module[{f = ϕk1 + D4[{{k1}}], Δf},
Δf = f (1 - e_a e_μ μ_0) + (1 - f) e_a e_μ μ_0 - ϕk1 /. subsReferenceValues;
FullSimplify // Normal]

GetFrequencyChange[k_, Or[k == s1, k == s2], orderA_List, orderW_List] :=
Normal

GetAssociationChange[K_List, orderA_List, orderW_List] :=
Series[D5[K] (1 - e_a e_μ μ_0 - e_a e_μ μ_1) Count[{{k1, 0, 0}}, -D1[K]] /. LifeCycle /. Linearize,
orderW, orderA] // FullSimplify // Normal

3. QLE APPROXIMATION - RESULTS FOR LOOSE LINKAGE

Parameters

In this section, we consider a sex-antagonistic locus that is located between the two sex-determination loci. The results for this scenario are general if there are no epistatic interactions between the sex factors. In that case, the recombination rate between the sex-factors does not enter in any of equations, implying that the precise order of the genes on the chromosome is irrelevant.

geneOrder = {s1, a, s2}
{s1, a, s2}
recombinationRate = {τ(s1, a), τ(a, s2)}
{τ(s1, a), τ(a, s2)}

Sex-ratio selection

We start with a number of consistency checks. In the absence of the modifier allele, the ancestral sex-determination alleles should be at their equilibrium frequencies.

GetFrequencyChange[s1f, {e_a, 0, 0}, {e_w, 0, 0}]
0

GetFrequencyChange[s1a, {e_a, 0, 0}, {e_w, 0, 0}]
0

In the absence of sex-antagonistic selection (e_a → 0), the novel sex-determination allele is neutral. Its frequency should then neither increase nor decrease.

Δwf = GetFrequencyChange[s2f, {e_a, 0, 0}, {e_w, 0, 1}]
0

We proceed by calculating the change of the genetical association between the sex-determination loci in females.

ΔDwyf = GetAssociationChange[{s1f, s2f}, {e_a, 0, 0}, {e_w, 0, 1}]

w (-τ(s1, a) + τ(a, s2) (-1 + 2 τ(s1, a))) e_w (-1 + D_w [{s1f, s2f}])

An equilibrium that maintains an equal sex ratio is attained for the following value of the association:

solutionDwy = Solve[ΔDwyf == 0, D_w [{s1f, s2f}]] // Flatten // FullSimplify
{D_w [{s1f, s2f}] → 1}
Sex-antagonistic selection

- Frequencies of sex-antagonistic alleles and associations with the ancestral sex determination locus

Recurrence relations for the sex-antagonistic allele frequencies are given by

\[ \Delta p_t = \text{GetFrequencyChange}[a_t, \{e_a, 0, 1\}, \{e_w, 0, 0\}] \]

\[ \frac{1}{2} \left( -p_t + p_0 - 2 D_0 ([a_n, s_1]) \right) + \frac{1}{2} e_a (2 e_w (\mu_0 - (e_w (\mu_0 + \mu_1) + \alpha_{(n),i}) (1 + p_0 - 2 D_0 ([a_n, s_1]))) \]

\[ p_n (2 D_0 ([a_n, s_1])) - p_t^2 (2 \alpha_{(t),i} - 2 \alpha_{(t,n),i} D_0 ([a_n, s_1])) + \]

\[ \text{Pr} (\alpha_{(t),i} - e_w (\mu_0 + \mu_1) - 2 \alpha_{(t,n),i} D_0 ([a_n, s_1])) - 2 D_0 ([a_n, s_1])] \]

\[ \Delta p_n = \text{GetFrequencyChange}[a_n, \{e_a, 0, 1\}, \{e_w, 0, 0\}] \]

\[ \frac{1}{2} (p_t - p_n) + D_0 ([a_n, s_1]) + \frac{1}{2} e_a (-p_t (2\alpha_{(t),m} + 2 \alpha_{(t,n),m} D_0 ([a_n, s_1])) + \]

\[ \text{Pr} (\alpha_{(t),m} - e_w (\mu_0 + \mu_1) + 2 \alpha_{(t,n),m} D_0 ([a_n, s_1])) - (p_n + 2 D_0 ([a_n, s_1])) \]

\[ (e_w (\mu_0 + \mu_1) + \alpha_{(n),m}) (1 + p_n + 2 D_0 ([a_n, s_1])) + 2 (e_w (\mu_0 + D_0 ([a_n, s_1]))) \]

The only genetical association that appears in these recurrences is that between the sex-antagonistic allele and the \( \gamma \)-allele in male gametes. This association changes from one generation to the next as given by the recurrence

\[ \Delta D_{\gamma n} = \text{GetAssociationChange}([a_n, s_1], \{e_a, 0, 1\}, \{e_w, 0, 0\}) \]

\[ \frac{1}{4} \left( (p_t - p_n) (1 + 2 r_{(s1,a)}) - 2 (1 + 2 r_{(s1,a)}) D_0 ([a_n, s_1]) \right) + \]

\[ \frac{1}{4} e_a (-p_t^2 (1 + 2 r_{(s1,a)}) (2 \alpha_{(t),a} + 2 \alpha_{(t,n),a} D_0 ([a_n, s_1])) + \]

\[ \text{Pr} (-1 + 2 r_{(s1,a)}) (\alpha_{(t),a} - e_w (\mu_0 + \mu_1) + 2 \alpha_{(t,n),a} D_0 ([a_n, s_1])) + \]

\[ (1 + 2 r_{(s1,a)}) (\alpha_{(t),a} + 1 + p_n + 2 D_0 ([a_n, s_1])) - 2 (1 + 2 r_{(s1,a)}) D_0 ([a_n, s_1]) \]

If recombination between \( a \) and \( s_1 \) occurs frequently then the difference between allele frequencies in male and female gametes will be small. Accordingly, \( p_t \) and \( p_n \) can be expressed in terms of an average allele frequency \( \bar{p} \) and a sex-specific deviation from the mean that is related to the association between sex-of-origin and the sex-antagonistic allele.

\[ \{ \text{varSubs} = (p_t \rightarrow \bar{p} + 2 e_a D_0 ([a, sex]), p_n \rightarrow \bar{p} - 2 e_a D_0 ([a, sex])) \} // \text{ColumnForm} \]

\[ p_t \rightarrow \bar{p} + 2 e_a D_0 ([a, sex]) \]

\[ p_n \rightarrow \bar{p} - 2 e_a D_0 ([a, sex]) \]

If the difference between \( p_t \) and \( p_n \) is small, then also the association between \( a \) and \( s_1 \) will be small. Indeed, the leading order term in the expansion of this association is of order \( e_a \):

\[ \text{Solve} [0 \rightarrow \Delta D_{\gamma n} /. \text{varSubs} /. \{e_a \rightarrow 0\}, D_0 ([a_n, s_1])] \]

To leading order in \( e_a \), the selection coefficients reduce to the average effects of allele substitution.

\[ \text{definitionCoeffSimple} = \{
\]

\[ \hat{V} \rightarrow \bar{p} (1 - \bar{p}), \]

\[ \alpha_t \rightarrow \nu_t (h_t + \bar{p} (1 - 2 h_t)), \]

\[ \alpha_n \rightarrow \nu_n (h_n + \bar{p} (1 - 2 h_n)) \]

\} ;

Up to leading order in \( e_a \), the selection coefficients reduce to the average effects of allele substitution.
subsCoeffSimple = 
  Join[
    
    
    
    
    
    
    
    
    ];

The simplified recurrence equations are given by

```
simpleDp = 
  Series[Collect[Dp /. subsCoeffSimple, p] /. {p^2 -> p - V}, {e_a, 0, 1}] // Normal // FullSimplify
```

```
es_0 (V <D< - e_0 ((-1 + p) mu_0 + p mu_1) - 2 D_o [{a, sex}] - D_o [{a_n, s1_n}])
```

```
simpleDp = 
  Series[Collect[Dp /. subsCoeffSimple, p] /. {p^2 -> p - V}, {e_a, 0, 1}] // Normal // FullSimplify
```

```
e_0 (V <D< + e_0 (mu_0 - p mu_0 - p mu_1) + 2 D_o [{a, sex}] + D_o [{a_n, s1_n}])
```

```
simpleDyp = 
  Series[Collect[Dyp /. subsCoeffSimple, p], {e_a, 0, 1}] // Normal // FullSimplify
```

```
1/2 e_0 ((-2 + 4 r_{a,s1}) D_o [{a, sex}] - (1 + 2 r_{s1,a}) D_o [{a_n, s1_n}])
```

```
simpleDpAvg = simpleDp + simpleDp // FullSimplify
```

```
e_0 (V <D< + e_0 ((-1 + p) mu_0 + p mu_1))
```

The equilibrium solution for the associations is

```
(solutionDyp = Solve[{simpleDyp = 0, simpleDp - simpleDp = 0},
  {D_o [{a_n, s1_n}], D_o [{a, sex}]}]) // Flatten // FullSimplify) // ColumnForm
```

```
D_o [{a_n, s1_n}] -> V [-1 - 2 r_{a,s1} | (s1_a) ]
```

```
D_o [{a, sex}] -> V [1 - 2 r_{s1,a} | (a,s1) ]
```

• Associations with the novel determination locus

In the final step of the analysis for weak linkage, we need to calculate the associations between the sex-antagonistic locus and the novel sex-determination factor. The equation is complicated and difficult to interpret without simplification.

```
DeltaDwp = GetAssociationChange[{a_t, s2_t}, {e_a, 0, 1}, {e_w, 0, 1}];
```

All associations involving s_2 and a are of the order e_w when linkage is weak, and the lowest order terms in the expansion of these associations vanish.

```
solutionDwp = Solve[DeltaDwp = 0 /. subsCoeffSimple /. {e_a -> 0}, D_w [{a_t, s2_t}]][[1]]
```

```
D_w [{a_t, s2_t}] -> 0
```

Previous results are used to simplify the expressions.

```
subsCoeffSimple = Join[subsCoeffSimple, solutionDwp];
```
Indirect selection on the modifier

Using the previous QLE solutions for the associations, the geometric rate of increase of the modifier allele frequency is calculated as

\[
\text{finalResultWeakLinkage} = \frac{\text{GetFrequencyChange}[s2f, \{e_u, 0, 2\}, \{e_w, 0, 1\}]}{2 e_u \omega} \text{/. subsPreviousResults /. solutionDwp // FullSimplify}
\]

This expression is identical to equation (3) in the main text.

4. APPROXIMATIONS FOR TIGHT LINKAGE

If linkage between loci is tight, the allele-frequency differences between males and females can become large, as can the associations between sex-antagonistic alleles and the sex-determination factors. For tight linkage we can therefore no longer use the linearization and simplification techniques that we used for the loose-linkage case. Instead, we develop population-genetic recursions for the allele frequencies of the sex-antagonistic allele in the different types of gametes that contain a mutant allele. The recursions are then applied to calculate the fitness of a \( W \) allele (1) with a sex-antagonistic locus tightly linked to the new sex-determining gene, (2) with a sex-antagonistic locus tightly linked to the ancestral sex-determining gene, and (3) during a homologous transition.

Transition matrices and population-genetic recursions

The state variables of our model are the frequencies of the different types of gametes that contain a mutant allele. Mutant alleles can occur in gametes together with a \( x \)- or \( y \)-chromosome and they can occur with either one of the sex-antagonistic alleles. Mutant alleles are always maternally inherited. Accordingly, four variables are required to keep track of the genetic state of the mutant population. Let \( \mathbf{g}_t \) be the state vector of the mutant population at time \( t \). Our aim is to derive a linear recurrence equation \( \mathbf{g}_{t+1} = \mathbf{T} \mathbf{g}_t \) that describes the growth of the mutant population when the mutant allele is rare. We write the vector \( \mathbf{g} \) of mutant gamete frequencies as

\[
\mathbf{g} = \begin{pmatrix}
\psi(WX0) \\
\psi(WX1) \\
\psi(WY0) \\
\psi(WY1)
\end{pmatrix}
\]

where \( \psi(WX0) \) is the frequency in generation \( t \) of female gametes carrying the mutant allele \( W \), an \( x \)-chromosome, and allele 0 at the sex-antagonistic locus \( a \). The frequencies of other gamete types follow similarly. To construct the matrix \( \mathbf{T} \), we proceed in steps through the life cycle: fusion of gametes to produce female juveniles, viability selection on juveniles to produce adults, meiosis to produce unmutated haploid gametes, and finally mutation at the sex antagonistic locus to give mutated gametes.

- Fusion of gametes

In the first step, we map the frequencies of genotypes in gametes to those in zygotes. The zygote genotypes are divided into two classes, depending on whether or not the gamete from the resident parent contained a \( y \)-chromosome or not. The vector of frequencies for zygotes that inherited an \( x \)-chromosome from their nonmutant parent is written
where \( \psi(\zeta X 0 / W X 0) \) is the frequency of zygotes that inherited the haplotype \( \zeta X 0 \) from one parent and \( W X 0 \) from the other; the other zygote frequencies are denoted similarly. Then

\[
\mathbf{z}_x = \mathbf{G}_x \mathbf{g}_i
\]

where the matrix \( \mathbf{G}_x \) is defined as

\[
\begin{pmatrix}
1 - p_{x_0} & 0 & 0 & 0 \\
p_{x_0} & 0 & 0 & 0 \\
0 & 1 - p_{x_0} & 0 & 0 \\
0 & p_{x_0} & 0 & 0 \\
0 & 0 & 1 - p_{x_m} & 0 \\
0 & 0 & 0 & p_{x_m} \\
0 & 0 & 0 & 1 - p_{x_m} \\
0 & 0 & 0 & 0 & p_{x_m}
\end{pmatrix}
\]

The second class of zygotes are those that inherited a y-chromosome from their resident parent. The vector of their frequencies is:

\[
\mathbf{z}_y = \begin{pmatrix}
\psi(\zeta Y 0 / W X 0) \\
\psi(\zeta Y 1 / W X 0) \\
\psi(\zeta Y 0 / W X 1) \\
\psi(\zeta Y 1 / W X 1) \\
\psi(\zeta Y 0 / W Y 0) \\
\psi(\zeta Y 1 / W Y 0) \\
\psi(\zeta Y 0 / W Y 1) \\
\psi(\zeta Y 1 / W Y 1)
\end{pmatrix} = \mathbf{G}_y \mathbf{g}_i
\]

where matrix \( \mathbf{G}_y \) is given by

\[
\begin{pmatrix}
1 - p_{y_0} & 0 & 0 & 0 \\
p_{y_0} & 0 & 0 & 0 \\
0 & 1 - p_{y_0} & 0 & 0 \\
0 & p_{y_0} & 0 & 0 \\
0 & 0 & 1 - p_{y_m} & 0 \\
0 & 0 & 0 & p_{y_m} \\
0 & 0 & 0 & 1 - p_{y_m} \\
0 & 0 & 0 & 0 & p_{y_m}
\end{pmatrix}
\]

Viability selection acts on juveniles to produce surviving adults. The vectors of genotype frequencies for adult females that carry an x- or y-chromosome from the nonmutant parent are

\[
\mathbf{a}_{y} = \mathbf{A}_y \mathbf{z}_y, \quad \mathbf{a}_{x} = \mathbf{A}_x \mathbf{z}_x
\]
where $A_f$ is a diagonal matrix with the viabilities of female genotypes:

$$
A_f = \text{DiagonalMatrix}[(1, 1 + h_f v_f e_a, 1 + h_f v_f e_a, 1 + v_f e_a, 1 + h_f v_f e_a, 1 + v_f e_a)] / (1 + e_a \tilde{v}_f) \end{equation}$$

\[
\begin{pmatrix}
\frac{1}{1 + e_a v_f} & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1 - h_f v_f e_a}{1 + e_a v_f} & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1 - h_f v_f e_a}{1 + e_a v_f} & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{1 + e_a v_f} & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1 - h_f v_f e_a}{1 + e_a v_f} & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1 - h_f v_f e_a}{1 + e_a v_f} \\
0 & 0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}
\]

- **Formation of gametes**

The next step is meiosis, which produces unmutated haploid gametes. The vector of gamete frequencies produced by females is written $h_f$, with elements corresponding to the first four elements of the vector $g_t$ defined earlier:

$$
h_f = R_X a_f + R_Y a_f,
$$

The matrices $R_X$ and $R_Y$ account for recombination, and are calculated by the following procedures.

```plaintext
indexY[i_Integer] := Mod[Quotient[i, 4], 2];
recombinationCoef[i_Integer, j_Integer] := Module[
  YfromResidentGamete = indexY[i],
  AfromMutantGamete = Mod[Quotient[i, 2], 2],
  AfromResidentGamete = Mod[i, 2],
  Yoffspring = Quotient[j, 2],
  Aoffspring = Mod[j, 2], I],
  I[parentAllele_, offspringAllele_] :=
    parentAllele offspringAllele + (1 - parentAllele) (1 - offspringAllele);
getTransmissionCoefficient[{a_, b_, c_}] :=
  I[YfromMutantGamete, Yoffspring] I[AfromMutantGamete, Aoffspring] +
  getTransmissionCoefficient[{a, b, c}] +
  I[YfromResidentGamete, Yoffspring] I[AfromResidentGamete, Aoffspring] +
  getTransmissionCoefficient[{a, b, c}] +
  I[YfromResidentGamete, Yoffspring] I[AfromMutantGamete, Aoffspring] +
  getTransmissionCoefficient[{a, b, c}] I[YfromResidentGamete, Yoffspring] // FullSimplify
R_X = Table[recombinationCoef[i, 0, j], {j, 0, 3}, {i, 0, 7}];
R_Y = Table[recombinationCoef[i, 1, j], {j, 0, 3}, {i, 0, 7}];
```

- **Mutation**

The final step to complete the life cycle is mutation at the sex-antagonistic locus, which then gives us the gamete frequencies in generation $t+1$:

$$
g_{t+1} = U h_f
$$

where the matrix $U$ is
The linkage between the ancestral sex-antagonistic locus and the novel sex-determination locus is tight if the recombination rate between them is low. When this is the case, the invasion fitness is calculated as

\[
W = \frac{1}{e_u} \text{NormalSeries}[(p_w - p_x) w_x (1 - w_x) / . \text{Linearize}, \{e_u, 0, 1\}] / . \text{FullSimplify}
\]

Putting these results together gives us a linear recursion equation for the vector of gamete frequencies:

\[
g_{n+1} = T g_n
\]

where the transition matrix \(T\) is:

\[
T = U (R_x A_f G_x + R_y A_f G_y)
\]

**Scenario 1: tight linkage between the sex-antagonistic locus and the novel sex-determination gene**

The linkage between the ancestral sex-antagonistic locus and the novel sex-determination locus is tight if the recombination rate between \(S_2\) and \(A\) is of the same order of magnitude as the selection coefficients.

\[
\text{tightLinkage1} = \{r_{(a,a2)} \rightarrow e_a \bar{F}_{(a,a2)} \}
\]

The association between the sex-antagonistic locus and the novel sex-determination locus depends on the difference between the allele frequencies \(\bar{p}\) and \(p_w\). The exact relationship can be found from the definition for the association. For example,

\[
D([a_x, s_x]) = w_x p_w \left(1-w_x\right) \left(1-p_x\right) + (1-w_x) p_x \left(1-w_x\right) \left(1-p_x\right) + w_x \left(1-p_w\right) \left(1-w_x\right) \left(1-p_x\right) \left(1-w_x\right) \left(1-p_x\right) \approx 2 w (p_w - \bar{p})
\]

where \(p_w\) is the frequency of allele \(A\) among haplotypes with a \(W\) allele. The approximation in the final step follows after the substitution of previous results:

\[
\text{solutionDwy} /. \text{varSubs1} /. \{e_a \rightarrow 0\} // \text{FullSimplify}
\]

The following substitutions help to simplify the equations:

\[
\text{subsCoeffSimple} = \text{Join[}
\{
D_0 [[a_x, s_x]] \rightarrow 0,
\alpha_{(f),f} \rightarrow \alpha_f,
\alpha_{(a),f} \rightarrow \alpha_f,
\alpha_{(f),n} \rightarrow \alpha_n,
\alpha_{(a),n} \rightarrow \alpha_n
\}]
\]

The invasion fitness is calculated as
The average allele frequency \( \bar{p} \) is set by selection and mutation. The balance between these evolutionary forces is reflected in the recursion for \( \bar{p} \) that was derived earlier

\[
\lambda_{\text{TightLinkage1}} = \lambda_{\text{TightLinkage1}} / . (p_w \to \Delta + \bar{p}) / .
\]

\[
\text{Flatten}[\text{Solve}[0 \to \text{simple}\Delta p_w / . (p_w \to \Delta + \bar{p}), \Delta]] / . \text{FullSimplify}
\]

\[
\alpha_x \epsilon_2 (\bar{v}_x \alpha_x + \epsilon_\mu \ (- (1 + \bar{p}) \mu_0 - \bar{p} \mu_1))
\]

\[
\Gamma_{(a,s2)} + \epsilon_\mu \epsilon_\mu \ (\mu_0 + \mu_1)
\]

The average allele frequency \( \bar{p} \) is set by selection and mutation. The balance between these evolutionary forces is reflected in the recursion for \( \bar{p} \) that was derived earlier

\[
\lambda_{\text{TightLinkage1}} = \lambda_{\text{TightLinkage1}} / . (p_w \to \Delta + \bar{p}) / .
\]

\[
\text{Flatten}[\text{Solve}[0 \to \text{simple}\Delta p_w / . (p_w \to \Delta + \bar{p}), \Delta]] / . \text{FullSimplify}
\]

\[
\alpha_x \epsilon_2 (\bar{v}_x \alpha_x + \epsilon_\mu \ (- (1 + \bar{p}) \mu_0 - \bar{p} \mu_1))
\]

\[
\Gamma_{(a,s2)} + \epsilon_\mu \epsilon_\mu \ (\mu_0 + \mu_1)
\]

The average allele frequency \( \bar{p} \) is set by selection and mutation. The balance between these evolutionary forces is reflected in the recursion for \( \bar{p} \) that was derived earlier

\[
\lambda_{\text{TightLinkage1}} = \lambda_{\text{TightLinkage1}} / . (p_w \to \Delta + \bar{p}) / .
\]

\[
\text{Flatten}[\text{Solve}[0 \to \text{simple}\Delta p_w / . (p_w \to \Delta + \bar{p}), \Delta]] / . \text{FullSimplify}
\]

\[
\alpha_x \epsilon_2 (\bar{v}_x \alpha_x + \epsilon_\mu \ (- (1 + \bar{p}) \mu_0 - \bar{p} \mu_1))
\]

\[
\Gamma_{(a,s2)} + \epsilon_\mu \epsilon_\mu \ (\mu_0 + \mu_1)
\]

\[
\text{solve} = \text{finalResultTightLinkage1} - \lambda_{\text{TightLinkage1}} / .
\]

\[
\text{Flatten}[\text{Solve}[\text{simple}\Delta pAvg = 0, \bar{p}] / . \text{FullSimplify}
\]

\[0\]

**Scenario 2: a sex-antagonistic locus that is closely linked to the ancestral sex-determination gene**

The linkage between the ancestral sex-antagonistic locus and the ancestral sex-determination locus is tight if the recombination rate between \( s2 \) and \( a \) is of the same order of magnitude as the selection coefficients.

\[
\lambda_{\text{TightLinkage2}} = \{r_{(a1,a)} \to \bar{v}_a \bar{r}_{(a1,a)}\}
\]

\[
\{r_{(a1,a)} \to \epsilon_\mu \bar{r}_{(a1,a)}\}
\]

Tight linkage with the ancestral sex-determination locus can result in significant differences in the frequencies of sex-antagonistic alleles between the \( x \) and the \( y \)-chromosome. Within the haplotypes that contain an \( X \) allele and within those that contain an \( Y \) allele, the allele frequencies will be roughly equal, irrespective of the presence or absence of an \( W \) allele. This leads us to define a
change of variables that approximates all allele frequencies in terms of the average allele frequencies on the x- and y-chromosome, $p_x$ and $p_y$.

$$\begin{align*}
\text{varSubs2} &= \text{Join}\left[\{p_{XW} \rightarrow p_x + e_a \ e_{pX}, \ p_{YW} \rightarrow p_t + e_a \ e_{pY}\}\right], \\
\text{Flatten}\left[\text{Solve}\left[\left\{p_t = p_{xt}, \ p_n = \frac{p_{xn} + p_t}{2}, \ 3 \ p_x = 2 \ p_{xt} + p_{xn}, \ e_a \ D_a([a, \ \text{sex}], X) = \frac{2}{9} \ (p_{xt} - p_{xn})\right\}\right]\right] // \text{ColumnForm}
\end{align*}$$

$D_a([a, \ \text{sex}], X) = (\frac{2}{9}) (2 \ p_x + 3 \ e_a \ D_a([a, \ \text{sex}], X))$

$D_a([a, \ \text{sex}], X) = (\frac{2}{9}) (p_x + 3 \ e_a \ D_a([a, \ \text{sex}], X))$

The coefficient $D_a([a, \ \text{sex}], X)$ is the association between $a$ and sex-of-carrier (coded as 1 for females and 0 for males) on the x-chromosome. This association is calculated as

$$D_a([a, \ \text{sex}], X) = \frac{2}{9} (p_x - p_{xt})$$

To simplify the expressions later on, we define genetic variances and selection effects specific for the x- and y-chromosome

$$\begin{align*}
\text{definitionCoeffSimple} = \{ \\
V_x \rightarrow p_x \ (1 - p_x), \ V_t \rightarrow p_t \ (1 - p_t), \ a_{xt} \rightarrow \sigma_{xt} \ (\sigma_{xt} + p_x \ (1 - 2 \ h_x) ), \\
\sigma_{xt} \rightarrow \sigma_{xt} \ (\sigma_{xt} + p_x \ (1 - 2 \ h_x) ), \ \sigma_{xn} \rightarrow \sigma_{xn} \ (\sigma_{xn} + p_x \ (1 - 2 \ h_n) ), \\
\sigma_x \rightarrow \sigma_x \ (\sigma_x + p_x \ (1 - 2 \ h_x) ), \ \sigma_t \rightarrow \sigma_t \ (\sigma_t + p_x \ (1 - 2 \ h_t) ) \} // \text{ColumnForm}
\end{align*}$$

The selection coefficients $\sigma$ are related to the sex-chromosome specific effects of allele substitution.
The association between the sex-antagonistic locus and the novel sex-determination factor is again derived from the definition of

$$a_i = \frac{\alpha_i \alpha_s + \alpha_i \alpha_n}{2},$$

for $i \in \{t, n\}$.

If the allele-frequency difference between x- and y-chromosome is large, then it is no longer true that the lowest order term in the

$$p_x = p_{xY} + p_{xW}$$

$$p_n = p_{xY} + p_{xW}$$

If the allele-frequency difference between x- and y-chromosome is large, then it is no longer true that the lowest order term in the expansion for $D[\{a_n, s_{1a}\}]$ vanishes. The association between the sex-antagonistic locus and the ancestral sex-determination factor is now calculated as

$$\text{solutionDyp} = \text{Solve}[[0 = \Delta_{Dyp}, \text{/}. \text{subsCoeffSimple} /\text{/}. \text{tightLinkage2}, 0 = D[\Delta_{Dyp}], \text{/}. \text{subsCoeffSimple} /\text{/}. \text{tightLinkage2}, e_n] /. \{e_n \to 0\}, \{D[\{a_n, s_{1a}\}], D[\{a_n, s_{1a}\}]\}] /\text{Flatten} /\text{FullSimplify}$$

$$[D_n[\{a_n, s_{1a}\}] = D[\{a_n, s_{1a}\}] /\text{Flatten} /\text{FullSimplify}$$

$$\frac{1}{4} (-p_x^2 + p_y) (\alpha disc + \alpha_n (\alpha s_{1a}) + p_x (-1 + p_x) p_{xY} + p_{xW}) = 0$$

The association between the sex-antagonistic locus and the novel sex-determination factor is again derived from the definition of these associations.

$$\text{solutionDwy} = \text{Solve}[[0 = \Delta_{Dwy}, \text{/}. \text{varSubs2} /\text{/}. \{e_n \to 0\}] /\text{FullSimplify}$$

$$[D_w[\{a_n, s_{2a}\}] = -p_{xY} + p_{xW}]$$

With this result, we can now derive an expression for the invasion fitness of the modifier allele.
The latter recursion yields a solution for the difference in allele frequencies between x- and y-chromosomes. In order to solve for this allele-frequency difference, we derive two recurrence relations, one for the average allele frequency of sex-antagonistic alleles, one for the allele-frequency difference between the sex chromosomes.

\[ \text{TightLinkage2} = \frac{\text{GetFrequencyChange}[s2t, \{e_u, 0, 1\}, \{e_u, 0, 1\}]}{2 e_u} \]  
\[ \text{solvedDyp} / \text{solvedDyp} / \text{solvedDyp} \text{// FullSimplify} \]

\[- (p_x - p_y) \alpha_T e_o\]

The invasion fitness is proportional to the allele-frequency difference between x- and y-chromosomes. In order to solve for this allele-frequency difference, we derive two recurrence relations, one for the average allele frequency of sex-antagonistic alleles, one for the allele-frequency difference between the sex chromosomes.

\[ \text{simplifiedDeltaAvg} = \]
\[ \text{Collect}\{\text{Series}[\Delta p_x + \Delta p_y], \text{solvedDyp} / \text{TightLinkage2}, \{e_o, 0, 1\}], \{p_x, p_y\}\} /.
\[ \{p_x \rightarrow p_x + \Delta, p_y \rightarrow p_y + \Delta\} /.
\[ \{p_x \rightarrow 4 \hat{p} - 3 p_x\} \text{// Normal} / \text{FullSimplify} \]

\[ \frac{1}{2} e_o (V_y \alpha_T + 2 V_x \alpha_{x_e} - (-1 + p_x) p_x \alpha_{x_a} - 4 e_u \left((-1 + \hat{p}) \mu_0 + \hat{p} \mu_1\right)) \]

\[ \text{simplifiedDeltaDiff} = \]
\[ \text{Collect}\{\text{Simplify}[\text{Series}[\Delta p_x - \Delta p_y], \text{solvedDyp} / \text{TightLinkage2}, \{e_o, 0, 1\}], \{p_x, p_y\}\} /.
\[ \{p_x \rightarrow p_x + \Delta, p_y \rightarrow p_y + \Delta\} /.
\[ \{p_x \rightarrow 4 \hat{p} - 3 p_x\} \text{// Normal} / \text{FullSimplify} \]

\[ \frac{1}{2} e_o \left(-3 V_y \alpha_T + V_x (2 \alpha_{x_e} + \alpha_{x_a}) - (p_x - p_y) (3 e_u (\mu_0 + \mu_1) + 4 T_{(s1,a)})\right) \]

The latter recursion yields a solution for the difference \( p_x - p_y \).

\[ \text{Flatten}[\text{Solve}[0 = \text{simplifiedDeltaDiff} / \text{solutionDyp} / \text{Simplify}, \{\Delta \rightarrow p_x - p_y\} / \text{FullSimplify} \]

\[ \left\{p_x - p_y \rightarrow -\frac{3 V_y \alpha_T + V_x (2 \alpha_{x_e} + \alpha_{x_a})}{3 e_u (\mu_0 + \mu_1) + 4 T_{(s1,a)}}\right\} \]

Substituting this result in the expression for the invasion fitness yields a final result for this scenario (equation 7 in the main text).

\[ \alpha_T \text{TightLinkage2} = \alpha_T \text{TightLinkage2} / \{p_x \rightarrow p_x + \Delta\} /.
\[ \text{Flatten}[\text{Solve}[0 = \text{simplifiedDeltaDiff} / \{p_x \rightarrow p_x + \Delta\}, \Delta]\} / \text{FullSimplify} \]

\[ \frac{\alpha_T (3 V_y \alpha_T - V_x (2 \alpha_{x_e} + \alpha_{x_a}) e_o)}{3 e_u (\mu_0 + \mu_1) + 4 T_{(s1,a)}} \]

\[ \text{finalResultTightLinkage2} = \frac{\alpha_T (3 V_y \alpha_T - V_x (2 \alpha_{x_e} + \alpha_{x_a})) e_o^2}{3 e_u e_v (\mu_0 + \mu_1) + 4 T_{(s1,a)}} \]

In the final part of this section, we examine the effect of deleterious alleles that may have accumulated on the y-chromosome. We consider a worst case scenario for the invasion of the modifier allele: the deleterious alleles are recessive and expressed only in females. The rate of deleterious mutations is \( \mu_0 \), and the reverse mutation rate is assumed to be negligibly small.

\[ \{\text{deleteriousAllele} = \{\alpha_{x_e} \rightarrow -v p_x, \alpha_T \rightarrow -v \hat{p}, \alpha_{x_a} \rightarrow 0, \alpha_{x_e} \rightarrow 0, \alpha_T \rightarrow 0, \mu_1 \rightarrow 0\}\} / \text{FullSimplify} \]

\[ \alpha_{x_e} \rightarrow -v p_x \]

\[ \alpha_T \rightarrow -v \hat{p} \]

\[ \alpha_{x_a} \rightarrow 0 \]

\[ \alpha_{x_e} \rightarrow 0 \]

\[ \alpha_T \rightarrow 0 \]

\[ \mu_1 \rightarrow 0 \]

Under these conditions, an explicit solution can be found for the recursions for \( p_x \) and \( p_y \).
A nonzero invasion fitness is realized only if the sex-antagonistic locus is tightly linked to the sex determination locus. Also the expression for the transition matrix has to be re-evaluated.

Analyzing homologous sex-determination mutations is equation (11) in the main text. The sex-determination factors and are no longer distinct loci. Accordingly the recombination rate between and is set to zero.

The loose-linkage results are independent of the recombination rate between the sex-determination loci. Loose-linkage results for homologous mutations can therefore be easily derived from the earlier loose-linkage approximation.

Substituting this result into the expression for the invasion fitness yields:

\[
\text{Series} [\lambda \text{TightLinkage2} / . \text{deleteriousAllele} / . \text{definitionCoeffSimple} / . \text{mutationSelectionBalance} / . \tilde{F}(s_1,a) \rightarrow \epsilon_{\mu} \tilde{F}(s_1,a) / . \epsilon_{\mu} \rightarrow \epsilon_{\mu}^2, \{\epsilon_{\mu}, 0, 0\}] \text{ // FullSimplify}
\]

which is equation (11) in the main text.

Scenario 3: a homologous mutant

This section concentrates on homologous sex-determination mutations. The evolutionary consequences of such mutations can be analyzed in our model by choosing the order of the genes on the chromosomes as

\[
geneOrder = \{s_1, s_2, a\}
\]

The sex-determination factors and are no longer distinct loci. Accordingly the recombination rate between and is set to zero.

The loose-linkage results are independent of the recombination rate between the sex-determination loci. Loose-linkage results for homologous mutations can therefore be easily derived from the earlier loose-linkage approximation.

\[
\text{finalResultWeakLinkage} / . \{\tilde{r}(s,a) \rightarrow \tilde{r}(s,a) / . \tilde{r}(s_1,a) \rightarrow \tilde{r}(s_1,a)\} \text{ // Simplify}
\]

A nonzero invasion fitness is realized only if the sex-antagonistic locus is tightly linked to the sex determination locus.

\[
\text{tightLinkage3} = \{\tilde{r}(s,a) \rightarrow \epsilon_{\mu} \tilde{r}(s,a)\}
\]

Since the order of the genes on the chromosome is different from what is was before, the recombination matrices have to be recalculated.
We now verify that these are the leading eigenvectors. We know that the leading eigenvalue of a dominant right eigenvector of \( \mathbf{T} \) is

\[
\lambda = \log[R] = \epsilon_u \frac{d \mathbf{T}}{d \epsilon_u} \mathbf{v},
\]

such that the invasion fitness is approximately

\[
\lambda = \log[R] = \epsilon_u \frac{d \mathbf{T}}{d \epsilon_u} \mathbf{v}
\]

In these expressions, \( \mathbf{T}_0 = \lim_{\epsilon_u \to 0} \mathbf{T} \), \( \mathbf{u} \) and \( \mathbf{v} \) are the dominant eigenvectors of \( \mathbf{T}_0 \) normalized such that \( \mathbf{u} \mathbf{v} = 1 \). The matrix derivative \( \frac{d \mathbf{T}}{d \epsilon_u} \) is evaluated at \( \epsilon_u = 0 \). \( \mathbf{T}_0 \) is:

\[
\mathbf{T}_0 = \mathbf{T} / . \text{varSubs3} / . \text{tightLinkage3} / . \{ \epsilon_u \to 0 \} // \text{Simplify}
\]

\[
\{ [1, 0, 0, 0], [0, 1, 0, 0], [0, 0, 1, 0], [0, 0, 0, 1] \}
\]

A dominant right eigenvector of \( \mathbf{T}_0 \) is given by

\[
\text{rightEv} = \{ \kappa (1 - \mu_u), \kappa \mu_u, (1 - \kappa) (1 - \mu_u), (1 - \kappa) \mu_u \};
\]

Likewise, a left eigenvector of \( \mathbf{T}_0 \) is:

\[
\text{leftEv} = \{ 1, 1, 1 \};
\]

Check that \( \text{rightEv} \) and \( \text{leftEv} \) are truly the right and left eigenvectors of \( \mathbf{T}_0 \) and confirm that the vectors are properly normalized

\[
\text{leftEv}.\mathbf{T}_0 - \text{leftEv} // \text{Simplify}
\]

\[
\mathbf{T}_0.\text{rightEv} - \text{rightEv} // \text{Simplify}
\]

\[
\text{rightEv}.\text{leftEv} // \text{Simplify}
\]

\[
\{ 0, 0, 0, 0 \}
\]

\[
\{ 0, 0, 0, 0 \}
\]

\[
1
\]

We now verify that these are the leading eigenvectors. We know that the leading eigenvalue of \( \mathbf{T}_0 \) is unity, which agrees with our eigenvectors:

\[
\mathbf{r}_0 = \text{leftEv}.\mathbf{T}_0.\text{rightEv} // \text{Simplify}
\]

\[
1
\]

As a preparatory step to calculating the fitness gradient, we first determine the mean fitness of males and females. To do this, we recalculate the association between the sex-determination locus and the sex-antagonistic locus

\[
\Delta \text{Dyp}_\alpha = \text{GetAssociationChange}([\{a_\alpha, s_1\}], \{\epsilon_u, 0, 1\}, \{\epsilon_u, 0, 0\})
\]

\[
\frac{1}{4} \{(p_x - p_\alpha) (1 + 2 p_x + 3 e_u \Delta_0 \{ \{a, sex\}, x \} ), p_x \to p_x - 3 e_u \Delta_0 \{ \{a, sex\}, x \}, p_x \to \frac{1}{2} \{ p_x + 3 e_u \Delta_0 \{ \{a, sex\}, x \} \};
\]

The invasion fitness of the mutant allele (i.e., its geometric rate of increase while it is rare) can be calculated from the dominant eigenvalue of the matrix \( \mathbf{T} \). For weak selection, this eigenvalue can be approximated as (H. Caswell (2001). Matrix Population Models. Sinauer Associates, Sunderland MA, USA.)

\[
R = \mathbf{u} \mathbf{T}_0 \mathbf{v} + \epsilon_u \frac{\partial \mathbf{T}}{\partial \epsilon_u} \mathbf{v} = 1 + \epsilon_u \frac{\partial \mathbf{T}}{\partial \epsilon_u} \mathbf{v},
\]

where \( d \mathbf{T} / d \epsilon_u \) is evaluated at \( \epsilon_u = 0 \).
Before proceeding, we must now calculate the mean fitness, which, so far, appears in the matrix $A_e$ as an unspecified coefficient $\bar{v}_t$.

$$\text{meanFitnessSubs} = \text{Simplify} \left[ (\bar{V}_t \to \bar{v}_t, \bar{V}_n \to \bar{v}_n) \right] / \text{.subsCoeffA / .SexDetermination / .Linearize / . \{e_\mu \to 0\} / \text{.solutionDyp / . (varSubs3 / . \{e_\mu \to 0\} / . \text{FullSimplify} \{v_t \to -(2 h_t (1 - p_X) + p_X v_t + v_n \to (p_X + p_Y) (1 - (2 p_X + p_Y))) v_n \} \right]$$

To calculate the fitness gradient, we take the derivative of $T$ with respect to $e_\mu$,

$$T_t = e_\mu \left( D_{[T / \varSubs3 / \text{.meanFitnessSubs / .tightLinkage3, e_\mu] / \{e_\mu \to 0\}} \right)$$

and then multiply on the right and left side with the eigenvectors.

$$\lambda \text{TightLinkage3Tmp} = \text{leftEv.T_1.rightEv // FullSimplify}$$

$$\frac{1}{2} \left( -2 p_X^2 + p_w (p_X + p_Y) + h_t (1 - 3 + 4 p_X) + p_Y - 2 p_w (-1 + p_X + p_Y)) \right) v_t e_\mu$$

The resulting expression is simplified to

$$\lambda \text{TightLinkage3} = \left( \frac{1}{2} \alpha_{a_t} (p_Y - p_X) + (-p_w - p_X) \alpha_{a_t} \right) e_\mu$$

$$\left( \alpha_{a_t} (p_Y - p_X) + \frac{1}{2} (-p_X + p_Y) \alpha_{a_t} \right) e_\mu$$

where $\alpha_{a_t} = e_t (h_t + p_Y (1 - 2 h_t))$ is the effect of allele substitution for sex-antagonistic alleles on the $w$-chromosome. The homologue of the $w$-chromosome in mutant females is a $x$- or $y$-chromosome, and each occur with probability one half. Accordingly, the frequency of sex-antagonistic alleles on the homologue is given by $p_Y = \frac{1}{2} (p_X + p_Y)$.

$$\text{check} = \lambda \text{TightLinkage3Tmp} \to \lambda \text{TightLinkage3} / \{e_{a_t} \to v_t (h_t + p_X (1 - 2 h_t))\}$$
$$p_Y \to \frac{p_X + p_Y}{2} \} / \text{.definitionCoeffSimple // FullSimplify}$$

$0$

As before, a final result is obtained by solving for the allele-frequency differences between different types of sex chromosomes. The variable $p_w$ that appears in the expressions for the fitness represents the frequency of the sex-antagonistic allele 1 on the mutant sex chromosome. An equilibrium solution for $p_w$ can be obtained from the recursion

$$\text{Series} \left[ \{0, 1, 0, 1\}.T.rightEv - p_w / \text{.varSubs3 / .tightLinkage3,} \right.$$
$$(1, 1, 1, 1).T.rightEv \{e_{a_t}, 0, 1\}, \{e_{\mu}, 0, 1\} \} / \text{Normal // FullSimplify}$$

$$\frac{1}{2} e_\mu (2 e_\mu \mu_0 +$$
$$p_w ((-1 + p_X) (-p_X - p_Y + 2 h_t (-1 + p_X + p_Y)) v_t - 2 e_\mu (\mu_0 + \mu_1)) - (2 p_w - p_X - p_Y) \bar{X}_{(a, a)}$$

Rewriting this recursion yields a simpler expression

$$\text{simpleDelta} = e_\mu (V_0 \alpha_{a_t} + (p_w - p_X) \bar{X}_{(a, a)} + e_{\mu} ((1 - p_w) \mu_0 - p_w \mu_1))$$
$$e_\mu (V_0 \alpha_{a_t} + e_{\mu} (1 - p_w) \mu_0 - p_w \mu_1) - (p_w - p_X) \bar{X}_{(a, a)}$$
The equilibrium solution for the difference \( p_W - p_T \) is obtained from the recursion for \( p_W \):

\[
eq_{\Delta W} = p_W - p_T . Flatten[Solve[simple\Delta p_w = 0, p_W]] \cdot \{p_T \to \frac{P_X + P_T}{2}\} \quad \text{// FullSimplify}
\]

An expression for the allele-frequency difference between the ancestral sex chromosomes can be derived from the recursion

\[
simple\Delta p_{\text{diff}} = \text{Collect}[\text{Simplify}[\text{Series}[\Delta p_{\text{diff}}, \{\epsilon_\mu, 0, 1\}], \{p_T, p_T\}] \cdot \{p_T \to p_T + \Delta p_{\text{diff}}, \Delta p_{\text{diff}}\}]
\]

This solution is substituted into the expression for the invasion fitness

\[
\lambda_{\text{tightLinkage3}} = \lambda_{\text{tightLinkage3}} \cdot \{p_W \to p_T + eq_{\Delta W}\} \cdot \{p_T \to p_T - \Delta p_{\text{diff}}\} \cdot \text{// FullSimplify}
\]

and the resulting expression is slightly rewritten. This following solution (equation (13) in the main text) is for small mutation rate (\( \epsilon_\mu \to 0 \)), but can be extended, if desired.

\[
\text{finalResultTightLinkageHomologous} = \epsilon_\alpha^2 (1 - 2 \sigma_{\epsilon}) - \frac{V_u \alpha_f}{4} \frac{3 Y \alpha_y - Y (2 \alpha_y + \alpha_u)}{\sigma_{\epsilon}} \sigma_{\epsilon} ;
\]

A consistency check for this result is obtained by a comparison with the weak-linkage approximation. If linkage is weak \( \alpha_f \) and \( \alpha_u \) converge to \( \alpha_u \), and \( \alpha_y \) converges to \( \alpha_y \). Similarly, the chromosome specific variances approach the average variance. Under these circumstances, the invasion fitness should reduces to zero, as predicted by the weak-linkage results
5. FIXATION OF THE MODIFIER

Successful invasion of the modifier allele does not necessarily imply that the ancestral sex determination system will eventually be lost. In some cases, a protected polymorphism of sex-determination factors could be established. This occurs if the rare ancestral sex-determination allele can increase in frequency when the novel W-allele has nearly reached its maximal frequency. If W is dominant over Y, such that XYZW and YYZW develop as females, then a completed heterogamety switch would lead to a population with ZZW sex determination and a loss of the ancestral X-chromosome. The following analysis therefore considers the fitness of the X allele in a population where female heterogamety is the predominant mode of sex determination. If the invasion fitness of the X is negative, then we expect the heterogamety transition to continue, leading to a complete loss of the X. If its fitness is positive, a stable polymorphism of sex-determination factors will be established.

The analysis is complicated somewhat by the fact that the X is recessive in the context of a population with epistatically dominant female heterogamety. X has a feminizing effect only if it occurs in a homozygous state, and only in individuals that are homozygous for the Z allele at the novel sex-determination locus. We can deal with this problem by expanding the expressions for the associations up to second order in the frequency of the X allele. The allele frequencies in gametes are close to one for the Y allele, and close to one half for the W allele in female gametes. The frequency of W in sperm is zero.

\[
\text{Linearize} = \{D_1[K\_List] \Rightarrow \\
\quad \text{Module}[\{n = \text{Count}[K, s1_1] + \text{Count}[K, s1_n], \ k = \text{Count}[K, s2_n]\}, \\
\quad \text{If}[k = 0, \text{If}[n = 0, D_0[K] + e_n D_a[K] + e^2_n (D_x[K] + e_n D_{ax}[K]) + e^3_n (1 + e_n) e_k, \\
\quad \quad e^2_n (D_x[K] + e_n D_{ax}[K]) + e^3_n (1 + e_n) e_k], \ e^3_n (1 + e_n) e_k]], \\
\quad \ w_a = 0, w_f = \frac{1}{2} - e^2_x e_{u_x}, \ y_f = 1 - e_x \bar{x} - e^2_x \delta_x, \ y_a = 1 - e_x \bar{x} + e^2_x \delta_x\};
\]

It is necessary to recalculate the sex-determination coefficients \(\sigma_k\) for the new population.

\[
\text{GetEquationS}[\text{Genotype} \_\text{List}, \text{fractionFemale}_\_] := \\
\text{Module}[\{K = \text{Subsets}[\$], \text{lhs} = 1, i\}, \\
\text{Do}[\text{lhs} += \sigma_k[[i]] (\text{GetMoment}[K[[i]], \text{Genotype}, \$] - D_i[K[[i]]]), \\
\{i, 1, \text{Length}[K]\}]]; \\
(\text{Normal}[\text{Series}[\text{lhs} /. \text{ExpandIntersexualAssociations} /. \text{subsReferenceValues} /. \text{Linearize} /. \{e_n \to 0\}, \{e_x, 0, 1\}]] = \text{fractionFemale} / \$]
\]
As before, we proceed in three steps. First we calculate the associations between the sex determination alleles, then the associations with the sex-antagonistic alleles, and finally we substitute the QLE values for the associations into an expression for the invasion fitness of the X allele. In what follows we present only the main derivation, omitting some of the explanatory intermediary steps from the previous analysis.

The parameters are changed back to the default values.

\[
\text{geneOrder} = \{s1, a, s2\}
\]

\[
\text{recombinationRate} = \{\text{r}_{(s1, a)}, \text{r}_{(a, s2)}\}
\]

Sex-ratio selection

Two recursions specify the distribution of the X allele in male and female gametes in the absence of sex-antagonistic selection \((e_a \rightarrow 0)\). In that case the novel sex-determination allele is neutral. Its total frequency neither increases nor decreases.

\[
\Delta y_f = \text{GetFrequencyChange}[s1_f, \{e_a, 0, 0\}, \{e_x, 0, 2\}]
\]

\[
\Delta y_m = \text{GetFrequencyChange}[s1_m, \{e_a, 0, 0\}, \{e_x, 0, 2\}]
\]

\[
e_x^2 \left( -x^2 + \delta_x + \Delta_x \left[ \{s1_f, s2_f\} \right] \right)
\]

\[
e_x^2 \left( x^2 - \delta_x - \Delta_x \left[ \{s1_f, s2_f\} \right] \right)
\]

\[
\Delta y_f + \Delta y_m // \text{Simplify}
\]

0

An additional recursion gives the change of the genetical association between the sex-determination loci in females.
\[ \Delta \text{Dwy}_f = \text{GetAssociationChange}([s_{1f}, s_{2f}], \{e_a, 0, 0\}, \{e_s, 0, 2\}) \]

\[ \frac{1}{2} e_s^2 (x^2 - (-1 + 2 r_{a,s2})) (-1 + 2 r_{s_{1f}, s_{1a}}) \delta_x + \]

\[ (-1 - 2 r_{s_{1f}, s_{1a}} + r_{a,s2}) (-2 + 4 r_{s_{1f}, s_{1a}})) D_x ([s_{1f}, s_{2f}]) \]

Equilibrium is attained for the following value of the association and the difference between the frequency of the \( X \) allele in female and male gametes, \( \delta_x = x_t - x_m : \)

\text{solutionDwy} = \text{Solve}([\Delta \text{Dwy}_f = 0, \Delta y_f - \Delta y_m = 0], [D_x([s_{1f}, s_{2f}]), \delta_x]) // \text{Flatten} // \text{FullSimplify}

\[ \{ D_x ([s_{1f}, s_{2f}]) \rightarrow x^2 / 2, \delta_x \rightarrow x^2 / 2 \} \]

Sex-antagonistic selection

Recurrence relations for the sex-antagonistic allele frequencies are given by

\[ \Delta p_t = \text{GetFrequencyChange}([a_t, \{e_a, 0, 1\}, \{e_s, 0, 1\}] \]

\[ \frac{1}{2} (-p_t + p_n) + D_0 ([a_t, s_{2f}] ) + \frac{1}{2} e_s (-p_t^2 \alpha_{(t), t} - (-1 + p_n) p_n \alpha_{(n), t} - p_n e_u (\mu_0 + \mu_1) -
\]

\[ 2 (-\alpha_{(t), t} + (-1 + p_n) p_n \alpha_{(n), t} + e_u (\mu_0 + \mu_1) ) D_0 ([a_t, s_{2f}]) -
\]

\[ 4 \alpha_{(t), t} D_0 ([a_t, s_{2f}])^2 - p_t (e_u (\mu_0 + \mu_1) + \alpha_{(t), t} (-1 + 4 D_0 ([a_t, s_{2f}]))) +
\]

\[ 2 (e_u, \mu_0 + p_n ([a_t, s_{2f}])) \]

\[ \Delta p_n = \text{GetFrequencyChange}([a_n, \{e_a, 0, 1\}, \{e_s, 0, 1\}] \]

\[ \frac{1}{2} (-p_t - p_n - 2 D_0 ([a_t, s_{2f}])) +
\]

\[ \frac{1}{2} e_s (-p_t^2 \alpha_{(t), n} - (-1 + p_n) p_n \alpha_{(n), n} + 2 e_u \mu_0 - p_n e_u (\mu_0 + \mu_1) +
\]

\[ 2 (-\alpha_{(n), n} + (-1 + p_n) p_n \alpha_{(n), n} + e_u (\mu_0 + \mu_1) ) D_0 ([a_t, s_{2f}]) -
\]

\[ 4 \alpha_{(n), n} D_0 ([a_t, s_{2f}])^2 +
\]

\[ p_r (-e_u (\mu_0 + \mu_1) + \alpha_{(n), n} (1 + 4 D_0 ([a_t, s_{2f}])) - 2 D_0 ([a_t, s_{2f}])) \]

Again only one genetical association appears in these recurrences. This time it is the association between the sex-antagonistic allele and the \( W \)-allele in female gametes. This association changes from one generation to the next as given by the recurrence

\[ \Delta \text{Dwp}_f = \text{GetAssociationChange}([a_t, s_{2f}], \{e_a, 0, 1\}, \{e_s, 0, 1\}] \]

\[ \frac{1}{4} (-p_t - p_n) (-1 + 2 r_{a,s2}) - 2 (1 + 2 r_{a,s2} ) D_0 ([a_t, s_{2f}]) +
\]

\[ \frac{1}{4} e_s (-p_t^2 (-1 + 2 r_{a,s2}) \alpha_{(t), t} + (-1 + 2 r_{a,s2}) (-p_n (-1 + p_n) \alpha_{(n), t} + e_u (\mu_0 + \mu_1) ) +
\]

\[ 2 (-\alpha_{(t), t} - (-1 + p_n) p_n \alpha_{(n), t} + e_u (\mu_0 + \mu_1) ) D_0 ([a_t, s_{2f}]) +
\]

\[ 4 \alpha_{(t), t} D_0 ([a_t, s_{2f}])^2 + p_r (-1 + 2 r_{a,s2})
\]

\[ (e_u (\mu_0 + \mu_1) + \alpha_{(t), t} (-1 + 4 D_0 ([a_t, s_{2f}]))) - 2 (1 + 2 r_{a,s2} ) D_0 ([a_t, s_{2f}]) \]

If the difference between \( p_t \) and \( p_n \) is small, then also the association between \( a \) and \( s_2 \) will be small. Indeed, the leading order term in the expansion of this association is of order \( e_a \):

\text{Solve}([0 = \Delta \text{Dwp}_f /. \text{varSubs} /. \{e_a \rightarrow 0\}, D_0 ([a_t, s_{2f}])]

\[ \{ [D_0 ([a_t, s_{2f}]) \rightarrow 0 ] \}

To simplify the expressions, we use the same notation as before
All associations involving depends on the genotype at the novel sex-determination locus. Previous results are used to simplify the expressions. Simplified recurrence equations for the mean allele frequency \( \bar{p} \), the allele frequency difference \( p_t - p_m \), and the association \( D_a([a_t, s2f]) \) are given by

```math
\begin{align*}
\text{simplifyDeltaAvg} = & \quad \text{Series}[\text{Collect}[\Delta p_t + \Delta p_m / \text{subsCoeffSimple}, \bar{p}] / . \{p^2 \rightarrow \bar{p} - \bar{v}, \{e_a, 0, 1\}] // \\
\text{Normal} // \text{FullSimplify}
\end{align*}
```

For the associations between the sex-antagonistic locus and the ancestral sex-determination factor, we have to solve for three

The equilibrium solution for the associations is

```math
\begin{align*}
\text{simplifyDeltaDiff} = & \quad \text{Series}[\text{Collect}[\Delta p_t - \Delta p_m / \text{subsCoeffSimple}, \bar{p}] / . \{p^2 \rightarrow \bar{p} - \bar{v}, \{e_a, 0, 1\}] // \\
\text{Normal} // \text{FullSimplify}
\end{align*}
```

For the associations between the sex-antagonistic locus and the ancestral sex-determination factor, we have to solve for three
different associations: the pairwise associations between the sex-antagonistic allele and the ancestral sex-determination allele
in male and female gametes, \( D_a([a_t, s1f]) \) and \( D_a([a_t, s1m]) \), as well as the three-way association
\( D_x([a_t, s1f, s2f]) \). The latter appears because the \( W \) allele is dominant such that the phenotypic effect of the \( X \) allele
depends on the genotype at the novel sex-determination locus.

\[ \Delta D_{ypf} = \text{GetAssociationChange}([a_t, s1f], \{e_a, 0, 1\}, \{e_x, 0, 2\}); \]

\[ \Delta D_{ypm} = \text{GetAssociationChange}([a_n, s1n], \{e_a, 0, 1\}, \{e_x, 0, 2\}); \]

\[ \Delta D_{wpf} = \text{GetAssociationChange}([a_t, s1f, s2f], \{e_a, 0, 1\}, \{e_x, 0, 2\}); \]

All associations involving \( s2f \) and \( a \) are of the order \( e_a \) when linkage is weak, and the lowest order terms in the expansion of these associations vanish.

\[ \text{simplifyDelta0} = \quad \text{Solve}[\{\Delta D_{ypf} = 0, \Delta D_{ypm} = 0, \Delta D_{wpf} = 0\} / . \text{subsCoeffSimple} / . \{e_a \rightarrow 0\}, \\
\{D_a([a_t, s1f]), D_a([a_n, s1n]), D_x([a_t, s1f, s2f])\}[1]] // \text{ColumnForm} \]

Previous results are used to simplify the expressions.
subsCoeffSimple = Join[subsCoeffSimple, solutionDwp0];
subsPreviousResults :=
eq_\rightarrow (eq /. subsCoeffSimple /. Flatten[Solve[simpleDwp_\_\_ == 0, D_a[[a, sex]]]]) /. solutionDwy

solutionDyp = Solve[{0 = Series[\Delta Dyp_\_\_ /. subsPreviousResults, \{e_\_\_a, 0, 1\}],
0 = Series[\Delta Dyp_\_\_a /. subsPreviousResults, \{e_\_\_a, 0, 1\}],
0 = Series[\Delta Dyp_\_\_e /. subsPreviousResults, \{e_\_\_a, 0, 1\}],
{D_\_\_o \[\{a_\_\_e, s_1f\}], D_\_\_o \[\{a_\_\_a, s_1m\}], D_\_\_o \[\{a_\_\_e, s_1f, s_2f\}]}[[1]] // FullSimplify

[D_\_\_o \[\{a_\_\_e, s_1f\}] ->
- (2 x^2 (2 \Gamma_\{a, s_2\} (1 - 3 \Gamma_\{s_1, a\}) \Gamma_\{s_1, a\} - (-3 + \Gamma_\{s_1, a\}) \Gamma_\{s_1, a\} + 4 \Gamma_\{a, s_2\} (-1 + \Gamma_\{s_1, a\})
+ (1 + (-1 + \Gamma_\{s_1, a\}) \Gamma_\{s_1, a\}) D_0 \[\{a_\_\_e, s_2f\}]) / ((-1 + 2 \Gamma_\{a, s_2\}) \Gamma_\{s_1, a\})
(2 \Gamma_\{a, s_2\} (-2 + \Gamma_\{s_1, a\}) (-1 + \Gamma_\{s_1, a\}) - (-3 + \Gamma_\{s_1, a\}) \Gamma_\{s_1, a\} )] // FullSimplify

[D_\_\_o \[\{a_\_\_e, s_1f, s_2f\}] ->
2 x^2 \Gamma_\{a, s_2\} (-1 + \Gamma_\{s_1, a\}) (-1 + 2 \Gamma_\{s_1, a\}) D_0 \[\{a_\_\_e, s_2f\}]
2 \Gamma_\{a, s_2\} (-2 + \Gamma_\{s_1, a\}) (-1 + \Gamma_\{s_1, a\}) - (-3 + \Gamma_\{s_1, a\}) \Gamma_\{s_1, a\} )

The three-way linkage disequilibrium vanishes in the case of a non-homologous transition. In that case, the solution simplifies to

(solutionDypNonHomologous =
Solve[{0 = Series[\Delta Dyp_\_\_ /. subsPreviousResults, \{e_\_\_a, 0, 1\}],
0 = Series[\Delta Dyp_\_\_a /. subsPreviousResults, \{e_\_\_a, 0, 1\}],
0 = D_\_\_o \[\{a_\_\_e, s_1f, s_2f\}], D_\_\_o \[\{a_\_\_e, s_1f\}], D_\_\_o \[\{a_\_\_a, s_1m\}],
D_\_\_o \[\{a_\_\_e, s_1f, s_2f\}]}[[1]] // FullSimplify] // ColumnForm

[D_\_\_o \[\{a_\_\_e, s_1f\}] ->
2 x^2 \Gamma_\{a, s_2\} (-1 + \Gamma_\{s_1, a\}) (-1 + 2 \Gamma_\{s_1, a\}) D_0 \[\{a_\_\_e, s_2f\}] / (-1 - 2 \Gamma_\{a, s_2\}) \Gamma_\{s_1, a\}

[D_\_\_o \[\{a_\_\_a, s_1m\}] ->
2 x^2 \Gamma_\{a, s_2\} (-1 + \Gamma_\{s_1, a\}) (-1 + 2 \Gamma_\{s_1, a\}) D_0 \[\{a_\_\_e, s_2f\}] / (-1 - 2 \Gamma_\{a, s_2\}) \Gamma_\{s_1, a\}

[D_\_\_o \[\{a_\_\_e, s_1f, s_2f\}] -> 0

Invasion Fitness

During a non-homologous transition, the mean frequency of the X allele changes from one generation to the next according to the recursion

\[ \Delta X = \frac{1}{2} \left( \text{GetFrequencyChange}[s_1f, \{e_\_\_a, 0, 2\}] + \text{GetFrequencyChange}[s_1m, \{e_\_\_a, 0, 2\}] \right) / \text{subsPreviousResults} / \text{solutionDypNonHomologous} / \text{solutionDwp} / \text{FullSimplify} \]

Rewriting this expression yields the following result for the invasion fitness of the X allele (equation 9 in the main text):

\[ \text{finalResultFixationWeakLinkage} = \frac{1}{2} e_\_\_x \hat{x} e_\_\_x \left( -\Gamma_\{s_1, a\} \Gamma_\{s_1, a\} \frac{\text{V \_ \_ \_ \_}}{2 \Gamma_\{a, s_2\} \Gamma_\{s_1, a\}} \frac{\alpha_\_\_e + \alpha_\_\_a}{2} \right) \]

\[ \text{check} = \text{finalResultFixationWeakLinkage} - \frac{\Delta X}{e_\_\_x} / \text{Simplify} \]

0