

# Genetic signatures of evolutionary rescue by a selective sweep

Matthew M Osmond & Graham Coop

Center for Population Biology and Department of Evolution and Ecology, University of California - Davis  
mmosmond@ucdavis.edu

---

## Prerequisites

### Directories

```
In[75]:= SetDirectory[NotebookDirectory[]];  
(*sets current directory to be location of this file*)  
imagedir = "../plots/";(*where to save plots*)  
datadir = "../simulations/";(*where simulation data is*)
```

### Functions (derived in Methods below, and in manuscript)

#### Recursions

```
In[78]:= wbar = (1 - d) (1 - q[t])^2 + 2 (1 - d) (1 + h s) (1 - q[t]) q[t] + (1 - d) (1 + s) q[t]^2;  
qnew = 
$$\frac{q[t] (-1 - h s + (-1 + h) s q[t])}{-1 - 2 h s q[t] + (-1 + 2 h) s q[t]^2};$$
  
qtadditive = 
$$\left\{ q[t] \rightarrow \frac{e^{\frac{s t}{2}} q_0}{1 + \left( -1 + e^{\frac{s t}{2}} \right) q_0} \right\};$$
  
ntadditive = 
$$\left\{ n[t] \rightarrow e^{-d t} N_0 \left( 1 + \left( -1 + e^{\frac{s t}{2}} \right) q_0 \right)^{2-2 d} \right\};$$
  
ntadditiveClonal = 
$$\left\{ \left\{ n[t] \rightarrow e^{-d t} N_0 \left( 1 + \left( -1 + e^{\frac{s t}{2}} \right) q_0 \right)^{1-d} \right\} \right\};$$

```

#### Probability of establishment

```
In[83]:= pest = 1 - Exp[-2 (w - 1) / v];
```

## Rescue from standing genetic variance

$$\begin{aligned} \text{In}[84]:= \text{PSGV} &= 1 - (1 - \rho)^{\kappa}; \\ \text{q0rescueSGV} &= \frac{\kappa}{2 N_0} \frac{1}{\text{PSGV}}; \end{aligned}$$

## Rescue from de novo mutation

$$\begin{aligned} \text{In}[86]:= \text{PDNM} &= 1 - e^{-\frac{2 N_0 \rho u}{d}}; \\ \text{q0rescueDNM} &= \frac{(-c)^{-\frac{\epsilon}{d}} \left( -\text{Gamma}\left[\frac{d+\epsilon}{d}\right] + \text{Gamma}\left[\frac{d+\epsilon}{d}, -c\right] \right)}{2 (-1 + e^c) N_0 \rho} /. c \rightarrow u 2 N_0 \rho / d; \\ \text{q0sweepDNM} &= \frac{1}{2 N_0 \rho}; \end{aligned}$$

## Rescue from migration

$$\begin{aligned} \text{In}[89]:= \text{PMIG} &= 1 - N_0^{-\frac{m \rho}{d}}; \\ \text{q0rescueMIG} &= \frac{1}{2 N_0} \frac{1}{\rho} \frac{m \rho \left( 1 - N_0^{-\frac{\epsilon + m \rho}{d}} \right)}{(\epsilon + m \rho) \left( 1 - N_0^{-\frac{m \rho}{d}} \right)}; \\ \text{q0sweepMIG} &= \frac{1}{2 N_0 \rho}; \end{aligned}$$

## Structured coalescent

```

In[92]:= pcoal[k_, τ_] := Binomial[k, 2]  $\frac{1}{2 n[\tau] x[\tau]}$ 

preco[k_, τ_] := k  $\frac{2 n[\tau] r (1 - x[\tau])}{2 n[\tau]}$ 

pmut[k_, τ_] := k  $\frac{2 n[\tau] u (1 - x[\tau])}{2 n[\tau] x[\tau]}$ 

pmig[k_, τ_] := k  $\frac{m}{2 n[\tau] x[\tau]}$ 

coalrate[n_, p_, τ_, l_, NeN_] :=  $\frac{1 (1 - l)}{4 n[[\tau]] NeN p[[\tau]]}$ 
coalEmp[n_, p_, l_, NeN_] := Table[coalrate[n, p, τ, l, NeN], {τ, Length[p]}};
(*prob of coal in gen τ*)

recreate[p_, τ_, l_, r_] := l r (1 - p[[τ]])
recEmp[p_, l_, r_] := Table[recreate[p, τ, l, r], {τ, Length[p]}};
(*prob of rec in gen τ*)

mutrate[p_, τ_, l_, u_] :=  $\frac{1 u (1 - p[[\tau]])}{p[[\tau]]}$ 
mutEmp[p_, l_, u_] := Table[mutrate[p, τ, l, u], {τ, Length[p]}};
(*prob of mut in gen τ*)

migrate[l_, m_, n_, τ_, p_] :=  $\frac{1 m}{2 n[[\tau]] p[[\tau]]}$ 
migEmp[n_, p_, l_, m_] := Table[migrate[l, m, n, τ, p], {τ, Length[p]}};
(*prob of mig in gen τ*)

```

## Backward-time dynamics

```

In[104]:= tfixadditive =  $\frac{2 \text{Log}\left[\frac{(-1+q0) qf}{q0 (-1+qf)}\right]}{s}$ ;

qtadditiveback[τ_] :=  $\frac{qf}{e^{\frac{s \tau}{2}} (1 - qf) + qf}$ 

ntadditiveback[τ_] :=  $e^{d \tau} N0 \left(\frac{(1 - q0) qf}{q0 (1 - qf)}\right)^{-\frac{2d}{s}} \left(1 - q0 + \frac{e^{-\frac{s \tau}{2}} (1 - q0) qf}{1 - qf}\right)^{2-2d}$ 

```

## Probability of no events

$$\begin{aligned}
\text{ln}[107] := & \frac{1}{4 \text{N0 NeN} (-1+q0)^2 qf^3} (-1+k) k (-1+qf)^2 \left( \frac{(-1+q0) qf}{q0 (-1+qf)} \right)^{\frac{2d}{s}} \\
& \left( \frac{-e^{\frac{s\tau}{2}} (-1+qf) + qf}{qf^2} \right)^{-2d} \left( \frac{1}{d-s+d s} qf \left( \frac{(-1+q0) \left( -e^{\frac{s\tau}{2}} (-1+qf) + qf \right)}{(-1+qf) qf} \right)^{2d} \right. \\
& \text{Hypergeometric2F1} \left[ 2-2d, 2-\frac{2d(1+s)}{s}, 3-\frac{2d(1+s)}{s}, \frac{-1+qf}{qf} \right] - \\
& \frac{1}{d-s+d s} e^{(-d+s)\tau} qf \left( \frac{(-1+q0) \left( 1 + \left( -1 + e^{-\frac{s\tau}{2}} \right) qf \right)}{(-1+qf) qf} \right)^{2d} \text{Hypergeometric2F1} \left[ 2-2d, \right. \\
& \left. 2-\frac{2d(1+s)}{s}, 3-\frac{2d(1+s)}{s}, \frac{e^{\frac{s\tau}{2}} (-1+qf)}{qf} \right] - 1 / (-3s+2d(1+s)) \quad 2 e^{-d\tau} \\
& (-1+qf) \left( e^{d\tau} \left( \frac{(-1+q0) \left( -e^{\frac{s\tau}{2}} (-1+qf) + qf \right)}{(-1+qf) qf} \right)^{2d} \text{Hypergeometric2F1} \left[ 2-2d, \right. \right. \\
& \left. \left. 3-\frac{2d(1+s)}{s}, 4-\frac{2d(1+s)}{s}, \frac{-1+qf}{qf} \right] - e^{\frac{3s\tau}{2}} \left( \frac{(-1+q0) \left( 1 + \left( -1 + e^{-\frac{s\tau}{2}} \right) qf \right)}{(-1+qf) qf} \right)^{2d} \right. \\
& \left. \left. \text{Hypergeometric2F1} \left[ 2-2d, 3-\frac{2d(1+s)}{s}, 4-\frac{2d(1+s)}{s}, \frac{e^{\frac{s\tau}{2}} (-1+qf)}{qf} \right] \right] \right) \\
\text{rbackapprox}[k_, \tau_] := & \frac{2 k r \left( -i \pi + \text{Log} \left[ e^{\frac{s\tau}{2}} (-1+qf) - qf \right] \right)}{s} \\
\text{mutbackapprox}[k_, \tau_] := & - \frac{2 \left( -1 + e^{\frac{s\tau}{2}} \right) k (-1+qf) u}{qf s} \\
\text{migbackapprox}[k_, \tau_] := & \frac{1}{2 \text{N0 NeN} (-1+q0)^2} k m (-1+qf)^2 \left( \frac{1}{qf} \right)^{3-2d} \left( \frac{(-1+q0) qf}{q0 (-1+qf)} \right)^{\frac{2d}{s}} \\
& \left( \frac{-e^{\frac{s\tau}{2}} (-1+qf) + qf}{qf} \right)^{-2d} \left( \frac{1}{d-s+d s} \left( \frac{-1+q0}{-1+qf} \right)^{2d} qf \left( \frac{-e^{\frac{s\tau}{2}} (-1+qf) + qf}{qf} \right)^{2d} \right. \\
& \text{Hypergeometric2F1} \left[ 2-2d, 2-\frac{2d(1+s)}{s}, 3-\frac{2d(1+s)}{s}, \frac{-1+qf}{qf} \right] - \\
& \frac{1}{d-s+d s} e^{(-d+s)\tau} \left( \frac{1}{qf} \right)^{-1+2d} \left( \frac{(-1+q0) \left( 1 + \left( -1 + e^{-\frac{s\tau}{2}} \right) qf \right)}{-1+qf} \right)^{2d}
\end{aligned}$$



$$\begin{aligned}
& \text{Hypergeometric2F1}\left[2-2d, 2-\frac{2d(1+s)}{s}, 3-\frac{2d(1+s)}{s}, \frac{e^{\frac{s\tau}{2}}(-1+qf)}{qf}\right] - \\
& 1/(-3s+2d(1+s)) \quad 2e^{-d\tau}(-1+qf) \left( e^{d\tau} \left( \frac{-1+q0}{-1+qf} \right)^{2d} \left( \frac{-e^{\frac{s\tau}{2}}(-1+qf)+qf}{qf} \right)^{2d} \right. \\
& \quad \text{Hypergeometric2F1}\left[2-2d, 3-\frac{2d(1+s)}{s}, 4-\frac{2d(1+s)}{s}, \frac{-1+qf}{qf}\right] - \\
& \quad \left. e^{\frac{3s\tau}{2}} \left( \frac{1}{qf} \right)^{2d} \left( \frac{(-1+q0) \left( 1 + \left( -1 + e^{-\frac{s\tau}{2}} \right) qf \right)}{-1+qf} \right)^{2d} \right. \\
& \quad \left. \text{Hypergeometric2F1}\left[2-2d, 3-\frac{2d(1+s)}{s}, 4-\frac{2d(1+s)}{s}, \frac{e^{\frac{s\tau}{2}}(-1+qf)}{qf}\right] \right) \Bigg) \\
& \text{cclassicbackapprox}[k_, \tau_] := \frac{(-1+k) k \left( -\frac{2 \left( -1 + e^{\frac{s\tau}{2}} \right) (-1+qf)}{s} + qf \tau \right)}{4 N0 \text{ NeN } qf} \\
& \text{migclassicbackapprox}[k_, \tau_] := \frac{k m \left( -\frac{2 \left( -1 + e^{\frac{s\tau}{2}} \right) (-1+qf)}{s} + qf \tau \right)}{2 N0 \text{ NeN } qf} \\
& \text{cbottlebackapprox}[k_, \tau_] := -\frac{1}{4 N0 \text{ NeN } (-1+q0)^2 (d-s+d s)} \\
& (-1+k) k (-1+qf)^2 \left( \frac{1}{qf} \right)^{2-2d} \left( \frac{(-1+q0) qf}{q0 (-1+qf)} \right)^{\frac{2d}{s}} \left( \frac{-e^{\frac{s\tau}{2}}(-1+qf)+qf}{qf} \right)^{-2d} \\
& \left( -\left( \frac{-1+q0}{-1+qf} \right)^{2d} \left( \frac{-e^{\frac{s\tau}{2}}(-1+qf)+qf}{qf} \right)^{2d} \text{Hypergeometric2F1}\left[2-2d, 2-\frac{2d(1+s)}{s}, \right. \right. \\
& \quad \left. \left. 3-\frac{2d(1+s)}{s}, \frac{-1+qf}{qf}\right] + e^{(-d+s)\tau} \left( \frac{1}{qf} \right)^{2d} \left( \frac{(-1+q0) \left( 1 + \left( -1 + e^{-\frac{s\tau}{2}} \right) qf \right)}{-1+qf} \right)^{2d} \right. \\
& \quad \left. \text{Hypergeometric2F1}\left[2-2d, 2-\frac{2d(1+s)}{s}, 3-\frac{2d(1+s)}{s}, \frac{e^{\frac{s\tau}{2}}(-1+qf)}{qf}\right] \right) \\
& \text{pcoalbottle}[k_, T_] := 1 - \text{Exp}[-\text{cbottlebackapprox}[k, T]] \\
& \text{nothingyetEmp}[c_, r_, m_, mig_, i_, \tau_] := \\
& \quad \text{Exp}[-\text{Accumulate}[c[[i]] + r[[i]] + m[[i]] + mig[[i]]][[\tau]]];
\end{aligned}$$

## Haldane's mapping function

```
In[116]:= haldane[r_] := 50 * Log[ $\frac{1}{1 - 2 \text{Abs}[r]}$ ] Sign[r]
```

## Plots

### Style

```
In[117]:= defaultcolors = ColorData[1, "ColorList"]; (*colors*)
labelstyle = Directive[FontSize → 12, FontFamily → "Helvetica"];
(*font for axis labels*)
letterstyle = Directive[FontSize → 14, FontFamily → "Helvetica"];
(*font for panel letters*)
letterposition = {0.05, 0.925}; (*position of panel letters*)
padding = {{60, 10}, {40, 10}}; (*padding around figures*)
Hpadding = {{70, 10}, {40, 10}}; (*special case of padding for diversity plots*)
ylabelposition = {-0.15, 0.5}; (*position of y-axis labels*)
Needs["ErrorBarPlots`"]; (*for plotting error bars*)
```

## Get backward-time simulation dynamics

```

In[125]:= getSimDynamics[Nval_, dval_, sval_,
    hval_, kval_, uval_, mval_, nreps_, maxt_, avg_] :=
(
    folder = StringForm["K`_d`_s`_k`_u`_m`", Nval, NumberForm[dval, {3, 2}],
        NumberForm[sval, {3, 2}], kval, If[0 < uval, NumberForm[uval, {6, 5}], 0],
        If[0 < mval < 1, NumberForm[mval, {2, 1}], mval]]; (*folder where data is*)
    Clear[data];
    Table[data[i] = Import[
        datadir <> ToString[folder] <> "/data/dynamics_" <> ToString[i - 1] <> ".txt",
        "Table", "FieldSeparators" -> " ", {i, nreps}]; (*data for each rep*)
    allp = Table[data[i][[2 ;; 3]], {i, nreps}]; (*allele frequencies, p*)
    alln = Table[data[i][[2 ;; 2]], {i, nreps}]; (*population sizes, n*)
    allpcut =
        Table[allp[[i]][[1 ;; Module[{n = 1}, While[allp[[i]][[n]] < 1, n++]; n - 1]]],
            {i, nreps}]; (*consider fixed when p=1*)
    allncut = Table[alln[[i]][[1 ;; Length[allpcut[[i]]]]], {i, nreps}];
    (*cut n when fixed*)
    allpreversed = Reverse[allpcut, 2]; (*reverse p*)
    allnreversed = Reverse[allncut, 2]; (*reverse n*)
    allppadded = PadRight[allpreversed];
    (*pad p with zeros for taking the mean/median*)
    allnpadded = PadRight[allnreversed - Nval] + Nval;
    (*pad n with N0 for taking the mean/median*)
    maxT = Min[maxt, Length[allppadded[[1]]]];
    (*consider times up to min of plotting range, maxt, or max fixation time*)
    listp = Table[Transpose[Join[{Table[-i, {i, maxT}],
        {allppadded[[j, 1 ;; maxT]]}], {j, nreps}]; (*add times to p for plotting*)
    listn = Table[Transpose[Join[{Table[-i, {i, maxT}],
        {allnpadded[[j, 1 ;; maxT]]}], {j, nreps}]; (*add times to n for plotting*)
    listpmean = avg[listp]; (*take mean/median*)
    listnmean = avg[listn]; (*take mean/median*)
    fixtimes = Table[Module[{n = 1},
        While[allpreversed[[i]][[n]] > 1 / (2 * Nval), n++] // Quiet; n - 1], {i, nreps}];
    (*fixation times: time between p=1/2N0 and p=1*)
    {listp, listn, listpmean, listnmean, fixtimes}
    (*return individual and mean/median dynamics and fixation times*)
)

```

```

In[126]:= getSimDynamicsClonal[Nval_, dval_, sval_,
    hval_, kval_, uval_, mval_, nreps_, maxt_, avg_] :=
(
    folder =
        StringForm["K`_d`s`_k`_u`_m`_clonal", Nval, NumberForm[dval, {3, 2}],
            NumberForm[sval, {3, 2}], kval, If[0 < uval, NumberForm[uval, {6, 5}], 0],
            If[0 < mval < 1, NumberForm[mval, {3, 2}], mval]]; (*folder where data is*)
    Clear[data];
    Table[data[i] = Import[
        datadir <> ToString[folder] <> "/data/dynamics_" <> ToString[i - 1] <> ".txt",
        "Table", "FieldSeparators" → " "], {i, nreps}]; (*data for each rep*)
    allp = Table[2 * data[i][[2 ;; 3]], {i, nreps}]; (*genotype frequencies, p*)
    alln = Table[data[i][[2 ;; 2]], {i, nreps}]; (*population sizes, n*)
    allpcut =
        Table[allp[[i]][[1 ;; Module[{n = 1}, While[allp[[i]][[n]] < 1, n++]; n - 1]]],
            {i, nreps}]; (*consider fixed when p=1*)
    allncut = Table[alln[[i]][[1 ;; Length[allpcut[[i]]]]], {i, nreps}];
    (*cut n when fixed*)
    allpreversed = Reverse[allpcut, 2]; (*reverse p*)
    allnreversed = Reverse[allncut, 2]; (*reverse n*)
    allppadded = PadRight[allpreversed];
    (*pad p with zeros for taking the mean/median*)
    allnpadded = PadRight[allnreversed - Nval] + Nval;
    (*pad n with N0 for taking the mean/median*)
    maxT = Min[maxt, Length[allppadded[[1]]]];
    (*consider times up to min of plotting range, maxt, or max fixation time*)
    listp = Table[Transpose[Join[{Table[-i, {i, maxT}],
        {allppadded[[j, 1 ;; maxT]]}], {j, nreps}]; (*add times to p for plotting*)
    listn = Table[Transpose[Join[{Table[-i, {i, maxT}],
        {allnpadded[[j, 1 ;; maxT]]}], {j, nreps}]; (*add times to n for plotting*)
    listpmean = avg[listp]; (*take mean/median*)
    listnmean = avg[listn]; (*take mean/median*)
    fixtimes =
        Table[Module[{n = 1}, While[allpreversed[[i]][[n]] > 1/Nval, n++] // Quiet; n - 1],
            {i, nreps}]; (*fixation times: time between p=1/N0 and p=1*)
    {listp, listn, listpmean, listnmean, fixtimes}
    (*return individual and mean/median dynamics and fixation times*)
)

```

## Predict backward-time theoretical dynamics

```

In[127]:= backwardsTheory[dval_, kval_, uval_, Nval_] :=
(
  q0 =
    (*effective initial frequency*)
    If[dval > 0,
      If[kval > 0, q0rescueSGV, If[uval > 0, q0rescueDNM, q0rescueMIG]],
      If[kval > 0, q0rescueSGV, If[uval > 0, q0sweepDNM, q0sweepMIG]]
    ] /.  $\epsilon \rightarrow w - 1$  /.  $\rho \rightarrow \text{pest}$  /.  $v \rightarrow w (3 + 4 B - 4 w) / 4$  /.  $w \rightarrow (1 + s h) (1 - dval)$ ;
  (*effective final frequency*)
  qf =  $1 - \frac{1}{2 N_0 \rho^2}$  /.  $\rho^2 \rightarrow \text{pest}$  /.  $v \rightarrow w \frac{(3 + 4 B - 4 W)}{4 \bar{W}^2}$  /.  $w \rightarrow \epsilon + 1$  /.  $\epsilon \rightarrow 1 - \frac{W}{\bar{W}}$  /.
     $\bar{W} \rightarrow (1 - dval) (1 + s)$  /.  $W \rightarrow (1 - dval) (1 + s h)$ ;
  (*backwards time allele frequency dynamics*)
  xsol = x[t_]  $\rightarrow$  qtadditiveback[t];
  (*backwards time population size dynamics*)
  If[dval > 0,
    nsol = n[t_]  $\rightarrow$  Min[Re[ntadditiveback[t]], N0],
    nsol = n[t_]  $\rightarrow$  N0
  ];
  (*fixation time*)
  tmax = Re[tfixadditive] /. N0  $\rightarrow$  Nval /. params;
  (*add times to p for plotting*)
  listp = Table[{-t, x[tmax - t] /. xsol /. N0  $\rightarrow$  Nval /. params}, {t, 0, tmax}];
  (*add times to n for plotting*)
  listn = Table[{-t, n[tmax - t] /. nsol /. N0  $\rightarrow$  Nval /. params}, {t, 0, tmax}];
  Clear[q0, qf];
  {listp, listn, tmax}
)

```

```

In[128]:= backwardsTheoryClonal[dval_, kval_, uval_, Nval_] :=
(
  q0 = 2 *
    (*effective initial frequency*)
    If[dval > 0,
      If[kval > 0, q0rescueSGV, If[uval > 0, q0rescueDNM, q0rescueMIG]],
      If[kval > 0, q0rescueSGV, If[uval > 0, q0sweepDNM, q0sweepMIG]]
    ] /.  $\epsilon \rightarrow w - 1$  /.  $\rho \rightarrow \text{pest}$  /.  $v \rightarrow (B - w) w$  /.  $w \rightarrow (1 + h s) (1 - dval)$ ;
  (*effective final frequency*)
  qf =  $1 - \frac{1}{N0 \rho^2}$  /.  $\rho^2 \rightarrow \text{pest}$  /.  $v \rightarrow \frac{(B - W) W}{Wbar^2}$  /.  $w \rightarrow \epsilon + 1$  /.  $\epsilon \rightarrow 1 - \frac{W}{Wbar}$  /.  $W \rightarrow (1 - dval)$  /.
    Wbar  $\rightarrow (1 + s h) (1 - dval)$ ;
  (*backwards time allele frequency dynamics*)
  xsol = x[t_]  $\rightarrow$  qtadditiveback[t];
  (*backwards time population size dynamics*)
  If[dval > 0,
    nsol = n[t_]  $\rightarrow$  Min[Re[ntadditivebackClonal[t]], N0],
    nsol = n[t_]  $\rightarrow$  N0
  ];
  (*fixation time*)
  tmax = Re[tfixadditive] /. N0  $\rightarrow$  Nval /. params;
  (*add times to p for plotting*)
  listp = Table[{-t, x[tmax - t] /. xsol /. N0  $\rightarrow$  Nval /. params}, {t, 0, tmax}];
  (*add times to n for plotting*)
  listn = Table[{-t, n[tmax - t] /. nsol /. N0  $\rightarrow$  Nval /. params}, {t, 0, tmax}];
  Clear[q0, qf];
  {listp, listn, tmax}
)

```

## Plot backwards-time dynamics, rescue vs constant population size

Plot dynamics backwards in time

```

In[129]:= plotDynamicsBackwards[Nval_, dval_, sval_, hval_,
  kval_, uval_, mval_, maxt_, letter_, save_, nreps_, avg_] :=
(
  params = {N0  $\rightarrow$  Nval, d  $\rightarrow$  dval, s  $\rightarrow$  sval, h  $\rightarrow$  hval,  $\kappa \rightarrow$  kval, u  $\rightarrow$  uval, m  $\rightarrow$  mval, B  $\rightarrow$  2};

  (*tick marks for x-axis*)
  ticks = {Table[{-x, x}, {x, 0, maxt, 50}], Automatic, None, None};

  (*rescue: simulations*)
  {pRescue, nRescue, pMeanRescue, nMeanRescue, fixTimesRescue} =
    getSimDynamics[Nval, dval, sval, hval, kval, uval, mval, nreps, maxt, avg];

  (*constant: simulations*)
  {pSweep, nSweep, pMeanSweep, nMeanSweep, fixTimesSweep} =
    getSimDynamics[Nval, 0, sval, hval, kval, uval, mval, nreps, maxt, avg];

  (*rescue: theory*)
  {pTheoryRescue, nTheoryRescue, fixTimeTheoryRescue} =

```

```

backwardsTheory[dval, kval, uval, Nval];

(*constant theory*)
{pTheorySweep, nTheorySweep, fixTimeTheorySweep} =
  backwardsTheory[0, kval, uval, Nval];

(*frame style and plotting fixation times*)
starheight = -0.025;
arrowheight = 0.05;
topstyle = Which[
  letter == "A",
  {
    FrameTicksStyle → {FontColor → White, Automatic, Automatic, Automatic},
    Epilog → {
      Text[Style[letter, letterstyle], Scaled@letterposition],
      Rotate[Text[
        Style["Allele frequency", labelstyle], Scaled@ylabelposition],  $\pi/2$ ],
      Text[Style[*, 16, defaultcolors[[1]]], {-fixTimeTheoryRescue, starheight}],
      Text[Style[*, 16, defaultcolors[[2]]], {-fixTimeTheorySweep, starheight}],
      Text[Style["↓", 16, defaultcolors[[1]]],
        {-Mean[fixTimesRescue], arrowheight}],
      Text[Style["↓", 16, defaultcolors[[2]]],
        {-Mean[fixTimesSweep], arrowheight}]
      ]
    },
  letter == "B",
  {
    FrameTicksStyle → {FontColor → White, FontColor → White, Automatic, Automatic},
    Epilog → {
      Text[Style[letter, letterstyle], Scaled@letterposition],
      Text[Style[*, 16, defaultcolors[[1]]],
        {-fixTimeTheoryRescue, starheight}],
      Text[Style[*, 16, defaultcolors[[2]]], {-fixTimeTheorySweep, starheight}],
      Text[Style["↓", 16, defaultcolors[[1]]],
        {-Mean[fixTimesRescue], arrowheight}],
      Text[Style["↓", 16, defaultcolors[[2]]],
        {-Mean[fixTimesSweep], arrowheight}]
      ]
    },
  letter == "C",
  {
    FrameTicksStyle → {FontColor → White, Automatic, Automatic, Automatic},
    Epilog → {
      Text[Style[letter, letterstyle], Scaled@letterposition],
      Rotate[Text[
        Style["Allele frequency", labelstyle], Scaled@ylabelposition],  $\pi/2$ ],
      Text[Style[*, 16, defaultcolors[[1]]], {-fixTimeTheoryRescue, starheight}],
      Text[Style[*, 16, defaultcolors[[2]]], {-fixTimeTheorySweep, starheight}],
      Text[Style["↓", 16, defaultcolors[[1]]],
        {-Mean[fixTimesRescue], arrowheight}],
      Text[Style["↓", 16, defaultcolors[[2]]],
        {-Mean[fixTimesSweep], arrowheight}]
      ]
    },
  letter == "D",

```

```

{
  FrameTicksStyle → {FontColor → White, FontColor → White, Automatic, Automatic},
  Epilog → {
    Text[Style[letter, letterstyle], Scaled@letterposition],
    Text[Style[*, 16, defaultcolors[[1]]],
      {-fixTimeTheoryRescue, starheight}],
    Text[Style[*, 16, defaultcolors[[2]]], {-fixTimeTheorySweep, starheight}],
    Text[Style["↓", 16, defaultcolors[[1]]],
      {-Mean[fixTimesRescue], arrowheight}],
    Text[Style["↓", 16, defaultcolors[[2]]],
      {-Mean[fixTimesSweep], arrowheight}]
  }
}
];
bottomstyle = Which[
  letter == "A",
  {
    FrameTicksStyle → {FontColor → White, Automatic, Automatic, Automatic},
    Epilog → {
      Rotate[
        Text[Style["Population size", labelstyle], Scaled@ylabelposition],  $\pi/2$ 
      ]
    },
  ],
  letter == "B",
  {
    FrameTicksStyle → {FontColor → White, FontColor → White, Automatic, Automatic}
  },
  letter == "C",
  {
    FrameTicksStyle → {Automatic, Automatic, Automatic, Automatic},
    Epilog → {
      Rotate[
        Text[Style["Population size", labelstyle], Scaled@ylabelposition],  $\pi/2$ 
      ],
      FrameLabel → {"Generations before fixation"}
    },
  ],
  letter == "D",
  {
    FrameTicksStyle → {Automatic, FontColor → White, Automatic, Automatic},
    FrameLabel → {"Generations before fixation"}
  }
];

(*plotting allele frequency dynamics*)
pdynamics = Show[
  ListPlot[
    pSweep,
    Joined → True,
    PlotStyle →
      Directive[defaultcolors[[2]], AbsoluteThickness[1], Opacity[0.1]],
    PlotRange → {{-maxt, 0}, {0, 1.01}}
  ],
  ListPlot[
    pRescue,
    Joined → True,

```



```

PlotStyle →
  Directive[defaultcolors[[1]], AbsoluteThickness[1], Opacity[0.1]],
PlotRange → {{-maxt, 0}, {0, 1.01}}
],
ListPlot[
  {pTheorySweep, pTheoryRescue},
  Joined → True,
  PlotStyle → {Directive[defaultcolors[[2]], AbsoluteThickness[3]],
    Directive[defaultcolors[[1]], AbsoluteThickness[3]]},
  PlotRange → {{-maxt, 0}, {0, 1.01}}
],
(*ListPlot[
  {pMeanSweep, pMeanRescue},
  Joined → True,
  PlotStyle → {Directive[defaultcolors[[2]], AbsoluteThickness[2], Opacity[0.8]],
    Directive[defaultcolors[[1]], AbsoluteThickness[2], Opacity[0.8]]},
  PlotRange → {{-maxt, 0}, {0, 1.01}}
], *)
Frame → {True, True, False, False},
FrameTicks → ticks,
PlotRange → {{-maxt, 0}, {0, 1.01}},
FrameStyle → labelstyle,
PlotRangePadding → None,
ImagePadding → padding,
topstyle,
PlotRangeClipping → False
];
If[save == 1,
  Export[imagedir <> ToString[StringForm["pdynamics_rescue_s``_k``_u``_m``.pdf",
    sval, kval, uval, mval]], pdynamics];
];

(*plotting population size dynamics*)
ndynamics = Show[
  ListPlot[
    nSweep,
    Joined → True,
    PlotStyle →
      Directive[defaultcolors[[2]], AbsoluteThickness[1], Opacity[0.1]],
    PlotRange → {{-maxt, 0}, {0, 1.01 * Nval}}
  ],
  ListPlot[
    nRescue,
    Joined → True,
    PlotStyle →
      Directive[defaultcolors[[1]], AbsoluteThickness[1], Opacity[0.1]],
    PlotRange → {{-maxt, 0}, {0, 1.01 * Nval}}
  ],
  ListPlot[
    {nTheorySweep, nTheoryRescue},
    Joined → True,
    PlotStyle → {Directive[defaultcolors[[2]], AbsoluteThickness[3]],
      Directive[defaultcolors[[1]], AbsoluteThickness[3]]},
    PlotRange → {{-maxt, 0}, {0, 1.01 * Nval}}
  ],

```

```

(*ListPlot[
  {nMeanSweep,nMeanRescue},
  Joined→True,
  PlotStyle→{Directive[defaultcolors[[2]],AbsoluteThickness[2],Opacity[0.8]],
    Directive[defaultcolors[[1]],AbsoluteThickness[2],Opacity[0.8]]},
  PlotRange→{{-maxt,0},{0,1.01*Nval}}
],*)
Frame → {True, True, False, False},
FrameTicks → ticks,
PlotRange → {{-maxt, 0}, {0, 1.01 * Nval}},
FrameStyle → labelstyle,
PlotRangePadding → None,
ImagePadding → padding,
bottomstyle,
PlotRangeClipping → False
];
If[save == 1,
  Export[imagedir <> ToString[StringForm["ndynamics_rescue_s``k``u``m``.pdf",
    sval, kval, uval, mval]], ndynamics];
];

GraphicsColumn[{pdynamics, ndynamics}, ImageSize → 400, Spacings → 0]
)

```

## Plot backwards-time dynamics, clonal

Plot dynamics backwards in time, with clonal model

```

In[130]:= plotDynamicsBackwardsClonal[Nval_, dval_, sval_, hval_,
  kval_, uval_, mval_, maxt_, letter_, save_, nreps_, avg_] :=
(
  params = {N0 → Nval, d → dval, s → sval, h → hval, κ → kval, u → uval, m → mval, B → 2};

  (*tick marks for x-axis*)
  ticks = {Table[{-x, x}, {x, 0, maxt, 50}], Automatic, None, None};

  (*simulations*)
  {pSweep, nSweep, pMeanSweep, nMeanSweep, fixTimesSweep} =
    getSimDynamicsClonal[Nval, dval, sval, hval, kval, uval, mval, nreps, maxt, avg];

  (*theory*)
  {pTheorySweep, nTheorySweep, fixTimeTheorySweep} =
    backwardsTheoryClonal[dval, kval, uval, Nval];

  (*frame style and plotting fixation times*)
  starheight = -0.025;
  arrowheight = 0.05;
  topstyle = Which[
    letter == "A",
    {
      FrameTicksStyle → {FontColor → White, Automatic, Automatic, Automatic},
      Epilog → {
        Text[Style[letter, letterstyle], Scaled@letterposition],
        Rotate[Text[
          Style["Allele frequency", labelstyle], Scaled@ylabelposition],  $\pi/2$ ],

```

```

Text[Style[*, 16, defaultcolors[[2]]], {-fixTimeTheorySweep, starheight}],
Text[Style["↓", 16, defaultcolors[[2]]],
  {-Mean[fixTimesSweep], arrowheight}]
}
},
letter == "B",
{
  FrameTicksStyle → {FontColor → White, FontColor → White, Automatic, Automatic},
  Epilog → {
    Text[Style[letter, letterstyle], Scaled@letterposition],
    Text[Style[*, 16, defaultcolors[[2]]], {-fixTimeTheorySweep, starheight}],
    Text[Style["↓", 16, defaultcolors[[2]]],
      {-Mean[fixTimesSweep], arrowheight}]
  }
},
letter == "C",
{
  FrameTicksStyle → {FontColor → White, Automatic, Automatic, Automatic},
  Epilog → {
    Text[Style[letter, letterstyle], Scaled@letterposition],
    Rotate[Text[
      Style["Allele frequency", labelstyle], Scaled@ylabelposition],  $\pi/2$ ],
    Text[Style[*, 16, defaultcolors[[2]]], {-fixTimeTheorySweep, starheight}],
    Text[Style["↓", 16, defaultcolors[[2]]],
      {-Mean[fixTimesSweep], arrowheight}]
  }
},
letter == "D",
{
  FrameTicksStyle → {FontColor → White, FontColor → White, Automatic, Automatic},
  Epilog → {
    Text[Style[letter, letterstyle], Scaled@letterposition],
    Text[Style[*, 16, defaultcolors[[2]]], {-fixTimeTheorySweep, starheight}],
    Text[Style["↓", 16, defaultcolors[[2]]],
      {-Mean[fixTimesSweep], arrowheight}]
  }
}
];
bottomstyle = Which[
  letter == "A",
  {
    FrameTicksStyle → {FontColor → White, Automatic, Automatic, Automatic},
    Epilog → {
      Rotate[
        Text[Style["Population size", labelstyle], Scaled@ylabelposition],  $\pi/2$ 
      ]
    }
  },
  letter == "B",
  {
    FrameTicksStyle → {FontColor → White, FontColor → White, Automatic, Automatic}
  },
  letter == "C",
  {
    FrameTicksStyle → {Automatic, Automatic, Automatic, Automatic},
    Epilog → {

```

```

        Rotate[
            Text[Style["Population size", labelstyle], Scaled@ylabelposition],  $\pi/2$ ]
    },
    FrameLabel → {"Generations before fixation"}
},
letter == "D",
{
    FrameTicksStyle → {Automatic, FontColor → White, Automatic, Automatic},
    FrameLabel → {"Generations before fixation"}
}
];

(*plotting allele frequency dynamics*)
pdynamics = Show[
    ListPlot[
        pSweep,
        Joined → True,
        PlotStyle →
            Directive[defaultcolors[[2]], AbsoluteThickness[1], Opacity[0.1]],
        PlotRange → {{-maxt, 0}, {0, 1.01}}
    ],
    ListPlot[
        pTheorySweep,
        Joined → True,
        PlotStyle → Directive[defaultcolors[[2]], AbsoluteThickness[3]],
        PlotRange → {{-maxt, 0}, {0, 1.01}}
    ],
    ListPlot[
        pMeanSweep,
        Joined → True,
        PlotStyle → Directive[Red, AbsoluteThickness[2], Opacity[0.8], Dashed],
        PlotRange → {{-maxt, 0}, {0, 1.01}}
    ],
    Frame → {True, True, False, False},
    FrameTicks → ticks,
    PlotRange → {{-maxt, 0}, {0, 1.01}},
    FrameStyle → labelstyle,
    PlotRangePadding → None,
    ImagePadding → padding,
    topstyle,
    PlotRangeClipping → False
];
If[save == 1,
    Export[
        imagedir <> ToString[StringForm["pdynamics_rescue_s``_k``_u``_m``_clonal.pdf",
            sval, kval, uval, mval]], pdynamics];
];

(*plotting population size dynamics*)
ndynamics = Show[
    ListPlot[
        nSweep,
        Joined → True,
        PlotStyle →
            Directive[defaultcolors[[2]], AbsoluteThickness[1], Opacity[0.1]],

```

```

    PlotRange → {{-maxt, 0}, {0, 1.01 * Nval}}
  ],
  ListPlot[
    nTheorySweep,
    Joined → True,
    PlotStyle → Directive[defaultcolors[[2]], AbsoluteThickness[3]],
    PlotRange → {{-maxt, 0}, {0, 1.01 * Nval}}
  ],
  ListPlot[
    nMeanSweep,
    Joined → True,
    PlotStyle → Directive[Red, AbsoluteThickness[2], Opacity[0.8], Dashed],
    PlotRange → {{-maxt, 0}, {0, 1.01 * Nval}}
  ],
  Frame → {True, True, False, False},
  FrameTicks → ticks,
  PlotRange → {{-maxt, 0}, {0, 1.01 * Nval}},
  FrameStyle → labelstyle,
  PlotRangePadding → None,
  ImagePadding → padding,
  bottomstyle,
  PlotRangeClipping → False
];
If[save == 1,
  Export[
    imagedir <> ToString[StringForm["ndynamics_rescue_s``k``u``m``clonal.pdf",
      sval, kval, uval, mval]], ndynamics];
];

GraphicsColumn[{pdynamics, ndynamics}, ImageSize → 400, Spacings → 0]
)

```

## Get coalescent estimates from simulations

```

In[131]:= getSimCoalescent[Nval_, dval_, sval_,
  kval_, uval_, mval_, nreps_, maxt_, lval_, rval_] :=
(
  (*get time reversed p and n dynamics, as in getSimDynamics[]*)
  folder = StringForm["K``d``s``k``u``m``", Nval, NumberForm[dval, {3, 2}],
    NumberForm[sval, {3, 2}], kval, If[0 < uval, NumberForm[uval, {6, 5}], 0],
    If[0 < mval < 1, NumberForm[mval, {2, 1}], mval]];
  Clear[data];
  Table[data[i] = Import[datadir <> ToString[folder] <> "/data/dynamics_" <>
    ToString[i - 1] <> ".txt", "Table", "FieldSeparators" → " ", {i, nreps}];
  allp = Table[data[i][[2 ;; 3]], {i, nreps}];
  alln = Table[data[i][[2 ;; 2]], {i, nreps}];
  allpcut = Table[
    allp[[i]][[1 ;; Module[{n = 1}, While[allp[[i]][[n]] < 1, n++]; n - 1]]],
    {i, nreps}];
  allncut = Table[alln[[i]][[1 ;; Length[allpcut[[i]]]]], {i, nreps}];
  allpreversed = Reverse[allpcut, 2];
  allnreversed = Reverse[allncut, 2];
  allpreversedcut = Table[
    allpreversed[[i]][[1 ;; Min[Module[{n = 1},

```

```

While[allpreversed[[i]][[n]] > 1 / (2 Nval), n++] // Quiet; n - 1], maxt]]],
{i, nreps}];
allnreversedcut =
Table[allnreversed[[i]][[1 ;; Length[allpreversedcut[[i]]]]], {i, nreps}];

(*calculate coalescent rates*)
coalEmpTabs =
Table[coalEmp[allnreversedcut[[i]], allpreversedcut[[i]], lval, 4 / 7],
{i, nreps}]; (*instantaneous coal rates*)
recEmpTabs = Table[recEmp[allpreversedcut[[i]], lval, rval], {i, nreps}];
(*instantaneous rec rates*)
mutEmpTabs = Table[mutEmp[allpreversedcut[[i]], lval, uval], {i, nreps}];
(*instantaneous mut rates*)
migEmpTabs = Table[migEmp[allnreversedcut[[i]], allpreversedcut[[i]], lval, mval],
{i, nreps}]; (*instantaneous mig rates*)

nothingyetTabs = Table[
Table[nothingyetEmp[coalEmpTabs, recEmpTabs, mutEmpTabs, migEmpTabs, i,  $\tau$ ],
{ $\tau$ , Length[coalEmpTabs[[i]]]}], {i, nreps}]; (*prob no events by gen  $\tau$ *)

coalEmpTabs = Table[Table[coalEmpTabs[[i,  $\tau$ ]] nothingyetTabs[[i,  $\tau$ ]],
{ $\tau$ , Length[coalEmpTabs[[i]]]}], {i, nreps}]; (*prob coal in gen  $\tau$ *)
recEmpTabs = Table[Table[recEmpTabs[[i,  $\tau$ ]] nothingyetTabs[[i,  $\tau$ ]],
{ $\tau$ , Length[recEmpTabs[[i]]]}], {i, nreps}]; (*prob rec in gen  $\tau$ *)
mutEmpTabs = Table[Table[mutEmpTabs[[i,  $\tau$ ]] nothingyetTabs[[i,  $\tau$ ]],
{ $\tau$ , Length[mutEmpTabs[[i]]]}], {i, nreps}]; (*prob mut in gen  $\tau$ *)
migEmpTabs = Table[Table[migEmpTabs[[i,  $\tau$ ]] nothingyetTabs[[i,  $\tau$ ]],
{ $\tau$ , Length[migEmpTabs[[i]]]}], {i, nreps}]; (*prob mig in gen  $\tau$ *)

coalEmpTabs = Table[{- $\tau$ , Mean[Select[PadRight[coalEmpTabs][[All,  $\tau$ ]], # > 0 &]]],
{ $\tau$ , Length[PadRight[coalEmpTabs][[1]]]}];
(*mean prob coal in gen  $\tau$ , with time added for plotting*)
recEmpTabs = Table[{- $\tau$ , Mean[Select[PadRight[recEmpTabs][[All,  $\tau$ ]], # > 0 &]]],
{ $\tau$ , Length[PadRight[recEmpTabs][[1]]]}];
(*mean prob rec in gen  $\tau$ , with time added for plotting*)
mutEmpTabs = Table[{- $\tau$ , Mean[Select[PadRight[mutEmpTabs][[All,  $\tau$ ]], # > 0 &]]],
{ $\tau$ , Length[PadRight[mutEmpTabs][[1]]]}];
(*mean prob mut in gen  $\tau$ , with time added for plotting*)
migEmpTabs = Table[{- $\tau$ , Mean[Select[PadRight[migEmpTabs][[All,  $\tau$ ]], # > 0 &]]],
{ $\tau$ , Length[PadRight[migEmpTabs][[1]]]}];
(*mean prob mig in gen  $\tau$ , with time added for plotting*)

{coalEmpTabs, recEmpTabs, mutEmpTabs, migEmpTabs}
(*return prob of events for plotting*)
)

```

## Plot the coalescent, rescue vs constant population size

```

In[132]:= plotCoalescentSims[Nval_, dval_, sval_, hval_, kval_, uval_,
mval_, lval_, rval_, maxt_, ymax_, letter_, save_, nreps_] :=
(
params = {N0 → Nval, d → dval, s → sval, h → hval,
 $\kappa$  → kval, u → uval, m → mval, l → lval, r → rval, NeN → 4 / 7, B → 2};

```

```

(*x-axis tick marks*)
ticks = {Table[{-x, x}, {x, 0, maxt, 50}], Automatic, None, None};

(*frame style*)
bottomstyle = Which[
  letter == "A",
  {
    FrameTicksStyle → {FontColor → White, Automatic, Automatic, Automatic},
    Epilog → {
      Text[Style[letter, letterstyle], Scaled@letterposition],
      Rotate[
        Text[Style["Probability", labelstyle], Scaled@ylabelposition],  $\pi/2$ ]
      ]
    },
  letter == "B",
  {
    FrameTicksStyle → {FontColor → White, FontColor → White, Automatic, Automatic},
    Epilog → {Text[Style[letter, letterstyle], Scaled@letterposition]}
  },
  letter == "C",
  {
    FrameTicksStyle → {Automatic, Automatic, Automatic, Automatic},
    Epilog → {
      Text[Style[letter, letterstyle], Scaled@letterposition],
      Rotate[
        Text[Style["Probability", labelstyle], Scaled@ylabelposition],  $\pi/2$ ]
      ],
    FrameLabel → {"Generations before fixation"}
  },
  letter == "D",
  {
    FrameTicksStyle → {Automatic, FontColor → White, Automatic, Automatic},
    Epilog → {Text[Style[letter, letterstyle], Scaled@letterposition]},
    FrameLabel → {"Generations before fixation"}
  }
];

(*rescue: theory*)
q0 = If[kval > 0, q0rescueSGV, If[uval > 0, q0rescueDNM, q0rescueMIG]] /.  $\epsilon \rightarrow w - 1$  /.
   $\rho \rightarrow \text{pest} /. v \rightarrow w (3 + 4 B - 4 w) / 4 /. w \rightarrow (1 + s h) (1 - d)$ ;
qf =  $1 - \frac{1}{2 N_0 \rho^2} /. \rho^2 \rightarrow \text{pest} /. v \rightarrow w \frac{(3 + 4 B - 4 W)}{4 Wbar^2} /. w \rightarrow \epsilon + 1 /. \epsilon \rightarrow 1 - \frac{W}{Wbar} /. Wbar \rightarrow (1 - d) (1 + s) /. W \rightarrow (1 - d) (1 + s h)$ ;
nsol = n[t_] → Min[Re[ntadditiveback[t]], N0];
xsol = x[t_] → qtadditiveback[t];
tmax = Re[tfixadditive /. params]; tmaxrescue = tmax;
nothingyet[l_,  $\tau$ ] :=
  Exp[-cbackapprox[l,  $\tau$ ] -
    rbackapprox[l,  $\tau$ ] - mutbackapprox[l,  $\tau$ ] - migbackapprox[l,  $\tau$ ]];
rescuecoal = Table[{- $\tau$ , Re[pcoal[l,  $\tau$ ] nothingyet[l,  $\tau$ ] /. ne[t_] → n[t] NeN /.
  xsol /. nsol /. params]}, { $\tau$ , 0, tmax}];
rescuerec = Table[{- $\tau$ , Re[preco[l,  $\tau$ ] nothingyet[l,  $\tau$ ] /. xsol /. nsol /. params]},
  { $\tau$ , 0, tmax}];

```

```

rescuemut = Table[{-τ, Re[pmut[1, τ] nothingyet[1, τ] /. xsol /. nsol /. params]},
  {τ, 0, tmax}];
rescuemig = Table[{-τ, Re[pmig[1, τ] nothingyet[1, τ] /. xsol /. nsol /. params]},
  {τ, 0, tmax}];

(*constant theory*)
q0 = If[kval > 0, q0rescueSGV, If[uval > 0, q0sweepDNM, q0sweepMIG]] /. ε → w - 1 /.
  ρ → pest /. v → w (3 + 4 B - 4 w) / 4 /. w → (1 + s h) (1 - d) /. d → 0;
qf = 1 -  $\frac{1}{2 N_0 \rho_2}$  /. ρ2 → pest /. v → w  $\frac{(3 + 4 B - 4 W)}{4 Wbar^2}$  /. w → ε + 1 /. ε → 1 -  $\frac{W}{Wbar}$  /.
  Wbar → (1 - d) (1 + s) /. W → (1 - d) (1 + s h) /. d → 0;
nsol = n[t_] → N0;
xsol = x[t_] → qtadditiveback[t];
tmax = tfixadditive /. params;
nothingyetc[l_, τ_] :=
  Exp[-cclassicbackapprox[1, τ] - rbackapprox[1, τ] -
    migclassicbackapprox[1, τ] - mutbackapprox[1, τ]];
sweepcoal = Table[{-τ, Re[pcoal[1, τ] nothingyetc[1, τ] /. ne[t_] → n[t] NeN /.
  xsol /. nsol /. params]}, {τ, 0, tmax}];
sweeprec = Table[{-τ, Re[preco[1, τ] nothingyetc[1, τ] /. xsol /. nsol /. params]},
  {τ, 0, tmax}];
sweepmut = Table[{-τ, Re[pmut[1, τ] nothingyetc[1, τ] /. xsol /. nsol /. params]},
  {τ, 0, tmax}];
sweepmig = Table[{-τ, Re[pmig[1, τ] nothingyetc[1, τ] /. xsol /. nsol /. params]},
  {τ, 0, tmax}];

(*rescue: simulations*)
{coalEmpRescue, recEmpRescue, mutEmpRescue, migEmpRescue} =
  getSimCoalescent[Nval, dval, sval, kval, uval, mval, nreps, maxt, lval, rval];

(*constant: simulations*)
{coalEmpSweep, recEmpSweep, mutEmpSweep, migEmpSweep} =
  getSimCoalescent[Nval, 0, sval, kval, uval, mval, nreps, maxt, lval, rval];

coalescent = Show[
  ListPlot[
    {rescuecoal, sweepcoal},
    Joined → True,
    PlotRange → All,
    PlotStyle → AbsoluteThickness[3],
    Axes → False
  ],
  ListPlot[
    {rescuerec, sweeprec},
    Joined → True,
    PlotRange → All,
    PlotStyle → Directive[AbsoluteThickness[3], Dashing[Large]]
  ],
  ListPlot[
    {rescuemut, sweepmut},
    Joined → True,
    PlotRange → All,
    PlotStyle → If[uval > 0, Directive[AbsoluteThickness[3], Dotted],
      Directive[AbsoluteThickness[0], White]]
  ]
];

```



```

],
ListPlot[
  {rescuemig, sweepmig},
  Joined → True,
  PlotRange → All,
  PlotStyle → If[mval > 0, Directive[AbsoluteThickness[3], Dotted],
    Directive[AbsoluteThickness[0], White]]
],
ListPlot[
  coalEmpSweep,
  Joined → True,
  PlotRange → All,
  PlotStyle →
    Directive[defaultcolors[[2]], AbsoluteThickness[2], Opacity[0.5]],
  Axes → False
],
ListPlot[
  recEmpSweep,
  Joined → True,
  PlotRange → All,
  PlotStyle →
    Directive[defaultcolors[[2]], AbsoluteThickness[2], Opacity[0.5], Dashed],
  Axes → False
],
ListPlot[
  mutEmpSweep,
  Joined → True,
  PlotRange → All,
  PlotStyle →
    Directive[defaultcolors[[2]], AbsoluteThickness[2], Opacity[0.5], Dotted],
  Axes → False
],
ListPlot[
  migEmpSweep,
  Joined → True,
  PlotRange → All,
  PlotStyle →
    Directive[defaultcolors[[2]], AbsoluteThickness[2], Opacity[0.5], Dotted],
  Axes → False
],
ListPlot[
  coalEmpRescue,
  Joined → True,
  PlotRange → All,
  PlotStyle →
    Directive[defaultcolors[[1]], AbsoluteThickness[2], Opacity[0.5]],
  Axes → False
],
ListPlot[
  recEmpRescue,
  Joined → True,
  PlotRange → All,
  PlotStyle →
    Directive[defaultcolors[[1]], AbsoluteThickness[2], Opacity[0.5], Dashed],
  Axes → False
]

```

```

],
ListPlot[
  mutEmpRescue,
  Joined → True,
  PlotRange → All,
  PlotStyle →
    Directive[defaultcolors[[1]], AbsoluteThickness[2], Opacity[0.5], Dotted],
  Axes → False
],
ListPlot[
  migEmpRescue,
  Joined → True,
  PlotRange → All,
  PlotStyle →
    Directive[defaultcolors[[1]], AbsoluteThickness[2], Opacity[0.5], Dotted],
  Axes → False
],
Frame → {True, True, False, False},
FrameTicks → ticks,
PlotRange → {{-maxt, 0}, {0, ymax}},
FrameStyle → labelstyle,
PlotRangePadding → None,
ImagePadding → padding,
bottomstyle,
PlotRangeClipping → False
];
If[save == 1,
  Export[imagedir <> ToString[StringForm["coalescent_rescue_s``k``u``m``.pdf",
    sval, kval, uval, mval]], coalescent];
];

Clear[q0, qf];
GraphicsColumn[{coalescent}, ImageSize → 400, Spacings → 0]
)

```

## Get simulation diversity

```

In[133]:= getSimDiversity[Nval_, dval_, sval_, hval_, kval_, uval_, mval_, nreps_] :=
(
  folder = StringForm["K``d``s``k``u``m``", Nval, NumberForm[dval, {3, 2}],
    NumberForm[sval, {3, 2}], kval, If[uval > 0, NumberForm[uval, {6, 5}], 0], mval];
  Clear[data];
  Table[data[i] = Import[datadir <> ToString[folder] <> "/data/stats_" <>
    ToString[i - 1] <> ".csv", "Table", "FieldSeparators" → ",", {i, nreps}];
  allH = Table[data[i][[All, {1, 2}]], {i, nreps}];
  (*pairwise diversity at a given physical distance for all reps*)
  allH[[All, All, 1]] = haldane[allH[[All, All, 1]]];
  (*convert physical distance to cM*)
  meanH = Mean/@Flatten[allH, {{2}, {1}}]; (*take mean across reps*)
  {allH, meanH} (*return each rep and mean*)
)

```

## Plot relative pairwise diversity, rescue vs constant population size

Plot pairwise nucleotide diversity divided by predicted background diversity

```
In[597]:= plotDiversityRelative[Nval_, dval_, sval_,
    hval_, kval_, uval_, mval_, nreps_, letter_, save_, nrs_] :=
(
  (*parameters*)
  params = {N0 → Nval, d → dval, s → sval, h → hval, κ → kval,
    u → uval, m → mval, l → 2, U → 6 * 10-9, rbp → 2 * 10-8, B → 2, NeN → 4 / 7};
  (*recombination distances to calculate theory at*)
  rmin = 0; rmax = 0.2; rint =  $\frac{2 \text{ rmax}}{\text{nrs}}$ ;
  (*plotting ranges*)
  xmin = -20; xmax = 20;
  ymax = 2;

  (*frame styling*)
  xticks = Table[{x, Abs[x]}, {x, xmin, xmax, 10}];
  yticks = Automatic;
  style = Which[
    letter == "A",
    {
      FrameLabel → {, "Relative mean pairwise diversity,  $\pi/\bar{\pi}$ "},
      FrameTicksStyle → {FontColor → White, Automatic, Automatic, Automatic},
      Epilog → {
        Text[Style[letter, letterstyle], Scaled@letterposition]
      }
    },
    letter == "B",
    {
      FrameTicksStyle → {FontColor → White, FontColor → White, Automatic, Automatic},
      Epilog → {
        Text[Style[letter, letterstyle], Scaled@letterposition]
      }
    },
    letter == "C",
    {
      FrameLabel → {"Distance from selected site (cM)",
        "Relative mean pairwise diversity,  $\pi/\bar{\pi}$ "},
      Epilog → {
        Text[Style[letter, letterstyle], Scaled@letterposition]
      }
    },
    letter == "D",
    {
      FrameTicksStyle → {Automatic, FontColor → White, Automatic, Automatic},
      FrameLabel → {"Distance from selected site (cM)",},
      Epilog → {
        Text[Style[letter, letterstyle], Scaled@letterposition]
      }
    }
  ]
)
```

```

];

(*rescue: simulations*)
{allHrescue, meanHrescue} =
  getSimDiversity[Nval, dval, sval, hval, kval, uval, mval, nreps];
backgroundH = Mean[Select[meanHrescue, Abs[#[[1]]] > 5 &][[All, 2]]];
(*mean diversity away from selected site*)
allHrescue[[All, All, 2]] = allHrescue[[All, All, 2]] / backgroundH;
(*relativize individual replicates*)
meanHrescue[[All, 2]] = meanHrescue[[All, 2]] / backgroundH; (*relativize mean*)

(*constant: simulations*)
{allHsweep, meanHsweep} =
  getSimDiversity[Nval, 0, sval, hval, kval, uval, mval, nreps];
backgroundH = Mean[Select[meanHsweep, Abs[#[[1]]] > 5 &][[All, 2]]];
(*mean diversity away from selected site*)
allHsweep[[All, All, 2]] = allHsweep[[All, All, 2]] / backgroundH;
(*relativize individual replicates*)
meanHsweep[[All, 2]] = meanHsweep[[All, 2]] / backgroundH; (*relativize mean*)

(*rescue: theory*)
q0 = If[kval > 0, q0rescueSGV, If[uval > 0, q0rescueDNM, q0rescueMIG]] /.  $\epsilon \rightarrow w - 1$  /.
   $\rho \rightarrow \text{pest}$  /.  $v \rightarrow w(3 + 4B - 4w) / 4$  /.  $w \rightarrow (1 + sh)(1 - d)$  /.
  params; (*effective initial frequency*)
qf =  $1 - \frac{1}{2N_0\rho^2}$  /.  $\rho \rightarrow \text{pest}$  /.  $v \rightarrow w \frac{(3 + 4B - 4W)}{4Wbar^2}$  /.  $w \rightarrow \epsilon + 1$  /.  $\epsilon \rightarrow 1 - \frac{W}{Wbar}$  /. Wbar  $\rightarrow$ 
   $(1 - d)(1 + s) / W \rightarrow (1 - d)(1 + sh) / \text{params}$ ; (*effective final frequency*)
xsol = x[t_]  $\rightarrow$  qtadditiveback[t]; (*allele frequency
  dynamics backwards in time*)
nsol = n[t_]  $\rightarrow$  Min[Re[ntadditiveback[t]], N0];
(*population size dynamics backwards in time*)
tmax = Re[tfixadditive /. params]; (*fixation time*)
tab = Table[
  {
    haldane[r], (*CM*)
    Sum[
      Which[
        kval > 0, preco[l,  $\tau$ ] Exp[-cbackapprox[l,  $\tau$ ] - rbackapprox[l,  $\tau$ ]],
        uval > 0, (preco[l,  $\tau$ ] + pmut[l,  $\tau$ ])
          Exp[-cbackapprox[l,  $\tau$ ] - rbackapprox[l,  $\tau$ ] - mutbackapprox[l,  $\tau$ ]],
        mval > 0, (preco[l,  $\tau$ ] + pmig[l,  $\tau$ ]) Exp[-cbackapprox[l,  $\tau$ ] -
          rbackapprox[l,  $\tau$ ] - migbackapprox[l,  $\tau$ ]]
      ] /. xsol /. nsol /. params,
    { $\tau$ , 0, tmax}], (*prob of getting off the selected background*)
  If[kval > 1, Exp[-cbackapprox[l,  $\tau$ ] - rbackapprox[l,  $\tau$ ]] (1 - 1/kval), 0] /.
     $\tau \rightarrow tmax$  /. params (*probability of no events*)
  },
  {r, rmin, rmax, rint}];
(*predict with theoretical  $\theta$ *)
tab1 = Table[{tab[[i, 1]],
   $\left( \text{tab}[[i, 2]] + \frac{\text{tab}[[i, 3]]}{1 - \text{pcoalbottle}[1, tmax]} \right) /. \text{params}$ }, {i, Length[tab]}];
(*fraction of diversity remaining across all recombination distances*)

```

```

tab2 = Table[{-tab1[[Length[tab1] - i, 1]], tab1[[Length[tab1] - i, 2]]},
  {i, 0, Length[tab1] - 1}]; (*just mirror theory in negative range,
and put it in left to right order so that Joined→True works for plotting*)
rescue = Join[tab2, tab1]; (*join negative and positive*)

(*constant: theory*)
q0 = If[kval > 0, q0rescueSGV, If[uval > 0, q0sweepDNM, q0sweepMIG]] /.  $\epsilon \rightarrow w - 1$  /.
   $\rho \rightarrow \text{pest}$  /.  $v \rightarrow w (3 + 4 B - 4 w) / 4$  /.  $w \rightarrow (1 + s h) (1 - d)$  /.  $d \rightarrow 0$  /. params;
qf =  $1 - \frac{1}{2 N_0 \rho_2}$  /.  $\rho_2 \rightarrow \text{pest}$  /.  $v \rightarrow w \frac{(3 + 4 B - 4 W)}{4 Wbar^2}$  /.  $w \rightarrow \epsilon + 1$  /.  $\epsilon \rightarrow 1 - \frac{W}{Wbar}$  /.
   $Wbar \rightarrow (1 - d) (1 + s)$  /.  $W \rightarrow (1 - d) (1 + s h)$  /.  $d \rightarrow 0$  /. params;
xsol = x[t_] → qtadditiveback[t];
nsol = n[t_] → N0;
tmax = Re[tfixadditive /. params];
tab1 = Table[{
  haldane[r],
  Poff = Sum[
    Which[
      kval > 0, preco[l,  $\tau$ ] Exp[-cclassicbackapprox[l,  $\tau$ ] - rbackapprox[l,  $\tau$ ]],
      uval > 0, (preco[l,  $\tau$ ] + pmut[l,  $\tau$ ]) Exp[
        -cclassicbackapprox[l,  $\tau$ ] - rbackapprox[l,  $\tau$ ] - mutbackapprox[l,  $\tau$ ]],
      mval > 0, (preco[l,  $\tau$ ] + pmig[l,  $\tau$ ]) Exp[-cclassicbackapprox[l,  $\tau$ ] -
        rbackapprox[l,  $\tau$ ] - migclassicbackapprox[l,  $\tau$ ]]
    ] /. xsol /. nsol /. params,
  { $\tau$ , 0, tmax}];
Pnothing = If[kval > 0, Exp[-cclassicbackapprox[l,  $\tau$ ] - rbackapprox[l,  $\tau$ ]]
  (1 - 1/kval), 0] /.  $\tau \rightarrow tmax$ ;
Re[Poff + Pnothing /. params]
},
{r, rmin, rmax, rint}];
tab2 = Table[{-tab1[[Length[tab1] - i, 1]], tab1[[Length[tab1] - i, 2]]},
  {i, 0, Length[tab1] - 1}];
sweep = Join[tab2, tab1];

(*plot*)
plot = Show[
  ListPlot[
    allHrescue,
    Joined → True,
    PlotStyle → Directive[defaultcolors[[1]], Thickness[0.001], Opacity[0.1]],
    Axes → False
  ],
  ListPlot[
    {sweep, rescue},
    Joined → True,
    PlotStyle → {
      Directive[AbsoluteThickness[3], defaultcolors[[2]], Dashing[0]],
      Directive[AbsoluteThickness[3], defaultcolors[[1]], Dashing[0]]
    },
    Axes → False,
    PlotRange → All
  ],
  ListPlot[
    {meanHrescue, meanHsweep},

```

```

    Joined → True,
    Axes → False,
    PlotStyle → Directive[Thickness[0.005], Opacity[0.8]]
  ],
  PlotRange → {{xmin, xmax}, {0, ymax}},
  Frame → {True, True, False, False},
  FrameStyle → labelstyle,
  FrameTicks → {xticks, yticks, None, None},
  style,
  PlotRangePadding → None,
  ImagePadding → Hpadding
];
If[save == 1,
  Export[imagedir <> ToString[StringForm[
    "EH_rescue_s``_k``_u``_m``_relative.pdf", sval, kval, uval, mval]], plot];
];

Clear[q0, qf, allH, allHs];
Show[plot, ImageSize → 400]
)

```

## Plot absolute pairwise diversity, rescue vs constant population size

Plot pairwise nucleotide diversity

```

In[613]:= plotDiversity[Nval_, dval_, sval_, hval_,
  kval_, uval_, mval_, nreps_, letter_, save_, nrs_] :=
(
  (*parameters*)
  params = {N0 → Nval, d → dval, s → sval, h → hval, κ → kval,
    u → uval, m → mval, l → 2, U → 6 * 10-9, rbp → 2 * 10-8, B → 2, NeN → 4 / 7};
  (*recombination distances to calculate theory at*)
  rmin = 0; rmax = 0.2; rint =  $\frac{2 \text{ rmax}}{\text{nrs}}$ ;

  (*plotting ranges*)
  xmin = -20; xmax = 20;
  ymax = 1.5 * 4 N0 NeN U /. params;

  (*frame styling*)
  xticks = Table[{x, Abs[x]}, {x, xmin, xmax, 10}];
  yticks = Automatic;
  style = Which[
    letter == "A",
    {
      FrameLabel → {, "Mean pairwise diversity, π"},
      FrameTicksStyle → {FontColor → White, Automatic, Automatic, Automatic},
      Epilog → {
        Text[Style[letter, letterstyle], Scaled@letterposition]
      }
    },
    letter == "B",

```

```

{
  FrameTicksStyle → {FontColor → White, FontColor → White, Automatic, Automatic},
  Epilog → {
    Text[Style[letter, letterstyle], Scaled@letterposition]
  }
},
letter == "C",
{
  FrameLabel →
    {"Distance from selected site (cM)", "Mean pairwise diversity,  $\pi$ "},
  Epilog → {
    Text[Style[letter, letterstyle], Scaled@letterposition]
  }
},
letter == "D",
{
  FrameTicksStyle → {Automatic, FontColor → White, Automatic, Automatic},
  FrameLabel → {"Distance from selected site (cM)",},
  Epilog → {
    Text[Style[letter, letterstyle], Scaled@letterposition]
  }
}
];

(*rescue: simulations*)
{allHrescue, meanHrescue} =
  getSimDiversity[Nval, dval, sval, hval, kval, uval, mval, nreps];

(*constant: simulations*)
{allHsweep, meanHsweep} =
  getSimDiversity[Nval, 0, sval, hval, kval, uval, mval, nreps];

(*rescue: theory*)
q0 = If[kval > 0, q0rescueSGV, If[uval > 0, q0rescueDNM, q0rescueMIG]] /.  $\epsilon \rightarrow w - 1$  /.
   $\rho \rightarrow \text{pest}$  /.  $v \rightarrow w (3 + 4 B - 4 w) / 4$  /.  $w \rightarrow (1 + s h) (1 - d)$  /.
  params; (*effective initial frequency*)

$$qf = 1 - \frac{1}{2 N_0 \rho^2} \text{ /. } \rho^2 \rightarrow \text{pest} \text{ /. } v \rightarrow W \frac{(3 + 4 B - 4 W)}{4 Wbar^2} \text{ /. } w \rightarrow \epsilon + 1 \text{ /. } \epsilon \rightarrow 1 - \frac{W}{Wbar} \text{ /. } Wbar \rightarrow$$

   $(1 - d) (1 + s) \text{ /. } W \rightarrow (1 - d) (1 + s h) \text{ /. params;}$  (*effective final frequency*)
xsol = x[t_] → qtadditiveback[t]; (*allele frequency
  dynamics backwards in time*)
nsol = n[t_] → Min[Re[ntadditiveback[t]], N0];
(*population size dynamics backwards in time*)
tmax = Re[tfixadditive /. params]; (*fixation time*)
tab = Table[
  {
    haldane[r], (*cM*)
    Sum[
      Which[
        kval > 0, preco[l,  $\tau$ ] Exp[-cbackapprox[l,  $\tau$ ] - rbackapprox[l,  $\tau$ ]],
        uval > 0, (preco[l,  $\tau$ ] + pmut[l,  $\tau$ ])
          Exp[-cbackapprox[l,  $\tau$ ] - rbackapprox[l,  $\tau$ ] - mutbackapprox[l,  $\tau$ ]],
        mval > 0, (preco[l,  $\tau$ ] + pmig[l,  $\tau$ ]) Exp[-cbackapprox[l,  $\tau$ ] -
          rbackapprox[l,  $\tau$ ] - migbackapprox[l,  $\tau$ ]]
      ]
    ]
  }
];

```

```

] /. xsol /. nsol /. params,
{τ, 0, tmax}], (*prob of getting off the selected background*)
If[kval > 1, Exp[-cbackapprox[1, τ] - rbackapprox[1, τ]] (1 - 1 / kval), 0] /.
τ → tmax /. params (*probability of no events*)
},
{r, rmin, rmax, rint}];
(*predict with theoretical θ:*)
tab1 = Table[{tab[[i, 1]],
  (tab[[i, 2]] * (1 - pcoalbottle[1, tmax]) + tab[[i, 3]]) 4 N0 NeN U /. params},
{i, Length[tab]}]; (*fraction of diversity
remaining across all recombination distances*)
tab2 = Table[{-tab1[[Length[tab1] - i, 1]], tab1[[Length[tab1] - i, 2]]},
{i, 0, Length[tab1] - 1}]; (*just mirror theory in negative range,
and put it in left to right order so that Joined→True works for plotting*)
rescueTheory = Join[tab2, tab1]; (*join negative and positive*)
(*predict with observed θ:*)
backgroundHrescue = Mean[Select[meanHrescue, Abs[#[[1]]] > 5 &][[All, 2]]];
(*estimate of background diversity
excluding sites within 5cM of selected site*)
tab1 = Table[{tab[[i, 1]], (tab[[i, 2]] * backgroundHrescue +
  tab[[i, 3]] * 4 N0 NeN U) /. params}, {i, Length[tab]}];
(*fraction of diversity remaining across all recombination distances*)
tab2 = Table[{-tab1[[Length[tab1] - i, 1]], tab1[[Length[tab1] - i, 2]]},
{i, 0, Length[tab1] - 1}]; (*just mirror theory in negative range,
and put it in left to right order so that Joined→True works for plotting*)
rescueEmp = Join[tab2, tab1];

(*constant: theory*)
q0 = If[kval > 0, q0rescueSGV, If[uval > 0, q0sweepDNM, q0sweepMIG]] /. ε → w - 1 /.
ρ → pest /. v → w (3 + 4 B - 4 w) / 4 /. w → (1 + s h) (1 - d) /. d → 0 /. params;
qf = 1 -  $\frac{1}{2 N_0 \rho_2}$  /. ρ2 → pest /. v → W  $\frac{(3 + 4 B - 4 W)}{4 Wbar^2}$  /. w → ε + 1 /. ε → 1 -  $\frac{W}{Wbar}$  /.
Wbar → (1 - d) (1 + s) /. W → (1 - d) (1 + s h) /. d → 0 /. params;
xsol = x[t_] → qtadditiveback[t];
nsol = n[t_] → N0;
tmax = Re[tfixadditive /. params];
tab1 = Table[{
  haldane[r],
  Poff = Sum[
    Which[
      kval > 0, preco[1, τ] Exp[-cclassicbackapprox[1, τ] - rbackapprox[1, τ]],
      uval > 0, (preco[1, τ] + pmut[1, τ]) Exp[
        -cclassicbackapprox[1, τ] - rbackapprox[1, τ] - mutbackapprox[1, τ]],
      mval > 0, (preco[1, τ] + pmig[1, τ]) Exp[-cclassicbackapprox[1, τ] -
        rbackapprox[1, τ] - migclassicbackapprox[1, τ]]
    ] /. xsol /. nsol /. params,
  {τ, 0, tmax}];
Pnothing = If[kval > 1, Exp[-cclassicbackapprox[1, τ] - rbackapprox[1, τ]]
(1 - 1 / kval), 0] /. τ → tmax;
Re[ $Poff * 4 N_0 U NeN * \text{Exp}\left[\frac{-tmax}{2 N_0 NeN}\right] + Pnothing * 4 N_0 U NeN$  /. params]
},

```



```

    {r, rmin, rmax, rint}];
tab2 = Table[{-tab1[Length[tab1] - i, 1]}, tab1[Length[tab1] - i, 2]],
    {i, 0, Length[tab1] - 1}];
sweep = Join[tab2, tab1];

(*plot*)
plot = Show[
  ListPlot[
    allHrescue,
    Joined → True,
    PlotStyle → Directive[defaultcolors[[1]], Thickness[0.001], Opacity[0.1]],
    Axes → False
  ],
  ListPlot[
    {sweep, rescueTheory},
    Joined → True,
    PlotStyle → {
      Directive[AbsoluteThickness[3], defaultcolors[[2]], Dashing[0]],
      Directive[AbsoluteThickness[3], defaultcolors[[1]], Dashing[0]]
    },
    Axes → False,
    PlotRange → All
  ],
  ListPlot[
    rescueEmp,
    Joined → True,
    PlotStyle →
      Directive[AbsoluteThickness[3], defaultcolors[[1]], Dashing[Large]],
    Axes → False,
    PlotRange → All
  ],
  ListPlot[
    {meanHrescue, meanHsweep},
    Joined → True,
    Axes → False,
    PlotStyle → Directive[Thickness[0.005], Opacity[0.8]]
  ],
  PlotRange → {{xmin, xmax}, {0, ymax}},
  Frame → {True, True, False, False},
  FrameStyle → labelstyle,
  FrameTicks → {xticks, yticks, None, None},
  style,
  PlotRangePadding → None,
  ImagePadding → Hpadding
];
If[save == 1,
  Export[imagedir <> ToString[
    StringForm["EH_rescue_s``_k``_u``_m``.pdf", sval, kval, uval, mval]], plot];
];

Clear[q0, qf, allH, allHs];
Show[plot, ImageSize → 400]

```

)

## Get simulation Tajima's D

```
In[136]:= getSimD[Nval_, dval_, sval_, kval_, uval_, mval_, nreps_] :=
(
  folder = StringForm["K`_d`_s`_k`_u`_m`_", Nval, NumberForm[dval, {3, 2}],
    NumberForm[sval, {3, 2}], kval, If[uval > 0, NumberForm[uval, {6, 5}], 0], mval];
  Clear[data];
  Table[data[i] = Import[datadir <> ToString[folder] <> "/data/stats_" <>
    ToString[i - 1] <> ".csv", "Table", "FieldSeparators" -> ",", {i, nreps}];
  allD = Table[data[i][[All, {1, 3}]], {i, nreps}];
  allD[[All, All, 1]] = haldane[allD[[All, All, 1]]];
  meanD = Mean /@ Flatten[allD, {{2}, {1}}];
  {allD, meanD}
)
```

## Plot Tajima's D, rescue vs constant population size

```
In[137]:= plotTajimasD[Nval_, dval_, sval_, hval_,
  kval_, uval_, mval_, nreps_, letter_, save_] :=
(
  params =
    {N0 -> Nval, d -> dval, s -> sval, h -> hval, k -> kval, u -> uval, m -> mval};
  xmin = -20; xmax = 20;
  ymin = -4; ymax = 4;
  paramposition = {0.85, 0.1};

  ticks = {Table[{x, Abs[x]}, {x, xmin, xmax, 10}], Automatic, None, None};
  style = Which[
    letter == "A",
    {
      FrameLabel -> {, "Tajima's D"},
      FrameTicksStyle -> {FontColor -> White, Automatic, Automatic, Automatic},
      Epilog -> {
        Text[Style[letter, letterstyle], Scaled@letterposition]
      }
    },
    letter == "B",
    {
      FrameTicksStyle ->
        {FontColor -> White, FontColor -> White, Automatic, Automatic},
      Epilog -> {
        Text[Style[letter, letterstyle], Scaled@letterposition]
      }
    },
    letter == "C",
    {
      FrameLabel -> {"Distance from selected site (cM)", "Tajima's D"},
      Epilog -> {
        Text[Style[letter, letterstyle], Scaled@letterposition]
      }
    },
  ],
```

```

letter == "D",
{
  FrameTicksStyle -> {Automatic, FontColor -> White, Automatic, Automatic},
  FrameLabel -> {"Distance from selected site (cM)",},
  Epilog -> {
    Text[Style[letter, letterstyle], Scaled@letterposition]
  }
}
];

(*rescue: simulations*)
{allDrescue, meanDrescue} = getSimD[Nval, dval, sval, kval, uval, mval, nreps];

(*constant: simulations*)
{allDsweep, meanDsweep} = getSimD[Nval, 0, sval, kval, uval, mval, nreps];

(*plot*)
plot = Show[
  ListPlot[allDrescue, Joined -> True, PlotStyle -> Directive[
    defaultcolors[[1]], Thickness[0.001], Opacity[0.1]], Axes -> False],
  ListPlot[{meanDrescue, meanDsweep}, Joined -> True, Axes -> False,
    PlotStyle -> Directive[Thickness[0.005], Opacity[0.8]]
  ],
  PlotRange -> {{xmin, xmax}, {ymin, ymax}},
  Frame -> {True, True, False, False},
  FrameStyle -> labelstyle,
  FrameTicks -> ticks,
  PlotRangePadding -> None,
  style,
  ImagePadding -> Hpadding
];
If[save == 1,
  Export[imagedir <> ToString[StringForm[
    "TajimasD_rescue_s``k``u``m``.pdf", sval, kval, uval, mval]], plot]
];

Show[plot, ImageSize -> 400]
)

```

## Plot backwards-time dynamics, rescue vs bottleneck

Plot dynamics backwards in time

```

In[138]:= plotDynamicsBackwardsBottleneck[Nval_, dval_, sval_, hval_, kval_,
  uval_, mval_, maxt_, letter_, save_, nreps_, avg_, Nvalbottle_] :=
(
  params = {N0 -> Nval, d -> dval, s -> sval, h -> hval, k -> kval, u -> uval, m -> mval, B -> 2};

  (*tick marks for x-axis*)
  ticks = {Table[{-x, x}, {x, 0, maxt, 50}], Automatic, None, None};

  (*rescue: simulations*)
  {pRescue, nRescue, pMeanRescue, nMeanRescue, fixTimesRescue} =
    getSimDynamics[Nval, dval, sval, hval, kval, uval, mval, nreps, maxt, avg];

```

```

(*constant: simulations*)
{pSweep, nSweep, pMeanSweep, nMeanSweep, fixTimesSweep} =
  getSimDynamics[Nvalbottle, 0, sval, hval, kval, uval, mval, nreps, maxt, avg];

(*rescue: theory*)
{pTheoryRescue, nTheoryRescue, fixTimeTheoryRescue} =
  backwardsTheory[dval, kval, uval, Nval];

(*constant theory*)
{pTheorySweep, nTheorySweep, fixTimeTheorySweep} =
  backwardsTheory[0, kval, uval, Nvalbottle];

(*frame style and plotting fixation times*)
starheight = -0.025;
arrowheight = 0.05;
topstyle = Which[
  letter == "A",
  {
    FrameTicksStyle → {FontColor → White, Automatic, Automatic, Automatic},
    Epilog → {
      Text[Style[letter, letterstyle], Scaled@letterposition],
      Rotate[Text[
        Style["Allele frequency", labelstyle], Scaled@ylabelposition],  $\pi/2$ ],
      Text[Style[*, 16, defaultcolors[[1]]], {-fixTimeTheoryRescue, starheight}],
      Text[Style[*, 16, defaultcolors[[3]]], {-fixTimeTheorySweep, starheight}],
      Text[Style["↓", 16, defaultcolors[[1]]],
        {-Mean[fixTimesRescue], arrowheight}],
      Text[Style["↓", 16, defaultcolors[[3]]],
        {-Mean[fixTimesSweep], arrowheight}]
    ]
  },
  letter == "B",
  {
    FrameTicksStyle → {FontColor → White, FontColor → White, Automatic, Automatic},
    Epilog → {
      Text[Style[letter, letterstyle], Scaled@letterposition],
      Text[Style[*, 16, defaultcolors[[1]]],
        {-fixTimeTheoryRescue, starheight}],
      Text[Style[*, 16, defaultcolors[[3]]], {-fixTimeTheorySweep, starheight}],
      Text[Style["↓", 16, defaultcolors[[1]]],
        {-Mean[fixTimesRescue], arrowheight}],
      Text[Style["↓", 16, defaultcolors[[3]]],
        {-Mean[fixTimesSweep], arrowheight}]
    ]
  },
  letter == "C",
  {
    FrameTicksStyle → {FontColor → White, Automatic, Automatic, Automatic},
    Epilog → {
      Text[Style[letter, letterstyle], Scaled@letterposition],
      Rotate[Text[
        Style["Allele frequency", labelstyle], Scaled@ylabelposition],  $\pi/2$ ],
      Text[Style[*, 16, defaultcolors[[1]]], {-fixTimeTheoryRescue, starheight}],
      Text[Style[*, 16, defaultcolors[[3]]], {-fixTimeTheorySweep, starheight}],
      Text[Style["↓", 16, defaultcolors[[1]]],

```

```

        {-Mean[fixTimesRescue], arrowheight}},
        Text[Style["↓", 16, defaultcolors[[3]]],
        {-Mean[fixTimesSweep], arrowheight}}
    }
},
letter == "D",
{
    FrameTicksStyle → {FontColor → White, FontColor → White, Automatic, Automatic},
    Epilog → {
        Text[Style[letter, letterstyle], Scaled@letterposition],
        Text[Style[*, 16, defaultcolors[[1]]],
        {-fixTimeTheoryRescue, starheight}},
        Text[Style[*, 16, defaultcolors[[3]]], {-fixTimeTheorySweep, starheight}],
        Text[Style["↓", 16, defaultcolors[[1]]],
        {-Mean[fixTimesRescue], arrowheight}},
        Text[Style["↓", 16, defaultcolors[[3]]],
        {-Mean[fixTimesSweep], arrowheight}}
    }
}
];
bottomstyle = Which[
    letter == "A",
    {
        FrameTicksStyle → {FontColor → White, Automatic, Automatic, Automatic},
        Epilog → {
            Rotate[
                Text[Style["Population size", labelstyle], Scaled@ylabelposition],  $\pi/2$ 
            ]
        },
    },
    letter == "B",
    {
        FrameTicksStyle → {FontColor → White, FontColor → White, Automatic, Automatic}
    },
    letter == "C",
    {
        FrameTicksStyle → {Automatic, Automatic, Automatic, Automatic},
        Epilog → {
            Rotate[
                Text[Style["Population size", labelstyle], Scaled@ylabelposition],  $\pi/2$ 
            ],
        },
        FrameLabel → {"Generations before fixation"}
    },
    letter == "D",
    {
        FrameTicksStyle → {Automatic, FontColor → White, Automatic, Automatic},
        FrameLabel → {"Generations before fixation"}
    }
];

(*plotting allele frequency dynamics*)
pdynamics = Show[
    ListPlot[
        pSweep,
        Joined → True,
        PlotStyle →

```

```

    Directive[defaultcolors[[3]], AbsoluteThickness[1], Opacity[0.1]],
    PlotRange → {{-maxt, 0}, {0, 1.01}}
],
ListPlot[
  pRescue,
  Joined → True,
  PlotStyle →
    Directive[defaultcolors[[1]], AbsoluteThickness[1], Opacity[0.1]],
    PlotRange → {{-maxt, 0}, {0, 1.01}}
],
ListPlot[
  {pTheorySweep, pTheoryRescue},
  Joined → True,
  PlotStyle → {Directive[defaultcolors[[3]], AbsoluteThickness[3]],
    Directive[defaultcolors[[1]], AbsoluteThickness[3]]},
  PlotRange → {{-maxt, 0}, {0, 1.01}}
],
(*ListPlot[
  {pMeanSweep, pMeanRescue},
  Joined → True,
  PlotStyle → {Directive[defaultcolors[[2]], AbsoluteThickness[2], Opacity[0.8]],
    Directive[defaultcolors[[1]], AbsoluteThickness[2], Opacity[0.8]]},
  PlotRange → {{-maxt, 0}, {0, 1.01}}
], *)
Frame → {True, True, False, False},
FrameTicks → ticks,
PlotRange → {{-maxt, 0}, {0, 1.01}},
FrameStyle → labelstyle,
PlotRangePadding → None,
ImagePadding → padding,
topstyle,
PlotRangeClipping → False
];
If[save == 1,
  Export[
    imagedir <> ToString[StringForm["pdynamics_rescue_s``_k``_u``_m``_bottle.pdf",
      sval, kval, uval, mval]], pdynamics];
];

(*plotting population size dynamics*)
ndynamics = Show[
  ListPlot[
    nSweep,
    Joined → True,
    PlotStyle →
      Directive[defaultcolors[[3]], AbsoluteThickness[1], Opacity[0.1]],
      PlotRange → {{-maxt, 0}, {0, 1.01 * Nval}}
  ],
  ListPlot[
    nRescue,
    Joined → True,
    PlotStyle →
      Directive[defaultcolors[[1]], AbsoluteThickness[1], Opacity[0.1]],
      PlotRange → {{-maxt, 0}, {0, 1.01 * Nval}}
  ],

```

```

ListPlot[
  {nTheorySweep, nTheoryRescue},
  Joined → True,
  PlotStyle → {Directive[defaultcolors[[3]], AbsoluteThickness[3]],
    Directive[defaultcolors[[1]], AbsoluteThickness[3]]},
  PlotRange → {{-maxt, 0}, {0, 1.01 * Nval}}
],
(*ListPlot[
  {nMeanSweep, nMeanRescue},
  Joined → True,
  PlotStyle → {Directive[defaultcolors[[2]], AbsoluteThickness[2], Opacity[0.8]],
    Directive[defaultcolors[[1]], AbsoluteThickness[2], Opacity[0.8]]},
  PlotRange → {{-maxt, 0}, {0, 1.01 * Nval}}
], *)
Frame → {True, True, False, False},
FrameTicks → ticks,
PlotRange → {{-maxt, 0}, {0, 1.01 * Nval}},
FrameStyle → labelstyle,
PlotRangePadding → None,
ImagePadding → padding,
bottomstyle,
PlotRangeClipping → False
];
If[save == 1,
  Export[
    imagedir <> ToString[StringForm["ndynamics_rescue_s``k``u``m``bottle.pdf",
      sval, kval, uval, mval]], ndynamics];
];

GraphicsColumn[{podynamics, ndynamics}, ImageSize → 400, Spacings → 0]
)

```

## Plot the coalescent, rescue vs bottleneck

```

In[139]:= plotCoalescentSimsBottleneck[Nval_, dval_, sval_, hval_, kval_, uval_,
  mval_, lval_, rval_, maxt_, ymax_, letter_, save_, nreps_, Nvalbottle_] :=
(
  params = {N0 → Nval, d → dval, s → sval, h → hval,
    x → kval, u → uval, m → mval, l → lval, r → rval, NeN → 4 / 7, B → 2};

  (*x-axis tick marks*)
  ticks = {Table[{-x, x}, {x, 0, maxt, 50}], Automatic, None, None};

  (*frame style*)
  bottomstyle = Which[
    letter == "A",
    {
      FrameTicksStyle → {FontColor → White, Automatic, Automatic, Automatic},
      Epilog → {
        Text[Style[letter, letterstyle], Scaled@letterposition],
        Rotate[
          Text[Style["Probability", labelstyle], Scaled@ylabelposition],  $\pi / 2$ ]
        }
    }
  ]
)

```

```

},
letter == "B",
{
  FrameTicksStyle → {FontColor → White, FontColor → White, Automatic, Automatic},
  Epilog → {Text[Style[letter, letterstyle], Scaled@letterposition]}
},
letter == "C",
{
  FrameTicksStyle → {Automatic, Automatic, Automatic, Automatic},
  Epilog → {
    Text[Style[letter, letterstyle], Scaled@letterposition],
    Rotate[
      Text[Style["Probability", labelstyle], Scaled@ylabelposition],  $\pi / 2$ 
    ]
  },
  FrameLabel → {"Generations before fixation"}
},
letter == "D",
{
  FrameTicksStyle → {Automatic, FontColor → White, Automatic, Automatic},
  Epilog → {Text[Style[letter, letterstyle], Scaled@letterposition]},
  FrameLabel → {"Generations before fixation"}
}
];

(*rescue: theory*)
q0 = If[kval > 0, q0rescueSGV, If[uval > 0, q0rescueDNM, q0rescueMIG]] /.  $\epsilon \rightarrow w - 1$  /.
 $\rho \rightarrow \text{pest} /. v \rightarrow w (3 + 4 B - 4 w) / 4 /. w \rightarrow (1 + s h) (1 - d)$ ;

$$qf = 1 - \frac{1}{2 N_0 \rho_2} /. \rho_2 \rightarrow \text{pest} /. v \rightarrow w \frac{(3 + 4 B - 4 W)}{4 Wbar^2} /. w \rightarrow \epsilon + 1 /. \epsilon \rightarrow 1 - \frac{W}{Wbar} /.$$

Wbar →  $(1 - d) (1 + s) /. W \rightarrow (1 - d) (1 + s h)$ ;
nsol = n[t_] → Min[Re[ntadditiveback[t]], N0];
xsol = x[t_] → qtadditiveback[t];
tmax = Re[tfixadditive /. params]; tmaxrescue = tmax;
nothingyet[l_,  $\tau$ ] :=
Exp[-cbackapprox[l,  $\tau$ ] -
  rbackapprox[l,  $\tau$ ] - mutbackapprox[l,  $\tau$ ] - migbackapprox[l,  $\tau$ ]];
rescuecoal = Table[{- $\tau$ , Re[pcoal[l,  $\tau$ ] nothingyet[l,  $\tau$ ] /. ne[t_] → n[t] NeN /.
  xsol /. nsol /. params]}, { $\tau$ , 0, tmax}];
rescuerec = Table[{- $\tau$ , Re[preco[l,  $\tau$ ] nothingyet[l,  $\tau$ ] /. xsol /. nsol /. params]},
  { $\tau$ , 0, tmax}];
rescuemut = Table[{- $\tau$ , Re[pmut[l,  $\tau$ ] nothingyet[l,  $\tau$ ] /. xsol /. nsol /. params]},
  { $\tau$ , 0, tmax}];
rescuemig = Table[{- $\tau$ , Re[pmig[l,  $\tau$ ] nothingyet[l,  $\tau$ ] /. xsol /. nsol /. params]},
  { $\tau$ , 0, tmax}];

(*constant theory*)
q0 = If[kval > 0, q0rescueSGV, If[uval > 0, q0sweepDNM, q0sweepMIG]] /.  $\epsilon \rightarrow w - 1$  /.
 $\rho \rightarrow \text{pest} /. v \rightarrow w (3 + 4 B - 4 w) / 4 /. w \rightarrow (1 + s h) (1 - d) /. d \rightarrow 0$ ;

$$qf = 1 - \frac{1}{2 N_0 \rho_2} /. \rho_2 \rightarrow \text{pest} /. v \rightarrow w \frac{(3 + 4 B - 4 W)}{4 Wbar^2} /. w \rightarrow \epsilon + 1 /. \epsilon \rightarrow 1 - \frac{W}{Wbar} /.$$

Wbar →  $(1 - d) (1 + s) /. W \rightarrow (1 - d) (1 + s h) /. d \rightarrow 0$ ;
nsol = n[t_] → N0;
xsol = x[t_] → qtadditiveback[t];

```



```

tmax = tfixadditive /. N0 → Nvalbottle /. params;
nothingyetc[l_, τ_] :=
  Exp[-cclassicbackapprox[l, τ] - rbackapprox[l, τ] -
    migclassicbackapprox[l, τ] - mutbackapprox[l, τ]];
sweepcoal = Table[{-τ, Re[pcoal[l, τ] nothingyetc[l, τ] /. ne[t_] → n[t] NeN /.
  xsol /. nsol /. N0 → Nvalbottle /. params]}, {τ, 0, tmax}];
sweeprec = Table[{-τ, Re[preco[l, τ] nothingyetc[l, τ] /. xsol /. nsol /.
  N0 → Nvalbottle /. params]}, {τ, 0, tmax}];
sweepmut = Table[{-τ, Re[pmut[l, τ] nothingyetc[l, τ] /. xsol /. nsol /.
  N0 → Nvalbottle /. params]}, {τ, 0, tmax}];
sweepmig = Table[{-τ, Re[pmig[l, τ] nothingyetc[l, τ] /. xsol /. nsol /.
  N0 → Nvalbottle /. params]}, {τ, 0, tmax}];

(*rescue: simulations*)
{coalEmpRescue, recEmpRescue, mutEmpRescue, migEmpRescue} =
  getSimCoalescent[Nval, dval, sval, kval, uval, mval, nreps, maxt, lval, rval];

(*constant: simulations*)
{coalEmpSweep, recEmpSweep, mutEmpSweep, migEmpSweep} =
  getSimCoalescent[Nvalbottle, 0, sval, kval, uval, mval, nreps, maxt, lval, rval];

coalescent = Show[
  ListPlot[
    {rescuecoal, , sweepcoal},
    Joined → True,
    PlotRange → All,
    PlotStyle → AbsoluteThickness[3],
    Axes → False
  ],
  ListPlot[
    {rescuerec, , sweeprec},
    Joined → True,
    PlotRange → All,
    PlotStyle → Directive[AbsoluteThickness[3], Dashing[Large]]
  ],
  ListPlot[
    {rescuemut, , sweepmut},
    Joined → True,
    PlotRange → All,
    PlotStyle → If[uval > 0, Directive[AbsoluteThickness[3], Dotted],
      Directive[AbsoluteThickness[0], White]]
  ],
  ListPlot[
    {rescuemig, , sweepmig},
    Joined → True,
    PlotRange → All,
    PlotStyle → If[mval > 0, Directive[AbsoluteThickness[3], Dotted],
      Directive[AbsoluteThickness[0], White]]
  ],
  ListPlot[
    coalEmpSweep,
    Joined → True,
    PlotRange → All,
    PlotStyle →
      Directive[defaultcolors[[3]], AbsoluteThickness[2], Opacity[0.5]],

```

```

    Axes → False
  ],
  ListPlot[
    recEmpSweep,
    Joined → True,
    PlotRange → All,
    PlotStyle →
      Directive[defaultcolors[[3]], AbsoluteThickness[2], Opacity[0.5], Dashed],
    Axes → False
  ],
  ListPlot[
    mutEmpSweep,
    Joined → True,
    PlotRange → All,
    PlotStyle →
      Directive[defaultcolors[[3]], AbsoluteThickness[2], Opacity[0.5], Dotted],
    Axes → False
  ],
  ListPlot[
    coalEmpRescue,
    Joined → True,
    PlotRange → All,
    PlotStyle →
      Directive[defaultcolors[[1]], AbsoluteThickness[2], Opacity[0.5]],
    Axes → False
  ],
  ListPlot[
    recEmpRescue,
    Joined → True,
    PlotRange → All,
    PlotStyle →
      Directive[defaultcolors[[1]], AbsoluteThickness[2], Opacity[0.5], Dashed],
    Axes → False
  ],
  ListPlot[
    mutEmpRescue,
    Joined → True,
    PlotRange → All,
    PlotStyle →
      Directive[defaultcolors[[1]], AbsoluteThickness[2], Opacity[0.5], Dotted],
    Axes → False
  ],
  Frame → {True, True, False, False},
  FrameTicks → ticks,
  PlotRange → {{-maxt, 0}, {0, ymax}},
  FrameStyle → labelstyle,
  PlotRangePadding → None,
  ImagePadding → padding,
  bottomstyle,
  PlotRangeClipping → False
];
If[save == 1,
  Export[imagedir <>
    ToString[StringForm["coalescent_rescue_s``_k``_u``_m``_bottle.pdf",
      sval, kval, uval, mval]], coalescent];

```

```
];

Clear[q0, qf];
GraphicsColumn[{coalescent}, ImageSize → 400, Spacings → 0]
)
```

## Plot relative pairwise diversity, rescue vs bottleneck

Plot pairwise nucleotide diversity divided by predicted background diversity

```
In[603]:= plotDiversityRelativeBottleneck[Nval_, dval_, sval_, hval_,
      kval_, uval_, mval_, nreps_, letter_, save_, nrs_, Nvalbottle_] :=
(
  (*parameters*)
  params = {N0 → Nval, d → dval, s → sval, h → hval, κ → kval,
    u → uval, m → mval, l → 2, U → 6 * 10-9, rbp → 2 * 10-8, B → 2, NeN → 4 / 7};
  (*recombination distances to calculate theory at*)
  rmin = 0; rmax = 0.2; rint =  $\frac{2 \text{ rmax}}{\text{nrs}}$ ;
  (*plotting ranges*)
  xmin = -20; xmax = 20;
  ymax = 1.5 /. params;

  (*frame styling*)
  xticks = Table[{x, Abs[x]}, {x, xmin, xmax, 10}];
  yticks = Automatic;
  style = Which[
    letter == "A",
    {
      FrameLabel → {"Relative mean pairwise diversity,  $\pi/\bar{\pi}$ "},
      FrameTicksStyle → {FontColor → White, Automatic, Automatic, Automatic},
      Epilog → {
        Text[Style[letter, letterstyle], Scaled@letterposition]
      }
    },
    letter == "B",
    {
      FrameTicksStyle → {FontColor → White, FontColor → White, Automatic, Automatic},
      Epilog → {
        Text[Style[letter, letterstyle], Scaled@letterposition]
      }
    },
    letter == "C",
    {
      FrameLabel → {"Distance from selected site (cM)",
        "Relative mean pairwise diversity,  $\pi/\bar{\pi}$ "},
      Epilog → {
        Text[Style[letter, letterstyle], Scaled@letterposition]
      }
    },
    letter == "D",
```

```

{
  FrameTicksStyle → {Automatic, FontColor → White, Automatic, Automatic},
  FrameLabel → {"Distance from selected site (cM)",},
  Epilog → {
    Text[Style[letter, letterstyle], Scaled@letterposition]
  }
}
];

(*rescue: simulations*)
{allHrescue, meanHrescue} =
  getSimDiversity[Nval, dval, sval, hval, kval, uval, mval, nreps];
backgroundH = Mean[Select[meanHrescue, Abs[#[[1]]] > 5 &][[All, 2]]];
(*mean diversity away from selected site*)
allHrescue[[All, All, 2]] = allHrescue[[All, All, 2]] / backgroundH;
(*relativize individual replicates*)
meanHrescue[[All, 2]] = meanHrescue[[All, 2]] / backgroundH; (*relativize mean*)

(*constant: simulations*)
{allHsweep, meanHsweep} =
  getSimDiversity[Nvalbottle, 0, sval, hval, kval, uval, mval, nreps];
backgroundH = Mean[Select[meanHsweep, Abs[#[[1]]] > 5 &][[All, 2]]];
(*mean diversity away from selected site*)
allHsweep[[All, All, 2]] = allHsweep[[All, All, 2]] / backgroundH;
(*relativize individual replicates*)
meanHsweep[[All, 2]] = meanHsweep[[All, 2]] / backgroundH; (*relativize mean*)

(*rescue: theory*)
q0 = If[kval > 0, q0rescueSGV, If[uval > 0, q0rescueDNM, q0rescueMIG]] /.  $\epsilon \rightarrow w - 1$  /.
   $\rho \rightarrow \text{pest}$  /.  $v \rightarrow w (3 + 4 B - 4 w) / 4$  /.  $w \rightarrow (1 + s h) (1 - d)$  /.
  params; (*effective initial frequency*)

$$qf = 1 - \frac{1}{2 N_0 \rho^2} \text{ /. } \rho^2 \rightarrow \text{pest} \text{ /. } v \rightarrow w \frac{(3 + 4 B - 4 W)}{4 Wbar^2} \text{ /. } w \rightarrow \epsilon + 1 \text{ /. } \epsilon \rightarrow 1 - \frac{W}{Wbar} \text{ /. } Wbar \rightarrow$$

   $(1 - d) (1 + s) \text{ /. } W \rightarrow (1 - d) (1 + s h) \text{ /. } params$ ; (*effective final frequency*)
xsol = x[t_] → qtadditiveback[t]; (*allele frequency
  dynamics backwards in time*)
nsol = n[t_] → Min[Re[ntadditiveback[t]], N0];
(*population size dynamics backwards in time*)
tmax = Re[tfixadditive /. params]; (*fixation time*)
tab = Table[
  {
    haldane[r], (*cM*)
    Sum[
      Which[
        kval > 0, preco[l,  $\tau$ ] Exp[-cbackapprox[l,  $\tau$ ] - rbackapprox[l,  $\tau$ ]],
        uval > 0, (preco[l,  $\tau$ ] + pmut[l,  $\tau$ ])
          Exp[-cbackapprox[l,  $\tau$ ] - rbackapprox[l,  $\tau$ ] - mutbackapprox[l,  $\tau$ ]],
        mval > 0, (preco[l,  $\tau$ ] + pmig[l,  $\tau$ ]) Exp[-cbackapprox[l,  $\tau$ ] -
          rbackapprox[l,  $\tau$ ] - migbackapprox[l,  $\tau$ ]]
      ] /. xsol /. nsol /. params,
    { $\tau$ , 0, tmax}}, (*prob of getting off the selected background*)
    If[kval > 1, Exp[-cbackapprox[l,  $\tau$ ] - rbackapprox[l,  $\tau$ ]] (1 - 1 / kval), 0] /.
     $\tau \rightarrow tmax$  /. params (*probability of no events*)
  },

```

```

{r, rmin, rmax, rint}];
(*predict with theoretical  $\theta$ *)
tab1 = Table[{tab[[i, 1]],

$$\left( \text{tab}[[i, 2]] + \frac{\text{tab}[[i, 3]]}{1 - \text{pcoalbottle}[1, \text{tmax}]} \right) /. \text{params}}, \{i, \text{Length}[\text{tab}]\}];
(*fraction of diversity remaining across all recombination distances*)
tab2 = Table[{-tab1[Length[tab1] - i, 1]], tab1[Length[tab1] - i, 2]],
  {i, 0, Length[tab1] - 1}]; (*just mirror theory in negative range,
and put it in left to right order so that Joined→True works for plotting*)
rescue = Join[tab2, tab1]; (*join negative and positive*)

(*constant: theory*)
q0 =
  If[kval > 0, q0rescueSGV, If[uval > 0, q0sweepDNM, q0sweepMIG]] /.  $\epsilon \rightarrow w - 1 /. \rho \rightarrow$ 
    pest /.  $v \rightarrow w (3 + 4 B - 4 w) / 4 /.$ 
     $w \rightarrow (1 + s h) (1 - d) /. d \rightarrow 0 /. N0 \rightarrow \text{Nvalbottle} /. \text{params};$ 
qf =  $1 - \frac{1}{2 N0 \rho^2} /. \rho^2 \rightarrow \text{pest} /. v \rightarrow w \frac{(3 + 4 B - 4 W)}{4 Wbar^2} /. w \rightarrow \epsilon + 1 /. \epsilon \rightarrow 1 - \frac{W}{Wbar} /. Wbar \rightarrow$ 
     $(1 - d) (1 + s) /. W \rightarrow (1 - d) (1 + s h) /. d \rightarrow 0 /. N0 \rightarrow \text{Nvalbottle} /. \text{params};$ 
xsol = x[t_] → qtadditiveback[t];
nsol = n[t_] → N0;
tmax = Re[tfixadditive /. params];
tab1 = Table[{
  haldane[r],
  Poff = Sum[
    Which[
      kval > 0,
      preco[l,  $\tau$ ] Exp[-cclassicbackapprox[l,  $\tau$ ] - rbackapprox[l,  $\tau$ ]],
      uval > 0, (preco[l,  $\tau$ ] + pmut[l,  $\tau$ ]) Exp[-cclassicbackapprox[l,  $\tau$ ] -
        rbackapprox[l,  $\tau$ ] - mutbackapprox[l,  $\tau$ ]],
      mval > 0, (preco[l,  $\tau$ ] + pmig[l,  $\tau$ ]) Exp[-cclassicbackapprox[l,  $\tau$ ] -
        rbackapprox[l,  $\tau$ ] - migclassicbackapprox[l,  $\tau$ ]]
    ] /. xsol /. nsol /. N0 → Nvalbottle /. params,
  { $\tau$ , 0, tmax}];
Pnothing = If[kval > 0, Exp[-cclassicbackapprox[l,  $\tau$ ] - rbackapprox[l,  $\tau$ ]]
  (1 - 1 / kval), 0] /.  $\tau \rightarrow \text{tmax} /. N0 \rightarrow \text{Nvalbottle} /. \text{params};$ 
Re[Poff + Pnothing]
},
{r, rmin, rmax, rint}];
tab2 = Table[{-tab1[Length[tab1] - i, 1]], tab1[Length[tab1] - i, 2]],
  {i, 0, Length[tab1] - 1}];
sweep = Join[tab2, tab1];

(*plot*)
plot = Show[
  ListPlot[
    allHrescue,
    Joined → True,
    PlotStyle → Directive[defaultcolors[[1]], Thickness[0.001], Opacity[0.1]],
    Axes → False
  ],
  ListPlot[
    {sweep, rescue},$$

```

```

Joined → True,
PlotStyle → {
  Directive[AbsoluteThickness[3], defaultcolors[[3]], Dashing[0]],
  Directive[AbsoluteThickness[3], defaultcolors[[1]], Dashing[0]]
},
Axes → False,
PlotRange → All
],
ListPlot[
  {meanHrescue, , meanHsweep},
  Joined → True,
  Axes → False,
  PlotStyle → Directive[Thickness[0.005], Opacity[0.8]],
  PlotRange → {0, All}
],
PlotRange → {{xmin, xmax}, {0, ymax}},
Frame → {True, True, False, False},
FrameStyle → labelstyle,
FrameTicks → {xticks, yticks, None, None},
style,
PlotRangePadding → None,
ImagePadding → Hpadding
];
If[save == 1,
  Export[imagedir <>
    ToString[StringForm["EH_rescue_s``k``u``m``_relative_bottle.pdf",
      sval, kval, uval, mval]], plot];
];

Clear[q0, qf, allH, allHs];
Show[plot, ImageSize → 400]
)

```

## Plot absolute pairwise diversity, rescue vs bottleneck

Plot pairwise nucleotide diversity

```

In[604]:= plotDiversityBottleneck[Nval_, dval_, sval_, hval_,
  kval_, uval_, mval_, nreps_, letter_, save_, nrs_, Nvalbottle_] :=
(
  (*parameters*)
  params = {N0 → Nval, d → dval, s → sval, h → hval, κ → kval,
    u → uval, m → mval, l → 2, U → 6 * 10-9, rbp → 2 * 10-8, B → 2, NeN → 4 / 7};
  (*recombination distances to calculate theory at*)
  rmin = 0; rmax = 0.2; rint =  $\frac{2 \text{ rmax}}{\text{nrs}}$ ;
  (*plotting ranges*)
  xmin = -20; xmax = 20;
  ymax = 1.5 * 4 N0 NeN U /. params;
  (*frame styling*)

```

```

xticks = Table[{x, Abs[x]}, {x, xmin, xmax, 10}];
yticks = Automatic;
style = Which[
  letter == "A",
  {
    FrameLabel → {"Mean pairwise diversity,  $\pi$ "},
    FrameTicksStyle → {FontColor → White, Automatic, Automatic, Automatic},
    Epilog → {
      Text[Style[letter, letterstyle], Scaled@letterposition]
    }
  },
  letter == "B",
  {
    FrameTicksStyle → {FontColor → White, FontColor → White, Automatic, Automatic},
    Epilog → {
      Text[Style[letter, letterstyle], Scaled@letterposition]
    }
  },
  letter == "C",
  {
    FrameLabel →
      {"Distance from selected site (cM)", "Mean pairwise diversity,  $\pi$ "},
    Epilog → {
      Text[Style[letter, letterstyle], Scaled@letterposition]
    }
  },
  letter == "D",
  {
    FrameTicksStyle → {Automatic, FontColor → White, Automatic, Automatic},
    FrameLabel → {"Distance from selected site (cM)",},
    Epilog → {
      Text[Style[letter, letterstyle], Scaled@letterposition]
    }
  }
];

(*rescue: simulations*)
{allHrescue, meanHrescue} =
  getSimDiversity[Nval, dval, sval, hval, kval, uval, mval, nreps];

(*constant: simulations*)
{allHsweep, meanHsweep} =
  getSimDiversity[Nvalbottle, 0, sval, hval, kval, uval, mval, nreps];

(*rescue: theory*)
q0 = If[kval > 0, q0rescueSGV, If[uval > 0, q0rescueDNM, q0rescueMIG]] /.  $\epsilon \rightarrow w - 1$  /.
   $\rho \rightarrow \text{pest}$  /.  $v \rightarrow w (3 + 4 B - 4 w) / 4$  /.  $w \rightarrow (1 + s h) (1 - d)$  /.
  params; (*effective initial frequency*)
qf =  $1 - \frac{1}{2 N_0 \rho^2}$  /.  $\rho^2 \rightarrow \text{pest}$  /.  $v \rightarrow w \frac{(3 + 4 B - 4 w)}{4 \bar{w}^2}$  /.  $w \rightarrow \epsilon + 1$  /.  $\epsilon \rightarrow 1 - \frac{w}{\bar{w}}$  /.  $\bar{w} \rightarrow$ 
   $(1 - d) (1 + s) / w \rightarrow (1 - d) (1 + s h) / w$  /. params; (*effective final frequency*)
xsol = x[t_] → qtadditiveback[t]; (*allele frequency
  dynamics backwards in time*)
nsol = n[t_] → Min[Re[ntadditiveback[t]], N0];

```

```

(*population size dynamics backwards in time*)
tmax = Re[tfixadditive /. params]; (*fixation time*)
tab = Table[
{
  haldane[r], (*cM*)
  Sum[
    Which[
      kval > 0, preco[l,  $\tau$ ] Exp[-cbackapprox[l,  $\tau$ ] - rbackapprox[l,  $\tau$ ],
      uval > 0, (preco[l,  $\tau$ ] + pmut[l,  $\tau$ ]
        Exp[-cbackapprox[l,  $\tau$ ] - rbackapprox[l,  $\tau$ ] - mutbackapprox[l,  $\tau$ ],
      mval > 0, (preco[l,  $\tau$ ] + pmig[l,  $\tau$ ]) Exp[-cbackapprox[l,  $\tau$ ] -
        rbackapprox[l,  $\tau$ ] - migbackapprox[l,  $\tau$ ]
    ] /. xsol /. nsol /. params,
    { $\tau$ , 0, tmax}], (*prob of getting off the selected background*)
  If[kval > 1, Exp[-cbackapprox[l,  $\tau$ ] - rbackapprox[l,  $\tau$ ] (1 - 1/kval), 0] /.
     $\tau \rightarrow$  tmax /. params (*probability of no events*)
},
{r, rmin, rmax, rint}];
(*predict with theoretical  $\theta$ *)
tab1 = Table[
  {tab[[i, 1]], (tab[[i, 2]] * (1 - pcoalbottle[l, tmax]) + tab[[i, 3]] (1 - 1/kval))
    4 N0 NeN U /. params}, {i, Length[tab]}];
(*fraction of diversity remaining across all recombination distances*)
tab2 = Table[{-tab1[[Length[tab1] - i, 1]], tab1[[Length[tab1] - i, 2]]},
  {i, 0, Length[tab1] - 1}]; (*just mirror theory in negative range,
and put it in left to right order so that Joined→True works for plotting*)
rescueTheory = Join[tab2, tab1]; (*join negative and positive*)
(*predict with observed  $\theta$ *)
backgroundHrescue = Mean[Select[meanHrescue, Abs[#[[1]]] > 5 &][[All, 2]]];
(*estimate of background diversity
excluding sites within 5cM of selected site*)
tab1 = Table[{tab[[i, 1]], (tab[[i, 2]] * backgroundHrescue +
  tab[[i, 3]] * 4 N0 NeN U) /. params}, {i, Length[tab]}];
(*fraction of diversity remaining across all recombination distances*)
tab2 = Table[{-tab1[[Length[tab1] - i, 1]], tab1[[Length[tab1] - i, 2]]},
  {i, 0, Length[tab1] - 1}]; (*just mirror theory in negative range,
and put it in left to right order so that Joined→True works for plotting*)
rescueEmp = Join[tab2, tab1];

(*constant: theory*)
q0 =
  If[kval > 0, q0rescueSGV, If[uval > 0, q0sweepDNM, q0sweepMIG]] /.  $\epsilon \rightarrow w - 1$  /.  $\rho \rightarrow$ 
    pest /.  $v \rightarrow w (3 + 4 B - 4 w) / 4$  /.
     $w \rightarrow (1 + s h) (1 - d) /. d \rightarrow 0 /. N0 \rightarrow Nvalbottle /. params$ ;

$$qf = 1 - \frac{1}{2 N0 \rho^2} /. \rho^2 \rightarrow pest /. v \rightarrow w \frac{(3 + 4 B - 4 W)}{4 Wbar^2} /. w \rightarrow \epsilon + 1 /. \epsilon \rightarrow 1 - \frac{W}{Wbar} /. Wbar \rightarrow$$

    (1 - d) (1 + s) /.  $W \rightarrow (1 - d) (1 + s h) /. d \rightarrow 0 /. N0 \rightarrow Nvalbottle /. params$ ;
xsol = x[t_] → qtadditiveback[t];
nsol = n[t_] → N0;
tmax = Re[tfixadditive /. params];
tab1 = Table[{
  haldane[r],
  Poff = Sum[

```



```

Which[
  kval > 0,
  preco[1,  $\tau$ ] Exp[-cclassicbackapprox[1,  $\tau$ ] - rbackapprox[1,  $\tau$ ]],
  uval > 0, (preco[1,  $\tau$ ] + pmut[1,  $\tau$ ]) Exp[-cclassicbackapprox[1,  $\tau$ ] -
    rbackapprox[1,  $\tau$ ] - mutbackapprox[1,  $\tau$ ]],
  mval > 0, (preco[1,  $\tau$ ] + pmig[1,  $\tau$ ]) Exp[-cclassicbackapprox[1,  $\tau$ ] -
    rbackapprox[1,  $\tau$ ] - migclassicbackapprox[1,  $\tau$ ]]
] /. xsol /. nsol /. N0 → Nvalbottle /. params,
{ $\tau$ , 0, tmax}];
Pnothing = If[kval > 0, Exp[-cclassicbackapprox[1,  $\tau$ ] - rbackapprox[1,  $\tau$ ]]
  (1 - 1/kval), 0] /.  $\tau$  → tmax /. N0 → Nvalbottle /. params;
Re[Poff Exp[- $\frac{tmax}{2 Nvalbottle NeN}$ ] + Pnothing] * 4 Nval NeN U /. params
},
{r, rmin, rmax, rint}];
tab2 = Table[{-tab1[Length[tab1] - i, 1]}, tab1[Length[tab1] - i, 2]],
  {i, 0, Length[tab1] - 1}];
sweep = Join[tab2, tab1];

(*plot*)
plot = Show[
  ListPlot[
    allHrescue,
    Joined → True,
    PlotStyle → Directive[defaultcolors[[1]], Thickness[0.001], Opacity[0.1]],
    Axes → False
  ],
  ListPlot[
    {sweep, rescueTheory},
    Joined → True,
    PlotStyle → {
      Directive[AbsoluteThickness[3], defaultcolors[[3]], Dashing[0]],
      Directive[AbsoluteThickness[3], defaultcolors[[1]], Dashing[0]]
    },
    Axes → False,
    PlotRange → All
  ],
  ListPlot[
    rescueEmp,
    Joined → True,
    PlotStyle →
      Directive[AbsoluteThickness[3], defaultcolors[[1]], Dashing[Large]],
    Axes → False,
    PlotRange → All
  ],
  ListPlot[
    {meanHrescue, , meanHsweep},
    Joined → True,
    Axes → False,
    PlotStyle → Directive[Thickness[0.005], Opacity[0.8]],
    PlotRange → All
  ],
  PlotRange → {{xmin, xmax}, {0, ymax}},

```

```

    Frame → {True, True, False, False},
    FrameStyle → labelstyle,
    FrameTicks → {xticks, yticks, None, None},
    style,
    PlotRangePadding → None,
    ImagePadding → Hpadding
  ];
  If[save == 1,
    Export[imagedir <> ToString[StringForm[
      "EH_rescue_s``_k``_u``_m``_bottle.pdf", sval, kval, uval, mval]], plot];
  ];

  Clear[q0, qf, allH, allHs];
  Show[plot, ImageSize → 400]
)

```

## Plot Tajima's D, rescue vs bottleneck

```

In[142]:= plotTajimasDBottleneck[Nval_, dval_, sval_, hval_,
  kval_, uval_, mval_, nreps_, letter_, save_, Nvalbottle_] :=
(
  params =
    {N0 → Nval, d → dval, s → sval, h → hval, k → kval, u → uval, m → mval};
  xmin = -20; xmax = 20;
  ymin = -4; ymax = 4;
  paramposition = {0.85, 0.1};

  ticks = {Table[{x, Abs[x]}, {x, xmin, xmax, 10}], Automatic, None, None};
  style = Which[
    letter == "A",
    {
      FrameLabel -> {, "Tajima's D"},
      FrameTicksStyle -> {FontColor -> White, Automatic, Automatic, Automatic},
      Epilog -> {
        Text[Style[letter, letterstyle], Scaled@letterposition]
      }
    },
    letter == "B",
    {
      FrameTicksStyle ->
        {FontColor -> White, FontColor -> White, Automatic, Automatic},
      Epilog -> {
        Text[Style[letter, letterstyle], Scaled@letterposition]
      }
    },
    letter == "C",
    {
      FrameLabel -> {"Distance from selected site (cM)", "Tajima's D"},
      Epilog -> {
        Text[Style[letter, letterstyle], Scaled@letterposition]
      }
    },
    letter == "D",

```

```

{
  FrameTicksStyle -> {Automatic, FontColor -> White, Automatic, Automatic},
  FrameLabel -> {"Distance from selected site (cM)",},
  Epilog -> {
    Text[Style[letter, letterstyle], Scaled@letterposition]
  }
}
];

(*rescue: simulations*)
{allDrescue, meanDrescue} = getSimD[Nval, dval, sval, kval, uval, mval, nreps];

(*constant: simulations*)
{allDsweep, meanDsweep} = getSimD[Nvalbottle, 0, sval, kval, uval, mval, nreps];

(*plot*)
plot = Show[
  ListPlot[allDrescue, Joined -> True, PlotStyle -> Directive[
    defaultcolors[[1]], Thickness[0.001], Opacity[0.1]], Axes -> False],
  ListPlot[{meanDrescue, meanDsweep}, Joined -> True, Axes -> False,
    PlotStyle -> Directive[Thickness[0.005], Opacity[0.8]], PlotRange -> All
  ],
  PlotRange -> {{xmin, xmax}, {ymin, ymax}},
  Frame -> {True, True, False, False},
  FrameStyle -> labelstyle,
  FrameTicks -> ticks,
  PlotRangePadding -> None,
  style,
  ImagePadding -> Hpadding
];
If[save == 1,
  Export[
    imagedir <> ToString[StringForm["TajimasD_rescue_s``k``u``m``bottle.pdf",
      sval, kval, uval, mval]], plot]
];

Show[plot, ImageSize -> 400]
)

```

## Plot SFS

```

In[143]:= getSimSFS[Nval_, dval_, sval_, kval_, uval_, mval_, nreps_] :=
(
  folder = StringForm["K``d``s``k``u``m``, Nval, NumberForm[dval, {3, 2}],
    NumberForm[sval, {3, 2}], kval, If[0 < uval, NumberForm[uval, {6, 5}], 0],
    If[0 < mval < 1, NumberForm[mval, {3, 2}], mval]]; (*folder where data is*)
  Clear[data];
  Table[data[i] = Import[datadir <> ToString[folder] <>
    "/data/sfs_" <> ToString[i - 1] <> ".txt", "Table"], {i, nreps}];
  sfs = Table[data[i], {i, nreps}]; (*sfs at a given recombination distance*)
  cMs = Import[datadir <> ToString[folder] <> "/data/stats_0.csv", "CSV"][[All, 1]];
  (*recombination distances*)
  {sfs, cMs}
)

```

```

In[144]:= plotSFS[Nval_, dval_, sval_, kval_, uval_, mval_, nreps_, Nvalbottle_, Uval_] :=
(
  (*rescue simulations*)
  {sfs1, rec1} = getSimSFS[Nval, dval, sval, kval, uval, mval, nreps];
  msfs1 = Mean[sfs1]; (*mean*)

  (*constant simulations*)
  {sfs2, rec2} = getSimSFS[Nval, 0, sval, kval, uval, mval, nreps];
  msfs2 = Mean[sfs2]; (*mean*)

  (*bottlenecked simulations*)
  {sfs3, rec3} = getSimSFS[Nvalbottle, 0, sval, kval, uval, mval, nreps];
  msfs3 = Mean[sfs3]; (*mean*)

  (*recombination distances to plot*)
  cMpicks = {10, 1, 0.1}; (*cMs you want to plot at*)
  rs = Table[
    list = -(haldane[rec1] - i)^2;
    Ordering[list, -1], (*entry with closest cM you can get*)
    {i, cMpicks}
  ] // Flatten;

  Table[
    Show[
      Plot[2 Nval Uval  $\frac{1}{x}$ , {x, 0, 100},
        PlotRange → {0, 10-4}, PlotStyle → Directive[Black]],
      ListPlot[msfs1[[r]], PlotRange → {0, 10-4}, Joined → True,
        PlotStyle → Directive[Thick]],
      ListPlot[msfs2[[r]], PlotRange → {0, 10-4}, Joined → True,
        PlotStyle → Directive[Thick, defaultcolors[[2]]]],
      ListPlot[msfs3[[r]], PlotRange → {0, 10-4}, Joined → True,
        PlotStyle → Directive[defaultcolors[[3]], Thick]],
      Frame → {True, True, False, False},
      FrameLabel → {
        "Copies of derived allele at site",
        If[r == rs[[1]], "Fraction of sites in window", ""]
      },
      FrameTicksStyle → {Automatic,
        If[r == rs[[1]], Automatic, FontColor → White], Automatic, Automatic},
      FrameStyle → labelstyle,
      Epilog → {
        Text[Style[ToString[NumberForm[haldane[rec1][[r]]], {3, 1}]] <> " cM",
          labelstyle], Scaled@{0.5, 0.9}]
      }
    ],
    {r, rs}
  ]

```

```

]
)

In[145]:= plotSFSlog[Nval_, dval_, sval_, kval_,
  uval_, mval_, nreps_, Nvalbottle_, Uval_, size_] :=
(
  (*rescue simulations*)
  {sfs1, rec1} = getSimSFS[Nval, dval, sval, kval, uval, mval, nreps];
  msfs1 = Mean[sfs1]; (*mean*)

  (*constant simulations*)
  {sfs2, rec2} = getSimSFS[Nval, 0, sval, kval, uval, mval, nreps];
  msfs2 = Mean[sfs2]; (*mean*)

  (*bottlenecked simulations*)
  {sfs3, rec3} = getSimSFS[Nvalbottle, 0, sval, kval, uval, mval, nreps];
  msfs3 = Mean[sfs3]; (*mean*)

  (*recombination distances to plot*)
  cMpicks = {10, 1, 0.1}; (*cMs you want to plot at*)
  rs = Table[
    list = -(haldane[rec1] - i)^2;
    Ordering[list, -1], (*entry with closest cM you can get*)
    {i, cMpicks}
  ] // Flatten;

  Table[
    Show[
      LogPlot[2 Nval Uval  $\frac{1}{x}$ , {x, 1, 100},
        PlotRange → {10-7, 10-3}, PlotStyle → Directive[Black]],
      ListLogPlot[msfs1[[r]], PlotRange → {10-7, 10-3},
        Joined → True, PlotStyle → Directive[Thick]],
      ListLogPlot[msfs2[[r]], PlotRange → {10-7, 10-3}, Joined → True,
        PlotStyle → Directive[Thick, defaultcolors[[2]]]],
      ListLogPlot[msfs3[[r]], PlotRange → {10-7, 10-3}, Joined → True,
        PlotStyle → Directive[Thick, defaultcolors[[3]]]],
      Frame → {True, True, False, False},
      FrameLabel → {
        "Copies of derived allele at site",
        If[r == rs[[1]], "Fraction of sites in window", ""]
      },
      FrameTicksStyle → {Automatic,
        If[r == rs[[1]], Automatic, FontColor → White], Automatic, Automatic},
      FrameStyle → Directive[FontSize → size, FontFamily → "Helvetica"],
      Epilog → {
        Text[Style[ToString[NumberForm[haldane[rec1][[r]]], {3, 1}]] <> " cM",

```

```

        Directive[FontSize → size, FontFamily → "Helvetica"]], Scaled@{0.5, 0.9}]
    }
  ],
  {r, rs}
]
)

```

## Plot linkage disequilibrium

```

In[146]:= getLD[Nval_, dval_, sval_, kval_, uval_, mval_, nreps_] :=
(
  folder = StringForm["K`_d`s`_k`_u`_m`", Nval, NumberForm[dval, {3, 2}],
    NumberForm[sval, {3, 2}], kval, If[0 < uval, NumberForm[uval, {6, 5}], 0],
    If[0 < mval < 1, NumberForm[mval, {3, 2}], mval]]; (*folder where data is*)
  Clear[data];
  data = Table[Import[datadir <> ToString[folder] <>
    "/data/ld_" <> ToString[i - 1] <> ".csv", "CSV"], {i, nreps}];
  data = Mean[data];
  data[[All, 1]] = haldane[data[[All, 1]]];
  data
)

plotLD[Nval_, dval_, sval_, kval_, uval_,
  mval_, nreps_, Nvalbottle_, save_, ymax_, letter_] :=
(
  xmin = -20; xmax = 20;
  (*frame styling*)
  xticks = Table[{x, Abs[x]}, {x, xmin, xmax, 10}];
  yticks = Automatic;
  style = Which[
    letter == "A",
    {
      FrameLabel → {"Linkage disequilibrium"},
      FrameTicksStyle → {FontColor → White, Automatic, Automatic, Automatic},
      Epilog → {
        Text[Style[letter, letterstyle], Scaled@letterposition]
      }
    },
    letter == "B",
    {
      FrameTicksStyle → {FontColor → White, FontColor → White, Automatic, Automatic},
      Epilog → {
        Text[Style[letter, letterstyle], Scaled@letterposition]
      }
    },
    letter == "C",
    {
      FrameLabel → {"Distance from selected site (cM)", "Linkage disequilibrium"},
      Epilog → {

```

```

        Text[Style[letter, letterstyle], Scaled@letterposition]
    }
},
letter == "D",
{
    FrameTicksStyle → {Automatic, FontColor → White, Automatic, Automatic},
    FrameLabel → {"Distance from selected site (cM)",},
    Epilog → {
        Text[Style[letter, letterstyle], Scaled@letterposition]
    }
}
];

data = {
    getLD[Nval, dval, sval, kval, uval, mval, nreps],
    getLD[Nval, 0, sval, kval, uval, mval, nreps],
    getLD[Nvalbottle, 0, sval, kval, uval, mval, nreps]
};

theory = 
$$\frac{4 N_e r + 10}{(4 N_e r)^2 + 13 \times 4 N_e r + 22}$$
 /. Ne → Nval * 4 / 7 /. r → 0.001;
(*eqn 7.31 in Wakeley 2009*)

plot =
Show[
    Plot[theory, {x, xmin, xmax},
        PlotStyle → Black, PlotRange → {{xmin, xmax}, {0, ymax}}],
    ListPlot[
        data,
        Joined → True,
        PlotStyle → Thick
    ],
    PlotRange → {{xmin, xmax}, {0, ymax}},
    Frame → {True, True, False, False},
    style,
    FrameStyle → labelstyle,
    PlotRangePadding → None,
    ImagePadding → Hpadding,
    Axes → False
];

If[save == 1,
    Export[imagedir <> ToString[
        StringForm["LD_s``k``u``m``.pdf", sval, kval, uval, mval]], plot]
];

Show[plot, ImageSize → 400]

)

```

Get Ne from simulations

```

In[148]:= getSimNe[Nval_, dval_, sval_, kval_, uval_, mval_, nreps_] :=
(
  (*get time reversed p and n dynamics, as in getSimDynamics[]*)
  folder = StringForm["K`_d`s`_k`_u`_m`", Nval, NumberForm[dval, {3, 2}],
    NumberForm[sval, {3, 2}], kval, If[0 < uval, NumberForm[uval, {6, 5}], 0],
    If[0 < mval < 1, NumberForm[mval, {2, 1}], mval]];
  Clear[data];
  Table[data[i] = Import[datadir <> ToString[folder] <> "/data/dynamics_" <>
    ToString[i - 1] <> ".txt", "Table", "FieldSeparators" -> " ", {i, nreps}];
  allp = Table[data[i][[2 ;; 3]], {i, nreps}];
  alln = Table[data[i][[2 ;; 2]], {i, nreps}];
  allpcut = Table[
    allp[[i]][[1 ;; Module[{n = 1}, While[allp[[i]][[n]] < 1, n++]; n - 1]]],
    {i, nreps}];
  allncut = Table[alln[[i]][[1 ;; Length[allpcut[[i]]]]], {i, nreps}];
  allpreversed = Reverse[allpcut, 2];
  allnreversed = Reverse[allncut, 2];
  allpreversedcut = Table[
    allpreversed[[i]][[1 ;; Module[{n = 1},
      While[allpreversed[[i]][[n]] > 1 / (2 Nval), n++] // Quiet; n - 1]]],
    {i, nreps}];
  allnreversedcut =
    Table[allnreversed[[i]][[1 ;; Length[allpreversedcut[[i]]]]], {i, nreps}];
  Mean[HarmonicMean /@ allnreversedcut // N]
)

```

## Methods

### Deterministic trajectories

Assume a lifecycle where viability selection is followed by random mating and reproduction. Let  $n_{aa}(t)$ ,  $n_{Aa}(t)$ ,  $n_{AA}(t)$  be the number of individuals of genotype aa, Aa, and AA at the beginning of generation  $t$ . Let the viability of these individuals be  $V_{aa} = \frac{W_{aa}}{B}$ ,  $V_{Aa} = \frac{W_{Aa}}{B}$ ,  $V_{AA} = \frac{W_{AA}}{B}$ , where  $B$  is the number of offspring produced by each surviving individual. We scale by  $B$  so that the absolute fitness of genotype  $ij$  is simply  $B V_{ij} = W_{ij}$ . After selection the expected number of individuals of each type are  $\tilde{n}_i(t) = n_i(t) V_i$   $\forall i \in \{aa, Aa, AA\}$ . Let  $p_{i,j}(k)$  be the probability a mating between genotype  $i$  and  $j$  makes genotype  $k$  following fair Mendelian segregation. Then the number of  $k$  genotypes produced by a surviving individual of genotype  $i$  is  $B \sum_j \frac{\tilde{n}_j(t)}{n(t)} p_{i,j}(k)$  where  $\tilde{n}(t) = \sum_i \tilde{n}_i(t)$  is the total population size after selection. Summing over all parents, the expected number of individuals of type  $i$  after reproduction (ie at the start of generation  $t+1$ ) is  $n_i(t+1) = \sum_k \tilde{n}_k(t) B \sum_j \frac{\tilde{n}_j(t)}{n(t)} p_{k,j}(i)$ . Expanding this out and simplifying gives



```

In[149]:= Table[Sum[n_i[t] B  $\frac{\tilde{n}_j[t]}{\tilde{n}[t]}$  p_{i,j}[k], {i, {aa, Aa, AA}}, {j, {aa, Aa, AA}}],
               {k, {aa, Aa, AA}}] /. p_{Aa,Aa}[Aa] -> 1/2 /. p_{Aa,Aa}[i_] -> 1/4 /.
               p_{i_,i_}[i_] -> 1 /. p_{i_,i_}[j_] -> 0 /. p_{aa,AA}[Aa] -> 1 /.
               p_{aa,AA}[i_] -> 0 /. p_{AA,aa}[Aa] -> 1 /. p_{AA,aa}[i_] -> 0 /.
               p_{aa,Aa}[aa] -> 1/2 /. p_{aa,Aa}[Aa] -> 1/2 /. p_{aa,Aa}[i_] -> 0 /.
               p_{Aa,aa}[aa] -> 1/2 /. p_{Aa,aa}[Aa] -> 1/2 /. p_{Aa,aa}[i_] -> 0 /.
               p_{AA,Aa}[AA] -> 1/2 /. p_{AA,Aa}[Aa] -> 1/2 /. p_{AA,Aa}[i_] -> 0 /.
               p_{Aa,AA}[AA] -> 1/2 /. p_{Aa,AA}[Aa] -> 1/2 /. p_{Aa,AA}[i_] ->
0 // Simplify;
nnew = % /. n[t] -> Sum[n_i[t], {i, {aa, Aa, AA}}] /. n_i[t] -> n_i[t] v_i (*W_i/B*) // Simplify
Out[150]= {  $\frac{B (2 V_{aa} n_{aa}[t] + V_{Aa} n_{Aa}[t])^2}{4 (V_{aa} n_{aa}[t] + V_{Aa} n_{Aa}[t] + V_{AA} n_{AA}[t])}$ ,
            (B (2 V_{aa} n_{aa}[t] + V_{Aa} n_{Aa}[t]) (V_{Aa} n_{Aa}[t] + 2 V_{AA} n_{AA}[t])) /
            (2 (V_{aa} n_{aa}[t] + V_{Aa} n_{Aa}[t] + V_{AA} n_{AA}[t])) ,  $\frac{B (V_{Aa} n_{Aa}[t] + 2 V_{AA} n_{AA}[t])^2}{4 (V_{aa} n_{aa}[t] + V_{Aa} n_{Aa}[t] + V_{AA} n_{AA}[t])}$  }

```

Now write the denominator as  $\tilde{n}(t) = n(t) \bar{W}(t)/B$ , where  $\bar{W}(t)$  is the population mean fitness in generation  $t$

```

In[151]:= nnew /. V_{aa} n_{aa}[t] + V_{Aa} n_{Aa}[t] + V_{AA} n_{AA}[t] -> n[t] w[t] / B
Out[151]= {  $\frac{B^2 (2 V_{aa} n_{aa}[t] + V_{Aa} n_{Aa}[t])^2}{4 n[t] \bar{W}[t]}$ ,
            (B^2 (2 V_{aa} n_{aa}[t] + V_{Aa} n_{Aa}[t]) (V_{Aa} n_{Aa}[t] + 2 V_{AA} n_{AA}[t])) / (2 n[t] \bar{W}[t]) ,
             $\frac{B^2 (V_{Aa} n_{Aa}[t] + 2 V_{AA} n_{AA}[t])^2}{4 n[t] \bar{W}[t]}$  }

```

We next divide both numerator and denominator by  $n(t)^2$  giving

```

In[152]:= nnew /. V_{aa} n_{aa}[t] + V_{Aa} n_{Aa}[t] + V_{AA} n_{AA}[t] -> n[t] w[t] / B /. n_i[t] -> p_i[t] /. n[t] -> 1/n[t]
Out[152]= {  $\frac{B^2 n[t] (2 V_{aa} p_{aa}[t] + V_{Aa} p_{Aa}[t])^2}{4 \bar{W}[t]}$ ,  $\frac{1}{2 \bar{W}[t]}$ ,
            B^2 n[t] (2 V_{aa} p_{aa}[t] + V_{Aa} p_{Aa}[t]) (V_{Aa} p_{Aa}[t] + 2 V_{AA} p_{AA}[t]) ,
             $\frac{B^2 n[t] (V_{Aa} p_{Aa}[t] + 2 V_{AA} p_{AA}[t])^2}{4 \bar{W}[t]}$  }

```

where  $p_i(t)$  is the frequency of genotype  $i$  at the start of generation  $t$ .

Using  $W = V B$  gives

```
In[153]:= nnew /. Vaa naa[t] + VAa nAa[t] + VAA nAA[t] → n[t]  $\bar{w}$ [t] / B /. ni → pi[t] /.
n[t] → 1 / n[t];
nnewsimp = % /. Vi → Wi / B // Simplify
```

$$\text{Out[154]} = \left\{ \frac{n[t] (2 W_{aa} p_{aa}[t] + W_{Aa} p_{Aa}[t])^2}{4 \bar{w}[t]}, \frac{1}{2 \bar{w}[t]} \right. \\ \left. n[t] (2 W_{aa} p_{aa}[t] + W_{Aa} p_{Aa}[t]) (W_{Aa} p_{Aa}[t] + 2 W_{AA} p_{AA}[t]), \frac{n[t] (W_{Aa} p_{Aa}[t] + 2 W_{AA} p_{AA}[t])^2}{4 \bar{w}[t]} \right\}$$

Thus the number of a and A alleles in generation t+1 are expected to be

```
In[155]:= {2 nnewsimp[[1]] + nnewsimp[[2]], nnewsimp[[2]] + 2 nnewsimp[[3]]} // FullSimplify
% /. (Waa paa[t] + WAa pAa[t] + WAA pAA[t]) →  $\bar{w}$ [t]
```

$$\text{Out[155]} = \left\{ \frac{1}{\bar{w}[t]} n[t] (2 W_{aa} p_{aa}[t] + W_{Aa} p_{Aa}[t]) (W_{Aa} p_{Aa}[t] + W_{Aa} p_{Aa}[t] + W_{AA} p_{AA}[t]), \right. \\ \left. \frac{1}{\bar{w}[t]} n[t] (W_{aa} p_{aa}[t] + W_{Aa} p_{Aa}[t] + W_{AA} p_{AA}[t]) (W_{Aa} p_{Aa}[t] + 2 W_{AA} p_{AA}[t]) \right\}$$

$$\text{Out[156]} = \{n[t] (2 W_{aa} p_{aa}[t] + W_{Aa} p_{Aa}[t]), n[t] (W_{Aa} p_{Aa}[t] + 2 W_{AA} p_{AA}[t])\}$$

Using  $n(t+1) = n(t) \bar{w}(t)$  the frequencies in generation t+1 are

```
In[157]:= {2 nnewsimp[[1]] + nnewsimp[[2]], nnewsimp[[2]] + 2 nnewsimp[[3]]} // FullSimplify;
% /. (Waa paa[t] + WAa pAa[t] + WAA pAA[t]) →  $\bar{w}$ [t]
% / (2 n[t+1]) /. n[t] → n[t+1] /  $\bar{w}$ [t]
```

$$\text{Out[158]} = \{n[t] (2 W_{aa} p_{aa}[t] + W_{Aa} p_{Aa}[t]), n[t] (W_{Aa} p_{Aa}[t] + 2 W_{AA} p_{AA}[t])\}$$

$$\text{Out[159]} = \left\{ \frac{2 W_{aa} p_{aa}[t] + W_{Aa} p_{Aa}[t]}{2 \bar{w}[t]}, \frac{W_{Aa} p_{Aa}[t] + 2 W_{AA} p_{AA}[t]}{2 \bar{w}[t]} \right\}$$

which we will call  $p(t+1)$ .

Thus the number of genotypes in generation t+1 can be written

```
In[160]:= nnewsimp /. (2 Waa paa[t] + WAa pAa[t]) → 2  $\bar{w}$ [t] p[t+1] /.
(WAa pAa[t] + 2 WAA pAA[t]) → 2  $\bar{w}$ [t] q[t+1] /. n[t] → n[t+1] /  $\bar{w}$ [t]
```

$$\text{Out[160]} = \{n[1+t] p[1+t]^2, 2 n[1+t] p[1+t] q[1+t], n[1+t] q[1+t]^2\}$$

giving frequencies

```
In[161]:= nnewsimp / n[t+1] /. (2 Waa paa[t] + WAa pAa[t]) → 2  $\bar{w}$ [t] p[t+1] /.
(WAa pAa[t] + 2 WAA pAA[t]) → 2  $\bar{w}$ [t] q[t+1] /. n[t] → n[t+1] /  $\bar{w}$ [t]
```

$$\text{Out[161]} = \{p[1+t]^2, 2 p[1+t] q[1+t], q[1+t]^2\}$$

Thus Hardy-Weinberg has been achieved and we can just track allele frequencies.

Defining  $p[t] W_{aa} + q[t] W_{Aa} = W_a[t]$  as the marginal fitness of allele a in generation t (similarly for A) the allele frequency recursions are

```
In[162]:= {2 nnewsimp[[1]] + nnewsimp[[2]], nnewsimp[[2]] + 2 nnewsimp[[3]]} // FullSimplify;
% /. (Waa Paa[t] + WAa PAa[t] + WAA PAA[t]) ->  $\bar{W}[t]$ ;
% / (2 n[t+1]) /. n[t] -> n[t+1] /  $\bar{W}[t]$ ;
% /.  $\bar{W}[t] \rightarrow W_{aa} P_{aa}[t] + W_{Aa} P_{Aa}[t] + W_{AA} P_{AA}[t]$  /. Paa[t] -> p[t]2 /. PAa[t] -> 2 p[t] q[t] /.
PAA[t] -> q[t]2 /. Waa ->  $\frac{W_a[t] - q[t] W_{Aa}}{p[t]}$  /. WAA ->  $\frac{W_A[t] - p[t] W_{Aa}}{q[t]}$  // Simplify
```

$$\text{Out[165]} = \left\{ \frac{p[t] W_a[t]}{p[t] W_a[t] + q[t] W_A[t]}, \frac{q[t] W_A[t]}{p[t] W_a[t] + q[t] W_A[t]} \right\}$$

Note that this simplifies (because marginal fitness is no longer time dependent) and reduces to a haploid model when  $W_{ij} = W_i W_j$

```
In[166]:= {2 nnewsimp[[1]] + nnewsimp[[2]], nnewsimp[[2]] + 2 nnewsimp[[3]]} // FullSimplify;
% /. (Waa Paa[t] + WAa PAa[t] + WAA PAA[t]) ->  $\bar{W}[t]$ ;
% / (2 n[t+1]) /. n[t] -> n[t+1] /  $\bar{W}[t]$ ;
% /.  $\bar{W}[t] \rightarrow W_{aa} P_{aa}[t] + W_{Aa} P_{Aa}[t] + W_{AA} P_{AA}[t]$  /. Paa[t] -> p[t]2 /.
PAa[t] -> 2 p[t] q[t] /. PAA[t] -> q[t]2;
% /. Waa -> Wa Wa /. WAa -> WA Wa /. WAA -> WA WA // Simplify
```

$$\text{Out[170]} = \left\{ \frac{p[t] W_a}{p[t] W_a + q[t] W_A}, \frac{q[t] W_A}{p[t] W_a + q[t] W_A} \right\}$$

meaning that the allele frequency dynamics in a haploid model with  $W_a = 1$  and  $W_A = 1 + s$  is equivalent to the allele frequency dynamics of a diploid model with  $W_{aa} = 1$ ,  $W_{Aa} = 1 + s$ ,  $W_{AA} = (1 + s)^2 \sim 1 + 2s$ , i.e., with additive fitness that is twice as strong in diploids (because selection is half as efficient there).

With our parameters the frequency of the beneficial allele in generation t+1 is

```
In[171]:= {2 nnewsimp[[1]] + nnewsimp[[2]], nnewsimp[[2]] + 2 nnewsimp[[3]]} // FullSimplify;
% /. (Waa Paa[t] + WAa PAa[t] + WAA PAA[t]) ->  $\bar{W}[t]$ ;
% / (2 n[t+1]) /. n[t] -> n[t+1] /  $\bar{W}[t]$ ;
% /.  $\bar{W}[t] \rightarrow W_{aa} P_{aa}[t] + W_{Aa} P_{Aa}[t] + W_{AA} P_{AA}[t]$  /. Paa[t] -> p[t]2 /. PAa[t] -> 2 p[t] q[t] /.
PAA[t] -> q[t]2
qnew = %[[2]] /. p[t] -> 1 - q[t] /. Waa -> 1 - d /. WAA -> (1 - d) (1 + s h) /.
WAa -> (1 - d) (1 + s) // Simplify
```

$$\text{Out[174]} = \left\{ \frac{2 p[t]^2 W_{aa} + 2 p[t] q[t] W_{Aa}}{2 (p[t]^2 W_{aa} + 2 p[t] q[t] W_{Aa} + q[t]^2 W_{AA})}, \frac{2 p[t] q[t] W_{Aa} + 2 q[t]^2 W_{AA}}{2 (p[t]^2 W_{aa} + 2 p[t] q[t] W_{Aa} + q[t]^2 W_{AA})} \right\}$$

$$\text{Out[175]} = \frac{q[t] (-1 - h s + (-1 + h) s q[t])}{-1 - 2 h s q[t] + (-1 + 2 h) s q[t]^2}$$

and the decline rate of the wildtype drops out (ie only relative fitness matters for allele frequency dynamics).

And the change in allele frequencies are (equation 3.1 in Gillespie, equations 5.2.13 and 5.2.18 in Crow & Kimura, equation 3.13a,b in Otto & Day)

```
In[176]:= {2 nnewsimp[[1]] + nnewsimp[[2]], nnewsimp[[2]] + 2 nnewsimp[[3]]} // FullSimplify;
% /. (Waa Paa[t] + Waa PAA[t] + WAA PAA[t]) -> W[t];
% / (2 n[t+1]) /. n[t] -> n[t+1] / W[t];
% - {p[t], q[t]} /. W[t] -> Waa Paa[t] + Waa PAA[t] + WAA PAA[t] /. Paa[t] -> p[t]^2 /.
PAA[t] -> 2 p[t] q[t] /. PAA[t] -> q[t]^2 // FullSimplify;
{(p[t] q[t] (p[t] (Waa - WAA) + q[t] (WAA - WAA))) / (p[t]^2 Waa + 2 p[t] q[t] WAA + q[t]^2 WAA),
(p[t] q[t] (p[t] (WAA - Waa) + q[t] (WAA - WAA))) / (p[t]^2 Waa + 2 p[t] q[t] WAA + q[t]^2 WAA)}
% == %% /. q[t] -> 1 - p[t] // Simplify
{(p[t] q[t] (p[t] (Waa - WAA) + q[t] (WAA - WAA))) / (p[t]^2 Waa + 2 p[t] q[t] WAA + q[t]^2 WAA),
(p[t] q[t] (p[t] (WAA - Waa) + q[t] (WAA - WAA))) /
(p[t]^2 Waa + 2 p[t] q[t] WAA + q[t]^2 WAA)} /. Waa -> 1 - d /.
WAA -> (1 - d) (1 + s h) /. WAA -> (1 - d) (1 + s) // Simplify
Out[180]= {(p[t] q[t] (p[t] (Waa - WAA) + q[t] (WAA - WAA))) / (p[t]^2 Waa + 2 p[t] q[t] WAA + q[t]^2 WAA),
(p[t] q[t] (p[t] (-Waa + WAA) + q[t] (-WAA + WAA))) / (p[t]^2 Waa + 2 p[t] q[t] WAA + q[t]^2 WAA)}
```

```
Out[181]= True
```

```
Out[182]= {- (s p[t] q[t] (h p[t] + q[t] - h q[t])) / (p[t]^2 + 2 (1 + h s) p[t] q[t] + (1 + s) q[t]^2), (s p[t] q[t] (h p[t] + q[t] - h q[t])) / (p[t]^2 + 2 (1 + h s) p[t] q[t] + (1 + s) q[t]^2)}
```

When the beneficial allele is additive the change in allele frequency is

```
In[183]:= qnew - q[t] /. h -> 1 / 2 // Simplify
```

```
Out[183]= - (s (-1 + q[t]) q[t]) / (2 + 2 s q[t])
```

which with  $s$  small is nearly (eqn 5.3.8 in Crow & Kimura)

```
In[184]:= qnew - q[t] /. h -> 1 / 2 // Simplify;
Series[%, {s, 0, 1}] // Normal // Simplify
```

```
Out[185]= - (1/2) s (-1 + q[t]) q[t]
```

and therefore has approximate solution (eqn 5.3.12 in Crow & Kimura)

```
In[186]:= qnew - q[t] /. h -> 1 / 2 // Simplify;
Series[%, {s, 0, 1}] // Normal;
qtadditive = DSolve[{D[q[t], t] == %, q[0] == q0}, q[t], t] // Simplify // Flatten
(q[t] /. qtadditive) == (1 + (1 - q0 / q0) Exp[-(s / 2) t])^-1 // Simplify
```

Solve::ifun : Inverse functions are being used by Solve, so

some solutions may not be found; use Reduce for complete solution information. >>

```
Out[188]= {q[t] -> (e^(s t / 2) q0) / (1 + (-1 + e^(s t / 2)) q0)}
```

```
Out[189]= True
```

The population size in the next generation, ignoring density-dependence, is mean fitness times the current size

```
In[190]:= wbar = Waa Paa[t] + Waa Paa[t] + WAA PAA[t] /. Paa[t] -> p[t]^2 /. Paa[t] -> 2 p[t] q[t] /.
          PAA[t] -> q[t]^2 /. p[t] -> 1 - q[t] /. Waa -> 1 - d /.
          WAA -> (1 - d) (1 + s h) /. WAA -> (1 - d) (1 + s) // Simplify;
wbar
n[
t]
```

```
Out[191]= (-1 + d) n[t] (-1 - 2 h s q[t] + (-1 + 2 h) s q[t]^2)
```

So that with additivity we have

```
In[192]:= wbar n[t];
% /. h -> 1 / 2
```

```
Out[193]= (-1 + d) n[t] (-1 - s q[t])
```

Subbing in the weak additive selection allele frequency dynamics gives

```
In[194]:= wbar n[t];
% /. h -> 1 / 2;
% /. qtadditive
```

```
Out[196]= (-1 + d) \left( -1 - \frac{e^{\frac{s t}{2}} q_0 s}{1 + \left( -1 + e^{\frac{s t}{2}} \right) q_0} \right) n[t]
```

So that population size at time t is roughly the minimum of N0 (the carrying capacity) and

```
In[197]:= wbar n[t];
% /. h -> 1 / 2;
% /. qtadditive;
% - n[t];
DSolve[{D[n[t], t] == %, n[0] == N0}, n[t], t] // Flatten // Simplify
Simplify[(n[t] /. %) == (N0 Exp[((1 + s) (1 - d) - 1) t] \left( \frac{q_0}{q[t]} \right)^{2 (1 - d)} /. qtadditive),
{0 < d < 1, s > 0, t > 0}]
ntadditive = {n[t] -> e^{-d t} N0 \left( 1 + q_0 \left( e^{\frac{s t}{2}} - 1 \right) \right)^{2 - 2 d}};
```

```
Out[201]= {n[t] -> e^{-d t} N0 \left( 1 + \left( -1 + e^{\frac{s t}{2}} \right) q_0 \right)^{2 - 2 d}}
```

```
Out[202]= True
```

## Probability of establishment

In a branching process the probability of establishment is (see Allen 2010 page 172)

```
In[204]:= pest = 1 - Exp[-2 (w - 1) / v];
```

where  $w > 1$  is the expected number of copies contributed to the next generation and  $v$  is the variance in this number.

## Genetic drift in the simulated lifecycle

### Wright-Fisher

In the Wright-Fisher with population size  $n$ , a rare allele with selective advantage  $s$  leaves  $X$  copies where  $X$  is a binomial with  $2n$  trials (the number of gametes in the next generation) and probability of success  $\frac{1+s}{n} \frac{1}{2}$  (the first factor is the fitness of this genotype divided by the sum of fitnesses over all individuals and the second factor is Mendelian inheritance). Thus it is expected to leave

```
In[205]:= Expectation[X, X ∈ BinomialDistribution[2 n,  $\frac{1+s}{n} \frac{1}{2}$ ]]
```

```
Out[205]= 1 + s
```

copies, with a variance of

```
In[206]:= Expectation[X, X ∈ BinomialDistribution[2 n,  $\frac{1+s}{n} \frac{1}{2}$ ]];
```

```
Expectation[X^2, X ∈ BinomialDistribution[2 n,  $\frac{1+s}{n} \frac{1}{2}$ ]] - %^2 // Simplify
```

```
Out[207]=  $\frac{(-1 + 2 n - s) (1 + s)}{2 n}$ 
```

which, when the population size is large, is

```
In[208]:= Expectation[X, X ∈ BinomialDistribution[2 n,  $\frac{1+s}{n} \frac{1}{2}$ ]];
```

```
Expectation[X^2, X ∈ BinomialDistribution[2 n,  $\frac{1+s}{n} \frac{1}{2}$ ]] - %^2 // Simplify;
```

```
Series[%, {n, ∞, 0}]
```

```
Out[210]=  $(1 + s) + O\left[\frac{1}{n}\right]^1$ 
```

which is then nearly 1 with weak selection.

### Our model, below carrying capacity

Let's say we start with  $n$  individuals at the census. The next event that takes place is selection, where an individual with viability  $V_i$  survives with probability  $V_i$ . The number of surviving individuals is then Poisson-binomially distributed with expectation  $n' = \sum_{i=1}^n V_i = n \bar{V}$ , where  $\bar{V}$  is the mean viability. Given that an individual survives it will have  $B$  offspring "as a mother" and will be chosen as a "father" for each of the  $n'' = n' B$  expected offspring with probability  $1/n'$ . Considering one allele in a diploid individual it will be passed down to each offspring that individual parents with probability  $1/2$ . Thus, the number of times this focal allele is passed on as a maternal allele is binomial with  $B$  trials and success probability  $1/2$ . The number of times it is passed down as a father is binomial with  $n'B$  trials and success probability  $1/(2n')$ . If  $n''$  is greater than the carrying capacity,  $n_0$ , then each offspring survives to the census with probability  $n''/n_0$ .

The expectation for  $n''$  is  $n' B = n \bar{V} B$ . We will use  $V_i = \frac{W_i}{B}$ , so that the expectation for  $n''$  becomes  $n \bar{W}$ . Thus if we start with a population of wildtypes, with  $W=1-d<1$ , we can ignore carrying capacity. Then

the expected number of copies produced by an allele in an individual with viability  $V=W/B$  is

```
In[211]:= Expectation[X (Y + Z),
  {X ∈ BernoulliDistribution[V], Y ∈ BinomialDistribution[B, 1 / 2],
   Z ∈ BinomialDistribution[np B,  $\frac{1}{np} \frac{1}{2}$ ]}] /. V → W / B
```

Out[211]=  $W$

and the variance is

```
In[212]:= Expectation[X (Y + Z), {X ∈ BernoulliDistribution[V],
  Y ∈ BinomialDistribution[B, 1 / 2], Z ∈ BinomialDistribution[np B,  $\frac{1}{np} \frac{1}{2}$ ]}];
Expectation[(X (Y + Z))^2, {X ∈ BernoulliDistribution[V],
  Y ∈ BinomialDistribution[B, 1 / 2], Z ∈ BinomialDistribution[np B,  $\frac{1}{np} \frac{1}{2}$ ]}] -
  %^2 /. V → W / B /. np → n Wbar / B // Simplify
```

Out[213]=  $\frac{1}{4} W \left( 3 - 4 W + B \left( 4 - \frac{1}{n Wbar} \right) \right)$

where we've used the expectation for  $n'$ .

Taking the large population limit this becomes

```
In[214]:= Expectation[X (Y + Z), {X ∈ BernoulliDistribution[V],
  Y ∈ BinomialDistribution[B, 1 / 2], Z ∈ BinomialDistribution[np B,  $\frac{1}{np} \frac{1}{2}$ ]}];
Expectation[(X (Y + Z))^2, {X ∈ BernoulliDistribution[V],
  Y ∈ BinomialDistribution[B, 1 / 2], Z ∈ BinomialDistribution[np B,  $\frac{1}{np} \frac{1}{2}$ ]}] -
  %^2 /. V → W / B /. np → n Wbar / B // Simplify;
Series[%, {n, ∞, 0}] // Simplify
```

Out[216]=  $\frac{1}{4} (3 + 4 B - 4 W) W + O\left[\frac{1}{n}\right]^1$

Under weak selection this is roughly

```
In[217]:= Expectation[X (Y + Z), {X ∈ BernoulliDistribution[V],
      Y ∈ BinomialDistribution[B, 1 / 2], Z ∈ BinomialDistribution[np B,  $\frac{1}{np} \frac{1}{2}$ ]}];
Expectation[(X (Y + Z))^2, {X ∈ BernoulliDistribution[V],
      Y ∈ BinomialDistribution[B, 1 / 2], Z ∈ BinomialDistribution[np B,  $\frac{1}{np} \frac{1}{2}$ ]}] -
      %^2 /. V → W / B /. np → n Wbar / B // Simplify;
Series[%, {n, ∞, 0}] // Simplify;
% /. W → 1
Out[220]=  $\frac{1}{4} (-1 + 4 B) + O\left[\frac{1}{n}\right]^1$ 
```

And we will primarily use B=2 in our simulations, giving

```
In[221]:= Expectation[X (Y + Z), {X ∈ BernoulliDistribution[V],
      Y ∈ BinomialDistribution[B, 1 / 2], Z ∈ BinomialDistribution[np B,  $\frac{1}{np} \frac{1}{2}$ ]}];
Expectation[(X (Y + Z))^2, {X ∈ BernoulliDistribution[V],
      Y ∈ BinomialDistribution[B, 1 / 2], Z ∈ BinomialDistribution[np B,  $\frac{1}{np} \frac{1}{2}$ ]}] -
      %^2 /. V → W / B /. np → n Wbar / B // Simplify;
Series[%, {n, ∞, 0}] // Simplify;
% /. W → 1;
% /. B → 2
Out[225]=  $\frac{7}{4} + O\left[\frac{1}{n}\right]^1$ 
```

meaning we have roughly 7/4 times as much drift in our model as compared to the Wright-Fisher. Thus our variance effective population size is 4/7 times the census size.

## Our model, at carrying capacity

Consider the same process but with carrying capacity (because  $\bar{W} > 1$ ). Then we can add the probability an offspring survives carrying capacity by multiplying the probability of success in each binomial by  $\frac{1}{W}$ . The expected number of alleles remaining after carrying capacity is then

```
In[226]:= Expectation[X (Y + Z),
      {X ∈ BernoulliDistribution[V], Y ∈ BinomialDistribution[B,  $\frac{1}{2} \frac{1}{Wbar}$ ],
      Z ∈ BinomialDistribution[np B,  $\frac{1}{np} \frac{1}{2} \frac{1}{Wbar}$ ]}] /. V → W / B
Out[226]=  $\frac{W}{Wbar}$ 
```

and the variance is



```
In[227]:= Expectation[X (Y + Z),
  {X ∈ BernoulliDistribution[V], Y ∈ BinomialDistribution[B, 1/2 Wbar],
   Z ∈ BinomialDistribution[np B, 1/2 Wbar]}];
Expectation[(X (Y + Z))^2, {X ∈ BernoulliDistribution[V],
  Y ∈ BinomialDistribution[B, 1/2 Wbar], Z ∈ BinomialDistribution[
   np B, 1/2 Wbar]}] - %^2 /. V → W / B /. np → n Wbar / B // Simplify
```

$$\text{Out[228]} = -\frac{W (B - 4 B n Wbar + n (1 + 4 W - 4 Wbar) Wbar)}{4 n Wbar^3}$$

Again taking the large population limit we have

```
In[229]:= Expectation[X (Y + Z),
  {X ∈ BernoulliDistribution[V], Y ∈ BinomialDistribution[B, 1/2 Wbar],
   Z ∈ BinomialDistribution[np B, 1/2 Wbar]}];
Expectation[(X (Y + Z))^2, {X ∈ BernoulliDistribution[V],
  Y ∈ BinomialDistribution[B, 1/2 Wbar], Z ∈ BinomialDistribution[
   np B, 1/2 Wbar]}] - %^2 /. V → W / B /. np → n Wbar / B // Simplify;
Series[%, {n, ∞, 0}] // Simplify
```

$$\text{Out[231]} = \frac{W (-1 + 4 B - 4 W + 4 Wbar)}{4 Wbar^2} + O\left[\frac{1}{n}\right]^1$$

Under weak selection this reduces to what we had in the absence of carrying capacity

```

In[232]:= Expectation[X (Y + Z),
  {X ∈ BernoulliDistribution[V], Y ∈ BinomialDistribution[B, 1/2 Wbar],
  Z ∈ BinomialDistribution[np B, 1/2 Wbar]}];
Expectation[(X (Y + Z))^2, {X ∈ BernoulliDistribution[V],
  Y ∈ BinomialDistribution[B, 1/2 Wbar], Z ∈ BinomialDistribution[
  np B, 1/2 Wbar]}] - %^2 /. V → W / B /. np → n Wbar / B // Simplify;
Series[%, {n, ∞, 0}] // Simplify;
% /. Wbar → 1 /. W → 1
% /. B → 2
Out[235]= 1/4 (-1 + 4 B) + O[1/n]^1
Out[236]= 7/4 + O[1/n]^1

```

## Backward-time dynamics

Fixation, when it happens, will actually tend to occur more quickly than expected. We can correct for this by setting the fixation frequency to be  $1 - \frac{1}{2N_0\rho_2}$ , where  $\rho_2$  is the backwards fixation probability of the heterozygote mutant in a population of homozygote mutants. Or more generally, we can set the final frequency as  $q_f$ , so that the fixation time is then

```

In[237]:= qf == q[t] /. qtadditive;
tfixadditive = t /. Simplify[Solve[%, t] /. C[1] → 0 // Flatten]
Out[238]= 2 Log[(-1+q0) qf / (q0 (-1+qf))] / s

```

and the backwards time allele frequency dynamics are

```

In[239]:= q[t] /. qtadditive /. t → tfixadditive - τ // Simplify
qtadditiveback[τ_] := (* (qf e^(-s τ/2) / ((1-qf) + e^(-s τ/2) qf)) *) (1 + ((1-qf) / qf) e^(s τ/2))^-1
qtadditiveback[τ] == %% // Simplify
Out[239]= - qf / (e^(s τ/2) (-1 + qf) - qf)
Out[241]= True

```

We can also compute the population size back in time

```

In[242]:= n[t] /. ntadditive /. t -> tfixadditive - τ // Simplify

ntadditiveback[τ_] := ed τ N0  $\left( \frac{(1 - q_0) q_f}{q_0 (1 - q_f)} \right)^{-\frac{2d}{s}} \left( (1 - q_0) \left( 1 + \frac{q_f}{1 - q_f} e^{-\frac{s \tau}{2}} \right) \right)^{2(1-d)}$ 

Simplify[ntadditiveback[τ] == %, {d < 1, 0 < q0 < 1}]
Simplify[%%% == N0 Exp[-d (tfixadditive - τ)] ((1 - q0) / (1 - qtadditiveback[τ]))2(1-d),
{0 < d < 1, s > 0, t > 0}]

Out[242]= ed τ N0  $\left( \frac{(-1 + q_0) q_f}{q_0 (-1 + q_f)} \right)^{-\frac{2d}{s}} \left( 1 - q_0 + \frac{e^{-\frac{s \tau}{2}} (-1 + q_0) q_f}{-1 + q_f} \right)^{2-2d}$ 

Out[244]= True

Out[245]= True

```

## Deriving rates of coalescence, recombination, mutation, and migration

### Setup

We have a lifecycle that has

1. census
  2. selection
  3. recombination
  4. syngamy
  5. mutation
  7. migration
- (see Figure S1).

We will be using backwards in time notation here (consistent with Pennings and Hermisson 2006 MBE; hereafter PH2). Given beneficial allele frequency  $x[\tau]$  and population size  $n[\tau]$  in generation  $\tau$  (in the main text we use  $p'(\tau)$  and  $N'(\tau)$ ) we will give recursions for the frequency of a beneficial allele in the next generation,  $x[\tau-1]$ . We denote the time of the artificial generation between selection and recombination as  $\tau-1/5$ , between recombination and syngamy as  $\tau-2/5$ , between syngamy and mutation as  $\tau-3/5$ , and between mutation and migration as  $\tau-4/5$ .

### Migration

The number of migrant alleles that arrive each generation is Poisson with mean  $m$ . Migrant alleles that arrive replace alleles already in the population. If the population size is  $n$ , the probability an allele is replaced by a migrant is therefore

```

In[246]:= Expectation[ $\frac{M}{n}$ , M ∈ PoissonDistribution[m]]

Out[246]=  $\frac{m}{n}$ 

```

The number of beneficial alleles immediately after migration is (which, ignoring carrying capacity is the number in the next generation)

$$\begin{aligned} \text{In[247]:= } 2 n[\tau - 1] x[\tau - 1] &= 2 n[\tau - 4 / 5] x[\tau - 4 / 5] \left( 1 - \frac{m}{2 n[\tau - 4 / 5]} \right) + \\ & 2 n[\tau - 4 / 5] x[\tau - 4 / 5] \left( \frac{m}{2 n[\tau - 4 / 5]} \right) + 2 n[\tau - 4 / 5] (1 - x[\tau - 4 / 5]) \left( \frac{m}{2 n[\tau - 4 / 5]} \right); \end{aligned}$$

where the first term is the previous migrants that have not been replaced, the second term is the migrants that have been replaced, and the third term is the non-migrants that have been replaced.

The probability a beneficial allele is a new migrant is therefore

$$\begin{aligned} \text{In[248]:= } & \left( 2 n[\tau - 4 / 5] x[\tau - 4 / 5] \left( \frac{m}{2 n[\tau - 4 / 5]} \right) + 2 n[\tau - 4 / 5] (1 - x[\tau - 4 / 5]) \left( \frac{m}{2 n[\tau - 4 / 5]} \right) \right) / \\ & (2 n[\tau - 1] x[\tau - 1]) // \text{Simplify} \\ \text{Out[248]= } & \frac{m}{2 n[-1 + \tau] x[-1 + \tau]} \end{aligned}$$

and assuming population size and allele frequency don't change much from generation to generation we have

$$\begin{aligned} \text{In[249]:= } \text{Pmig1} &= \frac{m}{2 n[-1 + \tau] x[-1 + \tau]} /. n[-1 + \tau] \rightarrow n[\tau] /. x[-1 + \tau] \rightarrow x[\tau] \\ \text{Out[249]= } \text{Pmig1} &= \frac{m}{2 n[\tau] x[\tau]} \end{aligned}$$

which is a diploid version of equation 15 in PH2, where they're  $m$  is our  $m/n[\tau]$  (they use  $m$  as the probability of being replaced).

The probability at least one of  $k$  beneficial alleles is a new migrant is then

$$\begin{aligned} \text{In[250]:= } 1 - (1 - \text{Pmig1})^k /. \text{Pmig1} &\rightarrow \frac{m}{2 n[\tau] x[\tau]} \\ \text{Series[\%, \{m, 0, 1\}} & \\ \text{Out[250]= } 1 - \left( 1 - \frac{m}{2 n[\tau] x[\tau]} \right)^k & \\ \text{Out[251]= } \frac{k m}{2 n[\tau] x[\tau]} + O[m]^2 & \end{aligned}$$

which gives equation 16 in PH2 (again with their  $m$  being replaced by  $m/n[\tau]$ ).

## Mutation

The number of beneficial alleles after mutation is the number before mutation plus the number of new mutants

$$\text{In[252]:= } 2 n[\tau - 4 / 5] x[\tau - 4 / 5] = 2 n[\tau - 3 / 5] x[\tau - 3 / 5] + u 2 n[\tau - 3 / 5] (1 - x[\tau - 3 / 5]);$$

so that the frequency of beneficial alleles after mutation is

$$\text{In[253]:= } x[\tau - 4 / 5] = \frac{n[\tau - 3 / 5]}{n[\tau - 4 / 5]} x[\tau - 3 / 5] + u \frac{n[\tau - 3 / 5]}{n[\tau - 4 / 5]} (1 - x[\tau - 3 / 5]);$$

Since the population size does not change during mutation, this simplifies to

In[254]:=  $x[\tau - 4/5] == x[\tau - 3/5] + u(1 - x[\tau - 3/5]);$

which is analogous to equation 2 in PH2.

The probability a beneficial allele is a new mutant (or the probability a neutral allele was on the ancestral background previously) is therefore

In[255]:= 
$$Pmut1 == \frac{u(1 - x[\tau - 3/5])}{x[\tau - 4/5]};$$

which, using the previous equation, is

In[256]:= `Solve[x[τ - 4/5] == x[τ - 3/5] + u(1 - x[τ - 3/5]), x[τ - 3/5]] // Flatten;`

$$Pmut1 == \frac{u(1 - x[\tau - 3/5])}{x[\tau - 4/5]} /. \% // Simplify$$

Out[257]= 
$$Pmut1 == \frac{u \left( -1 + x \left[ -\frac{4}{5} + \tau \right] \right)}{(-1 + u) x \left[ -\frac{4}{5} + \tau \right]}$$

which is analogous to equation 3 in PH2.

The probability that at least one of k beneficial alleles is a new mutant is

In[258]:= 
$$Pmut1 \rightarrow \frac{u \left( -1 + x \left[ -\frac{4}{5} + \tau \right] \right)}{(-1 + u) x \left[ -\frac{4}{5} + \tau \right]};$$

$$1 - (1 - Pmut1)^k /. \% // Simplify$$

Out[259]= 
$$1 - \left( \frac{u - x \left[ -\frac{4}{5} + \tau \right]}{(-1 + u) x \left[ -\frac{4}{5} + \tau \right]} \right)^k$$

When mutation is rare this is approximately

In[260]:= 
$$Pmut1 \rightarrow \frac{u \left( -1 + x \left[ -\frac{4}{5} + \tau \right] \right)}{(-1 + u) x \left[ -\frac{4}{5} + \tau \right]};$$

$$1 - (1 - Pmut1)^k /. \% // Simplify$$

`Normal@Series[%, {u, 0, 1}]`

$$\% == \frac{k u (1 - x[\tau - 4/5])}{x[\tau - 4/5]} // Simplify$$

Out[261]= 
$$1 - \left( \frac{u - x \left[ -\frac{4}{5} + \tau \right]}{(-1 + u) x \left[ -\frac{4}{5} + \tau \right]} \right)^k$$

Out[262]= 
$$u \left( -k + \frac{k}{x \left[ -\frac{4}{5} + \tau \right]} \right)$$

Out[263]= True

Assuming little change in allele frequency from one generation to the next gives Pmutk in equation 9 of PH2,

$$\text{In[264]:= } \frac{k u (1 - x[\tau - 4 / 5])}{x[\tau - 4 / 5]} /. x[\tau - 4 / 5] \rightarrow x[\tau]$$

$$\text{Out[264]:= } \frac{k u (1 - x[\tau])}{x[\tau]}$$

Note that while PH2 assume haploidy, with diploidy we both double the mutation rate as well as the number of alleles, which cancel.

## Coalescence

Given that none of the  $k$  beneficial alleles migrates or mutates the probability any two (and the neutral alleles on that background) coalesce is  $\frac{k(k-1)/2}{2 n e[\tau-2/5] x[\tau-2/5]}$ , where  $n_e$  is the effective population size. Thus the probability of coalescence is

$$\text{In[265]:= } \frac{\text{Binomial}[k, 2]}{2 n e[\tau - 2 / 5] x[\tau - 2 / 5]};$$

Assuming the effective population size and allele frequency change little from one generation to the next gives the diploid version of PcoalK given in equation 9 of PH2

$$\text{In[266]:= } \frac{\text{Binomial}[k, 2]}{2 n e[\tau - 2 / 5] x[\tau - 2 / 5]} /. n e[\tau - 2 / 5] \rightarrow n e[\tau] /. x[\tau - 2 / 5] \rightarrow x[\tau]$$

$$\text{Out[266]:= } \frac{(-1 + k) k}{4 n e[\tau] x[\tau]}$$

## Recombination

Consider a neutral locus recombination distance  $r$  from the selected site. Assuming weak selection such the survivors of viability selection are approximately in Hardy-Weinberg proportions, the number of alleles linked to the selected allele after recombination is

$$\text{In[267]:= } 2 n[\tau - 2 / 5] x[\tau - 2 / 5] == 2 n[\tau - 1 / 5] x[\tau - 1 / 5] (x[\tau - 1 / 5] + (1 - x[\tau - 1 / 5]) (1 - r)) + 2 n[\tau - 1 / 5] (1 - x[\tau - 1 / 5]) x[\tau - 1 / 5] r == 2 n[\tau - 1 / 5] x[\tau - 1 / 5];$$

where the first term on the RHS is the number currently linked to the selected allele multiplied by the probability of being in a homozygote plus the probability of being in a heterozygote and not recombining, and the second term on the RHS is the number not currently linked to the selected allele times the probability of being in a heterozygote and recombining.

Thus, the probability an allele on the selected background among the after recombination was not on that background before recombination is

$$\text{In[268]:= } \text{Precl} == \frac{2 n[\tau - 1 / 5] (1 - x[\tau - 1 / 5]) x[\tau - 1 / 5] r}{2 n[\tau - 2 / 5] x[\tau - 2 / 5]};$$

and since recombination doesn't change allele frequencies of population size this is

```
In[269]:= Prec1 == 
$$\frac{2n[\tau - 1/5] (1 - x[\tau - 1/5]) x[\tau - 1/5] r}{2n[\tau - 2/5] x[\tau - 2/5]} /. n[\tau - 2/5] \rightarrow n[\tau - 1/5] /. \\ x[\tau - 2/5] \rightarrow x[\tau - 1/5]$$

```

```
Out[269]= Prec1 == 
$$r \left( 1 - x \left[ -\frac{1}{5} + \tau \right] \right)$$

```

Therefore the probability at least one of k alleles on the selected background recombine off is

```
In[270]:= Prec1 -> 
$$r \left( 1 - x \left[ -\frac{1}{5} + \tau \right] \right);$$

1 - (1 - Prec1)^k /. % // Simplify
```

```
Out[271]= 
$$1 - \left( 1 - r + r x \left[ -\frac{1}{5} + \tau \right] \right)^k$$

```

Assuming recombination is rare gives

```
In[272]:= Prec1 -> 
$$r \left( 1 - x \left[ -\frac{1}{5} + \tau \right] \right);$$

1 - (1 - Prec1)^k /. % // Simplify;
Normal@Series[%, {r, 0, 1}]
% == 
$$r k \left( 1 - x \left[ -\frac{1}{5} + \tau \right] \right) // Simplify$$

```

```
Out[274]= 
$$r \left( k - k x \left[ -\frac{1}{5} + \tau \right] \right)$$

```

```
Out[275]= True
```

And assuming selection changes allele frequency slowly

```
In[276]:= Preck == 
$$r k \left( 1 - x \left[ -\frac{1}{5} + \tau \right] \right) /. x \left[ -\frac{1}{5} + \tau \right] \rightarrow x[\tau]$$

```

```
Out[276]= Preck == 
$$k r (1 - x[\tau])$$

```

which is analogous to the results found in Table 1 of Hudson & Kaplan 1988 (they consider 2 neutral loci so it's a bit more complicated).

## Event rates

A sample coalesces, recombines or mutates off the selected background, or migrates out of the population with probabilities that depend on the frequency of the beneficial background,  $x$ , the total number of chromosomes,  $2n$ , and the number of lineages remaining,  $k$ ,

```

In[277]:= pcoal[k_, τ_] := Binomial[k, 2]  $\frac{1}{2 n_e[\tau] x[\tau]}$ 

preco[k_, τ_] := k  $\frac{2 n[\tau] r (1 - x[\tau])}{2 n[\tau]}$ 

pmut[k_, τ_] := k  $\frac{2 n[\tau] u (1 - x[\tau])}{2 n[\tau] x[\tau]}$ 

pmig[k_, τ_] := k  $\frac{m}{2 n[\tau] x[\tau]}$ 

```

with  $\tau$  the number of generations before fixation (in the main text we replace  $n$  with  $N'$ ,  $n_e$  with  $N_e'$ , and  $x$  with  $p'$ ).

## Probability of no events until time $\tau$

The probability of no coalescence in  $\tau$  generations

```

In[281]:= cbackapprox[k_, τ_] :=  $\frac{1}{4 N_0 N_e N (-1 + q_0)^2 q f^3} (-1 + k) k (-1 + q f)^2 \left( \frac{(-1 + q_0) q f}{q_0 (-1 + q f)} \right)^{\frac{2d}{s}}$ 

 $\left( \frac{-e^{\frac{s\tau}{2}} (-1 + q f) + q f}{q f^2} \right)^{-2d} \left( \frac{1}{d - s + d s} q f \left( \frac{(-1 + q_0) (-e^{\frac{s\tau}{2}} (-1 + q f) + q f)}{(-1 + q f) q f} \right)^{2d} \right.$ 

 $\text{Hypergeometric2F1}\left[2 - 2d, 2 - \frac{2d(1+s)}{s}, 3 - \frac{2d(1+s)}{s}, \frac{-1 + q f}{q f}\right] - \frac{1}{d - s + d s}$ 

 $e^{(-d+s)\tau} q f \left( \frac{(-1 + q_0) \left(1 + (-1 + e^{-\frac{s\tau}{2}}) q f\right)}{(-1 + q f) q f} \right)^{2d} \text{Hypergeometric2F1}\left[2 - 2d,$ 

 $2 - \frac{2d(1+s)}{s}, 3 - \frac{2d(1+s)}{s}, \frac{e^{\frac{s\tau}{2}} (-1 + q f)}{q f}\right] - 1 / (-3s + 2d(1+s)) 2 e^{-d\tau}$ 

 $(-1 + q f) \left( e^{d\tau} \left( \frac{(-1 + q_0) (-e^{\frac{s\tau}{2}} (-1 + q f) + q f)}{(-1 + q f) q f} \right)^{2d} \text{Hypergeometric2F1}\left[2 - 2d,$ 

 $3 - \frac{2d(1+s)}{s}, 4 - \frac{2d(1+s)}{s}, \frac{-1 + q f}{q f}\right] - e^{\frac{3s\tau}{2}} \left( \frac{(-1 + q_0) \left(1 + (-1 + e^{-\frac{s\tau}{2}}) q f\right)}{(-1 + q f) q f} \right)^{2d}$ 

 $\left. \left. \text{Hypergeometric2F1}\left[2 - 2d, 3 - \frac{2d(1+s)}{s}, 4 - \frac{2d(1+s)}{s}, \frac{e^{\frac{s\tau}{2}} (-1 + q f)}{q f}\right] \right] \right)$ 

(*pcoal[k,t]/.x[t]→qtadditiveback[t]/.ne[t]→ntadditiveback[t]NeN
Integrate[%,{t,0,τ},Assumptions→{τ>0}](*)slow*)
(*cbackapprox[k,τ]==%/Simplify*)

```

The probability of no recombination in  $\tau$  generations



$$\text{In[282]:= rbackapprox}[k_, \tau_] := \frac{2 k r \left( -i \pi + \text{Log} \left[ e^{\frac{s \tau}{2}} (-1 + qf) - qf \right] \right)}{s}$$

```
(*preco[k,t]/.x[t]→qtadditiveback[t]
Integrate[%,{t,0,τ}]*)
(*rbackapprox[k,τ]==%/Simplify*)
```

The probability of no mutation in  $\tau$  generations

$$\text{In[283]:= mutbackapprox}[k_, \tau_] := - \frac{2 \left( -1 + e^{\frac{s \tau}{2}} \right) k (-1 + qf) u}{qf s}$$

```
(*pmut[k,t]/.x[t]→qtadditiveback[t]
Integrate[%,{t,0,τ},Assumptions→{τ>0}]*)
(*mutbackapprox[k,τ]==%/Simplify*)
```

The probability of no migration in  $\tau$  generations

$$\begin{aligned} \text{In[284]:= migbackapprox}[k_, \tau_] := & \frac{1}{2 N_0 N_{eN} (-1 + q_0)^2} k m (-1 + qf)^2 \left( \frac{1}{qf} \right)^{3-2d} \left( \frac{(-1 + q_0) qf}{q_0 (-1 + qf)} \right)^{\frac{2d}{s}} \\ & \left( \frac{-e^{\frac{s \tau}{2}} (-1 + qf) + qf}{qf} \right)^{-2d} \left( \frac{1}{d - s + d s} \left( \frac{-1 + q_0}{-1 + qf} \right)^{2d} qf \left( \frac{-e^{\frac{s \tau}{2}} (-1 + qf) + qf}{qf} \right)^{2d} \right. \\ & \quad \text{Hypergeometric2F1} \left[ 2 - 2d, 2 - \frac{2d(1+s)}{s}, 3 - \frac{2d(1+s)}{s}, \frac{-1 + qf}{qf} \right] - \frac{1}{d - s + d s} \\ & \quad e^{(-d+s)\tau} \left( \frac{1}{qf} \right)^{-1+2d} \left( \frac{(-1 + q_0) \left( 1 + \left( -1 + e^{-\frac{s \tau}{2}} \right) qf \right)}{-1 + qf} \right)^{2d} \text{Hypergeometric2F1} \left[ 2 - 2d, \right. \\ & \quad \left. 2 - \frac{2d(1+s)}{s}, 3 - \frac{2d(1+s)}{s}, \frac{e^{\frac{s \tau}{2}} (-1 + qf)}{qf} \right] - 1 / (-3s + 2d(1+s)) \quad 2 e^{-d\tau} (-1 + qf) \\ & \quad \left( e^{d\tau} \left( \frac{-1 + q_0}{-1 + qf} \right)^{2d} \left( \frac{-e^{\frac{s \tau}{2}} (-1 + qf) + qf}{qf} \right)^{2d} \text{Hypergeometric2F1} \left[ 2 - 2d, 3 - \frac{2d(1+s)}{s}, \right. \right. \\ & \quad \left. \left. 4 - \frac{2d(1+s)}{s}, \frac{-1 + qf}{qf} \right] - e^{\frac{3s\tau}{2}} \left( \frac{1}{qf} \right)^{2d} \left( \frac{(-1 + q_0) \left( 1 + \left( -1 + e^{-\frac{s \tau}{2}} \right) qf \right)}{-1 + qf} \right)^{2d} \right. \\ & \quad \left. \left. \text{Hypergeometric2F1} \left[ 2 - 2d, 3 - \frac{2d(1+s)}{s}, 4 - \frac{2d(1+s)}{s}, \frac{e^{\frac{s \tau}{2}} (-1 + qf)}{qf} \right] \right] \right) \end{aligned}$$

```
(*pmig[k,t]/.x[t]→qtadditiveback[t]/.n[t]→ntadditiveback[t]NeN
Integrate[%,{t,0,τ},Assumptions→{τ>0}]*slow*)
(*migbackapprox[k,τ]==%/Simplify*)
```

A population of constant size will have the same recombination and mutation terms (because these depend only on the allele frequency dynamics backwards in time, which are approximately the same,

just with different  $q_0$ ), but different coalescence and migration terms

$$\text{In[285]:= cclassicbackapprox}[k_, \tau_] := \frac{(-1+k) k \left( -\frac{2 \left( -1+e^{\frac{s \tau}{2}} \right) (-1+qf)}{s} + qf \tau \right)}{4 N_0 N_e N qf}$$

```
(*pcoal[k,t]/.x[t]→qtadditiveback[t]/.ne[t]→N0 NeN
Integrate[%,{t,0,τ},Assumptions→{τ>0}]
cclassicbackapprox[k,τ]==%/Simplify*)
```

$$\text{In[286]:= migclassicbackapprox}[k_, \tau_] := \frac{k m \left( -\frac{2 \left( -1+e^{\frac{s \tau}{2}} \right) (-1+qf)}{s} + qf \tau \right)}{2 N_0 N_e N qf}$$

```
(*pmig[k,t]/.x[t]→qtadditiveback[t]/.n[t]→N0 NeN
Integrate[%,{t,0,τ},Assumptions→{τ>0}]
migclassicbackapprox[k,τ]==%/Simplify*)
```

It is also useful to calculate the probability of coalescence by time  $T$  during the population bottleneck (ignoring the sweep)

$$\begin{aligned} \text{In[287]:= cbottlebackapprox}[k_, \tau_] := & -\frac{1}{4 N_0 N_e N (-1+q_0)^2 (d-s+d s)} \\ & (-1+k) k (-1+qf)^2 \left( \frac{1}{qf} \right)^{2-2d} \left( \frac{(-1+q_0) qf}{q_0 (-1+qf)} \right)^{\frac{2d}{s}} \left( \frac{-e^{\frac{s \tau}{2}} (-1+qf) + qf}{qf} \right)^{-2d} \\ & \left( -\left( \frac{-1+q_0}{-1+qf} \right)^{2d} \left( \frac{-e^{\frac{s \tau}{2}} (-1+qf) + qf}{qf} \right)^{2d} \text{Hypergeometric2F1} \left[ 2-2d, 2-\frac{2d(1+s)}{s}, \right. \right. \\ & \left. \left. 3-\frac{2d(1+s)}{s}, \frac{-1+qf}{qf} \right] + e^{(-d+s)\tau} \left( \frac{1}{qf} \right)^{2d} \left( \frac{(-1+q_0) \left( 1 + \left( -1+e^{-\frac{s \tau}{2}} \right) qf \right)}{-1+qf} \right)^{2d} \right. \\ & \left. \left. \text{Hypergeometric2F1} \left[ 2-2d, 2-\frac{2d(1+s)}{s}, 3-\frac{2d(1+s)}{s}, \frac{e^{\frac{s \tau}{2}} (-1+qf)}{qf} \right] \right) \end{aligned}$$

```
(*pcoal[k,t]/.x[t]→1/.ne[t]→ntadditiveback[t]NeN
Integrate[%,{t,0,τ},Assumptions→{τ>0}](*) (*slow*)
(*cbottlebackapprox[k,τ]==%/Simplify*)
pcoalbottle[k_, T_] := 1 - Exp[-cbottlebackapprox[k, T]]
```

# Results

## Rescue from standing genetic variance

### The probability of rescue and soft selective sweeps

We begin with  $\kappa$  copies of the beneficial allele. With  $\kappa \ll N(0)$  we can consider  $\kappa$  to be the initial number of heterozygotes and assume these heterozygotes do not interact. With  $\rho$  the probability any one copy establishes, the number of mutants that establish,  $X$ , is then binomially distributed with parameters  $\kappa$  and  $\rho$

```
In[293]:= Simplify[PDF[BinomialDistribution[κ, ρ], X], {0 ≤ X ≤ κ}]
```

```
Out[293]= (1 - ρ)-X+κ ρX Binomial[κ, X]
```

Check this with simulations [100 rescued replicates for each histogram, theory as lines, rescue (d=0.05) in blue and constant population size (d=0) in red]

```
In[294]:= plotNestablishSGV[sval_, dval_, Nval_, hval_, xmax_, ks_] :=
(
  params = {s → sval, d → dval, N0 → Nval, h → hval, B → 2};

  pdfgen = Simplify[PDF[BinomialDistribution[k, p], X], {0 ≤ X ≤ k}];

  plot = Table[
    (*rescue theory*)
    pdf = pdfgen /. p → pest /. v → w (3 + 4 B - 4 w) / 4 /. w → (1 + s h) (1 - d) /. params;
    theory = Table[{X, pdf}, {X, 0, xmax}];
    theoryplot = ListPlot[theory, Joined → True, PlotStyle → Thick];

    (*constant theory*)
    pdf =
      pdfgen /. p → pest /. v → w (3 + 4 B - 4 w) / 4 /. w → (1 + s h) (1 - d) /. d → 0 /. params;
    theory = Table[{X, pdf}, {X, 0, xmax}];
    theoryplotConstant = ListPlot[theory,
      Joined → True, PlotStyle → Directive[defaultcolors[[2]], Thick]];

    (*rescue simulations*)
    folder = StringForm["nestablish_SGV_d`_s`",
      NumberForm[dval, {3, 2}], NumberForm[sval, {3, 2}]];
    data = Import[datadir <> ToString[folder] <> "/data/nestablish_k" <>
      ToString[k] <> ".txt", "Table"] // Flatten;
    dataplot = Histogram[data, {-0.5, xmax - 0.5, 1}, "Probability"];

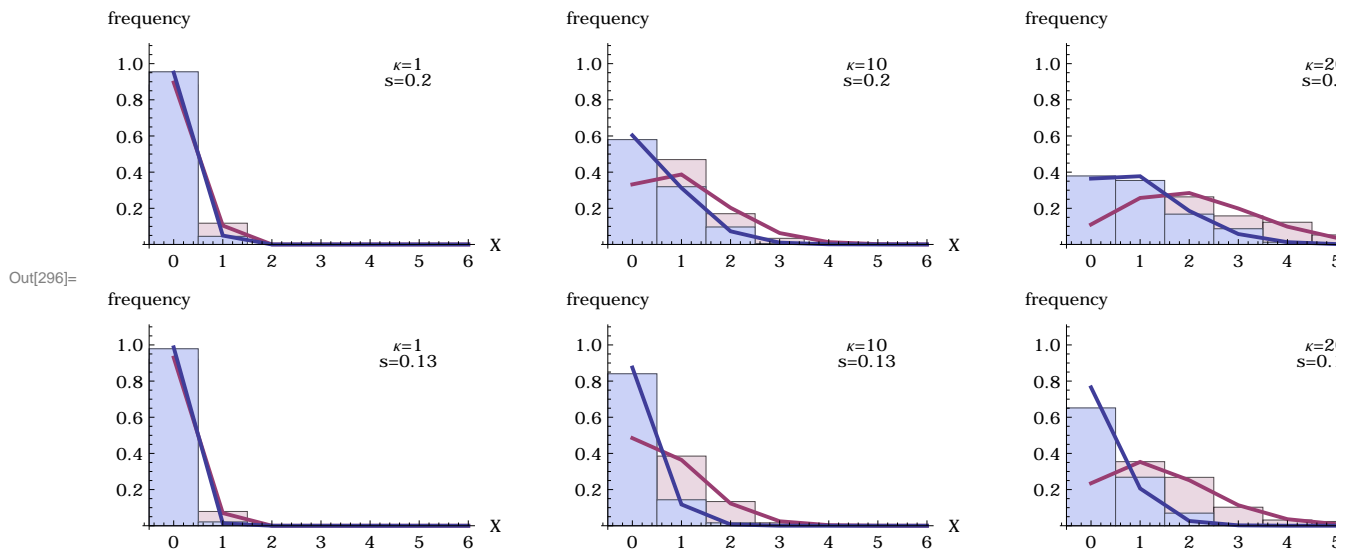
    (*constant simulations*)
    folder = StringForm["nestablish_SGV_d`_s`",
      NumberForm[0, {3, 2}], NumberForm[sval, {3, 2}]];
    data = Import[datadir <> ToString[folder] <> "/data/nestablish_k" <>
      ToString[k] <> ".txt", "Table"] // Flatten;
    dataplotConstant = Histogram[data, {-0.5, xmax - 0.5, 1}, "Probability",
      ChartStyle → Directive[defaultcolors[[2]], Opacity[0.2]]];

    (*plot*)
```

```
Show[
  dataplotConstant, dataplot, theoryplotConstant, theoryplot,
  PlotRange -> {0, 1},
  AxesLabel -> {"X", "frequency"},
  Epilog -> {
    Text[StringForm["κ=``", k], Scaled@{0.8, 0.9}],
    Text[StringForm["s=``", sval], Scaled@{0.8, 0.825}]
  }
],

{k, ks}
]

ss = {0.2, 0.13};
GraphicsGrid[
  Table[
    plotNestablishSGV[s, 0.05, 104, 0.5, 6, {1, 10, 20}],
    {s, ss}
  ],
  ImageSize -> 700
]
```



The probability of rescue is just the probability at least one establishes

```
In[298]:= PSGV = (1 - Simplify[PDF[BinomialDistribution[κ, ρ], X], {0 ≤ X ≤ κ}] /. X -> 0)
```

```
Out[298]:= 1 - (1 - ρ)κ
```

Let's check this with simulations (100 rescues for each dot, theory as lines, data as dots, rescue as solids, sweeps as open/dashed)

```
In[299]:= predictPrescue[sval_, dval_, hval_, ks_] :=
(
  params = {s -> sval, d -> dval, h -> hval, B -> 2};
  psgv = PSGV /. ρ -> pest /. v -> w (3 + 4 B - 4 w) / 4 /. w -> (1 + s h) (1 - d) /. params;
  Table[{κ, psgv}, {κ, ks[[1]], ks[[-1]]}]
)
```

```

)

getPrescue[dval_, sval_, ks_] :=
(
  folder = StringForm["nestablish_SGV_d`_s`",
    NumberForm[dval, {3, 2}], NumberForm[sval, {3, 2}]];
  Table[data = Import[datadir <> ToString[folder] <> "/data/nestablish_k" <>
    ToString[k] <> ".txt", "Table"] // Flatten;
    {k, 100 / Length[data] // N},
    {k, ks}
  ]
)

plotPrescuesSGV[svals_, dval_, Nval_, hval_, ks_] :=
(
  (*rescue theory*)
  theory = predictPrescue[svals[[1]], dval, hval, ks];
  (*rescue theory*)
  theory2 = predictPrescue[svals[[2]], dval, hval, ks];
  (*constant theory*)
  theorySweep = predictPrescue[svals[[1]], 0, hval, ks];
  (*constant theory*)
  theorySweep2 = predictPrescue[svals[[2]], 0, hval, ks];

  (*rescue simulations*)
  Prescue = getPrescue[dval, svals[[1]], ks];
  (*rescue simulations*)
  Prescue2 = getPrescue[dval, svals[[2]], ks];
  (*constant simulations*)
  Psweep = getPrescue[0, svals[[1]], ks];
  (*constant simulations*)
  Psweep2 = getPrescue[0, svals[[2]], ks];

  Show[
    ListPlot[Prescue, PlotStyle → AbsolutePointSize[5], PlotRange → {0, 1}],
    ListPlot[Psweep,
      PlotStyle → Directive[AbsolutePointSize[5], defaultcolors[[2]]],
      ListPlot[{theory, theorySweep}, Joined → True, PlotStyle → Thick],
      ListPlot[Prescue2,
        PlotMarkers → Graphics[{Thickness[0.4], Circle[]}, ImageSize → 6]],
      ListPlot[Psweep2, PlotMarkers → Graphics[
        {defaultcolors[[2]], Thickness[0.4], Circle[]}, ImageSize → 6]],
      ListPlot[{theory2, theorySweep2}, Joined → True,
        PlotStyle → Directive[Thick, Dashing[Large]]],
      PlotRange → {0, 1},
      Frame → {True, True, False, False},
      FrameLabel → {"Initial number of beneficial alleles"},
      LabelStyle → labelstyle,
      ImagePadding → {{50, 15}, {40, 10}},
      Epilog → {
        Rotate[Text[
          Style["Probability of rescue", labelstyle], Scaled@{-0.125, 0.5}],  $\pi/2$ ]
        },
      PlotRangeClipping → False,

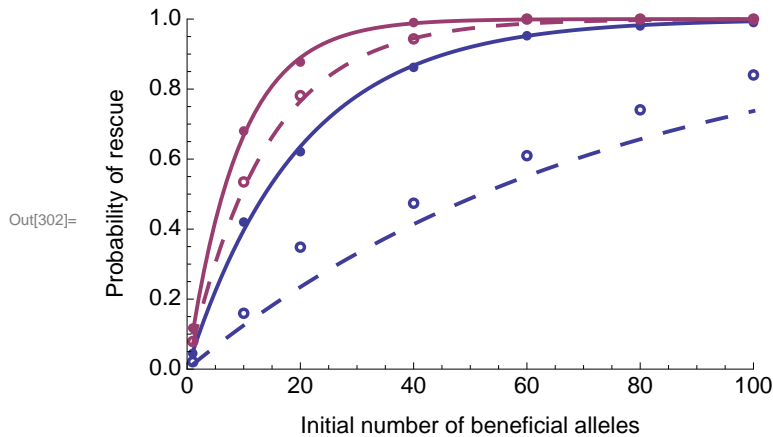
```

```

PlotRangePadding → None
]
)

plotPrescueSGV[{0.2, 0.13}, 0.05, 104, 0.5, {1, 10, 20, 40, 60, 80, 100}]

```



The expected number of copies that establish is

```
In[303]:= Expectation[X, X ∈ BinomialDistribution[κ, ρ]]
```

Out[303]=  $\kappa \rho$

so the expected number that establish given rescue is

```
In[304]:= nSGV =  $\frac{1}{\text{PSGV}}$  Expectation[X, X ∈ BinomialDistribution[κ, ρ]]
```

Out[304]=  $\frac{\kappa \rho}{1 - (1 - \rho)^\kappa}$

Check with simulations

```
In[305]:= predictNest[sval_, dval_, hval_, ks_] :=
(
  params = {s → sval, d → dval, h → hval, B → 2};
  nsgv = nSGV /. ρ → pest /. v → w (3 + 4 B - 4 w) / 4 /. w → (1 + s h) (1 - d) /. params;
  Table[{κ, nsgv}, {κ, ks[[1]], ks[[-1]]}]
)

```

```

getNest[dval_, sval_, ks_] :=
(
  folder = StringForm["nestablish_SGV_d`_s`"],
  NumberForm[dval, {3, 2}], NumberForm[sval, {3, 2}];
  Table[
    data = Import[datadir <> ToString[folder] <>
      "/data/nestablish_k" <> ToString[k] <> ".txt", "Table"] // Flatten;
    data = Select[data, # > 0 &];
    {k, Mean[data], StandardDeviation[data] / Sqrt[Length[data]}],
    {k, ks[[1]], ks[[-1]]}
  ]
)

```

```

    {k, ks}
  ]
)

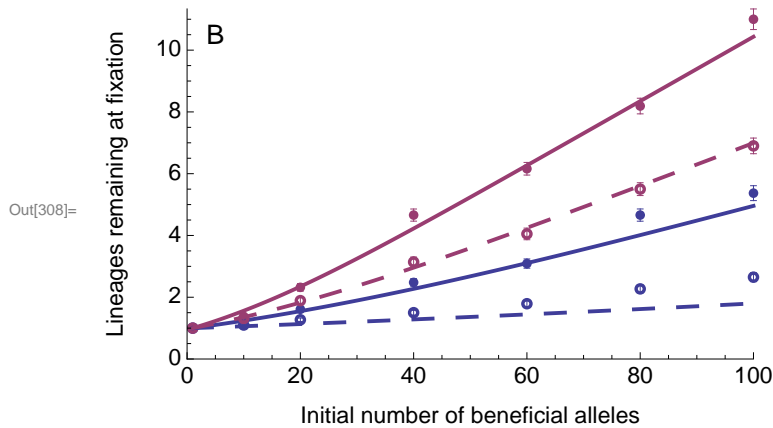
plotNestablishSGVrescue[svals_, dval_, Nval_, hval_, ks_] :=
(
  (*rescue theory*)
  theory = predictNest[svals[[1]], dval, hval, ks];
  (*rescue theory*)
  theory2 = predictNest[svals[[2]], dval, hval, ks];
  (*constant theory*)
  theorySweep = predictNest[svals[[1]], 0, hval, ks];
  (*constant theory*)
  theorySweep2 = predictNest[svals[[2]], 0, hval, ks];

  (*rescue simulations*)
  Prescue = getNest[dval, svals[[1]], ks];
  (*rescue simulations*)
  Prescue2 = getNest[dval, svals[[2]], ks];
  (*constant simulations*)
  Psweep = getNest[0, svals[[1]], ks];
  (*constant simulations*)
  Psweep2 = getNest[0, svals[[2]], ks];

  Show[
    ErrorListPlot[Prescue, PlotStyle → AbsolutePointSize[5], PlotRange → {0, 12}],
    ErrorListPlot[Psweep,
      PlotStyle → Directive[AbsolutePointSize[5], defaultcolors[[2]]]],
    ListPlot[{theory, theorySweep}, Joined → True, PlotStyle → Thick],
    ErrorListPlot[Prescue2,
      PlotMarkers → Graphics[{Thickness[0.4], Circle[]}, ImageSize → 6]],
    ErrorListPlot[Psweep2, PlotStyle → defaultcolors[[2]], PlotMarkers →
      Graphics[{defaultcolors[[2]], Thickness[0.4], Circle[]}, ImageSize → 6]],
    ListPlot[{theory2, theorySweep2}, Joined → True,
      PlotStyle → Directive[Thick, Dashing[Large]]],
    Frame → {True, True, False, False},
    FrameLabel → {"Initial number of beneficial alleles"},
    LabelStyle → labelstyle,
    ImagePadding → {{50, 15}, {40, 10}},
    PlotRange → All,
    Epilog → {
      Text[Style["B", letterstyle], Scaled@letterposition],
      Rotate[Text[Style["Lineages remaining at fixation", labelstyle],
        Scaled@{-0.125, 0.5}],  $\pi/2$ ],
    },
    PlotRangeClipping → False,
    PlotRangePadding → None
  ]
)

plotNestablishSGVrescue[{0.2, 0.13}, 0.05, 104, 0.5, {1, 10, 20, 40, 60, 80, 100}]

```



We see that we underestimate the number that establish for weaker selection coefficients, when the beneficial allele can drift for long enough that the wildtype is rare and thus homozygotes are made, increasing establishment probabilities.

The probability that more than one copy of the beneficial allele establishes is the probability of rescue minus the probability that only one establishes

```
In[310]:= Simplify[PDF[BinomialDistribution[κ, ρ], X], {0 ≤ X ≤ κ}];
PSGV - (% /. X → 1) // Simplify
```

```
Out[311]= 1 - (1 - ρ)κ - κ (1 - ρ)-1+κ ρ
```

Conditioning on rescue, the probability that more than one copy establishes is then

```
In[312]:= Simplify[PDF[BinomialDistribution[κ, ρ], X], {0 ≤ X ≤ κ}];
```

```
psoftSGV =  $\frac{\text{PSGV} - (\% /. X \rightarrow 1)}{\text{PSGV}}$  // Simplify
```

```
psoftSGV == 1 - nSGV  $\frac{1 - \text{PSGV}}{1 - \rho}$  // Simplify
```

```
Out[313]= -  $\frac{-1 + (1 - \rho)^\kappa + \kappa (1 - \rho)^{-1+\kappa} \rho}{1 - (1 - \rho)^\kappa}$ 
```

```
Out[314]= True
```

Check with simulations

```
In[315]:= predictPsoft[sval_, dval_, hval_, ks_] :=
(
  params = {s → sval, d → dval, h → hval, B → 2};
  p = psoftSGV /. ρ → pest /. v → w (3 + 4 B - 4 w) / 4 /. w → (1 + s h) (1 - d) /. params;
  Table[{κ, p}, {κ, ks[[1]], ks[[-1]]}]
)
```

```
getPsoft[dval_, sval_, ks_] :=
```

```
(
  folder = StringForm["nestablish_SGV_d`_s`"],
  NumberForm[dval, {3, 2}], NumberForm[sval, {3, 2}];
  Table[data = Import[datadir <> ToString[folder] <> "/data/nestablish_k" <>
    ToString[k] <> ".txt", "Table"] // Flatten;
    data = Length[Select[data, # > 1 &]];
  )
```



```

    {k, data / 100},
    {k, ks}
  ]
)

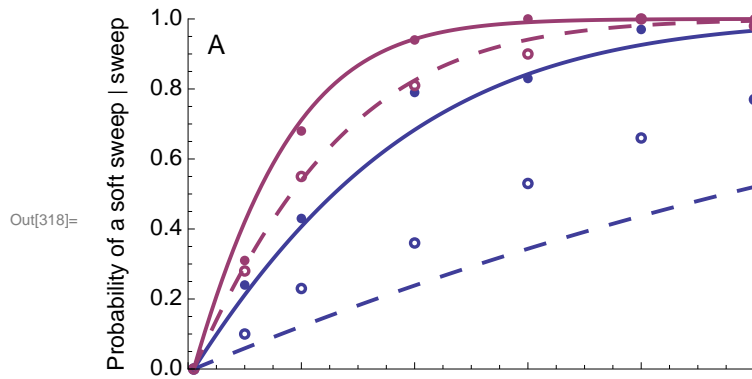
plotPsoftSGV[svals_, dval_, Nval_, hval_, ks_] :=
(
  (*rescue theory*)
  theory = predictPsoft[svals[[1]], dval, hval, ks];
  (*rescue theory*)
  theory2 = predictPsoft[svals[[2]], dval, hval, ks];
  (*constant theory*)
  theorySweep = predictPsoft[svals[[1]], 0, hval, ks];
  (*constant theory*)
  theorySweep2 = predictPsoft[svals[[2]], 0, hval, ks];

  (*rescue simulations*)
  Prescue = getPsoft[dval, svals[[1]], ks];
  (*rescue simulations*)
  Prescue2 = getPsoft[dval, svals[[2]], ks];
  (*constant simulations*)
  Psweep = getPsoft[0, svals[[1]], ks];
  (*constant simulations*)
  Psweep2 = getPsoft[0, svals[[2]], ks];

  Show[
    ListPlot[Prescue, PlotStyle → AbsolutePointSize[5], PlotRange → {0, 12}],
    ListPlot[Psweep,
      PlotStyle → Directive[AbsolutePointSize[5], defaultcolors[[2]]],
      ListPlot[{theory, theorySweep}, Joined → True, PlotStyle → Thick],
      ListPlot[Prescue2,
        PlotMarkers → Graphics[{Thickness[0.4], Circle[]}, ImageSize → 6]],
        ListPlot[Psweep2, PlotMarkers → Graphics[
          {defaultcolors[[2]], Thickness[0.4], Circle[]}, ImageSize → 6]],
          ListPlot[{theory2, theorySweep2}, Joined → True,
            PlotStyle → Directive[Thick, Dashing[Large]]],
            Frame → {True, True, False, False},
            FrameTicksStyle → {FontColor → White, Automatic, Automatic, Automatic},
            LabelStyle → labelstyle,
            ImagePadding → {{50, 15}, {40, 10}},
            PlotRange → All,
            Epilog → {
              Text[Style["A", letterstyle], Scaled@letterposition],
              Rotate[Text[Style["Probability of a soft sweep | sweep", labelstyle],
                Scaled@{-0.125, 0.5}],  $\pi/2$ ]
            },
            PlotRangeClipping → False,
            PlotRangePadding → None
          ]
        ]
      )

plotPsoftSGV[{0.2, 0.13}, 0.05, 104, 0.5, {1, 10, 20, 40, 60, 80, 100}]
Export[imagedir <> "PsoftSGV.pdf", %];

```



and we see that we do a good job unless  $s$  and  $\kappa$  are small and  $d > 0$ , which is where we underestimate the probability of establishment (due to deviations from a true branching process). If use the empirical measure of establishment probability (probability of rescue from the from the  $\kappa=1$  simulations) we do much better

```

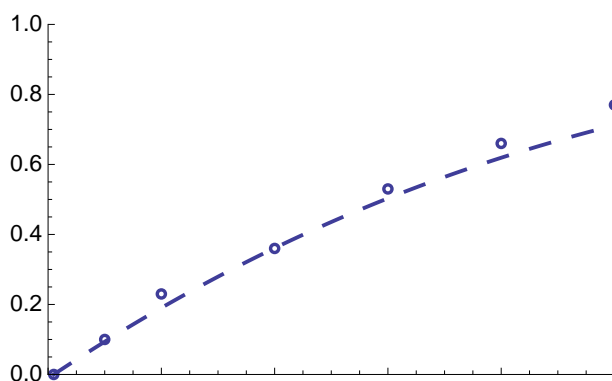
In[320]:= predictPsoftEmp[sval_, dval_, hval_, ks_] :=
(
  folder = StringForm["nestablish_SGV_d`_s`",
    NumberForm[dval, {3, 2}], NumberForm[sval, {3, 2}]];
  data = Import[datadir <> ToString[folder] <> "/data/nestablish_k1.txt", "Table"] //
    Flatten;
  Pest = 100 / Length[data] // N;
  psoft = 1 - nSGV  $\frac{1 - \text{PSGV}}{1 - \rho}$  /.  $\rho \rightarrow \text{Pest}$ ;
  Table[{ $\kappa$ , psoft}, { $\kappa$ , ks[[1]], ks[[-1]]}]
)

(*rescue theory*)
theory = predictPsoftEmp[0.13, 0.05, 0.5, {1, 10, 20, 40, 60, 80, 100}];
(*rescue simulations*)
Prescue = getPsoft[0.05, 0.13, {1, 10, 20, 40, 60, 80, 100}];

Show[
  ListPlot[theory, Joined -> True, PlotStyle -> Directive[Dashing[Large], Thick]],
  ListPlot[Prescue,
    PlotMarkers -> Graphics[{Thickness[0.4], Circle[]}, ImageSize -> 6]],
  Frame -> {True, True, False, False},
  LabelStyle -> labelstyle,
  ImagePadding -> {{50, 15}, {40, 10}},
  FrameTicksStyle -> {FontColor -> White, Automatic, Automatic, Automatic},
  PlotRangeClipping -> False,
  PlotRangePadding -> None,
  PlotRange -> {0, 1}
]

```

Out[323]=



## Effective initial allele frequency and the backward-time dynamics

Dividing the actual starting allele frequency by the probability of rescue, conditioned on rescue it is as if the beneficial allele started at frequency

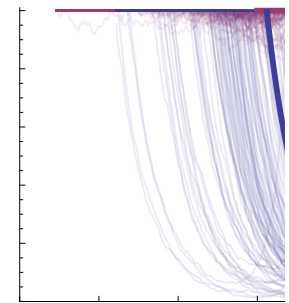
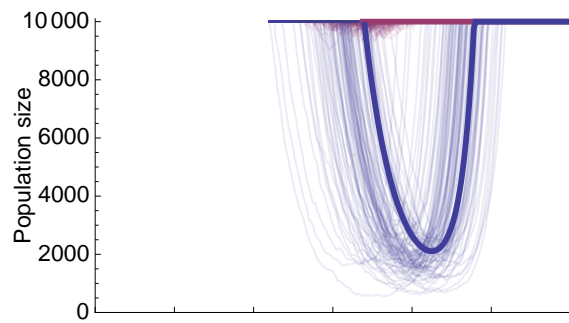
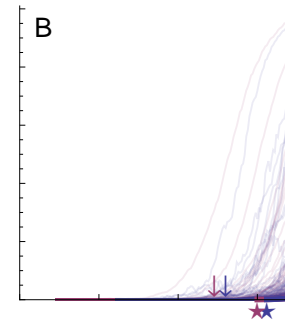
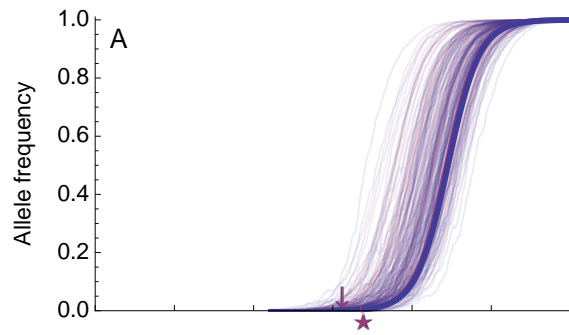
```
In[324]:= q0rescueSGV =  $\frac{\kappa}{2 N_0} \frac{1}{P_{SGV}};$ 
```

Since the probability of rescue is just the probability of sweep, this is the same in a constant population

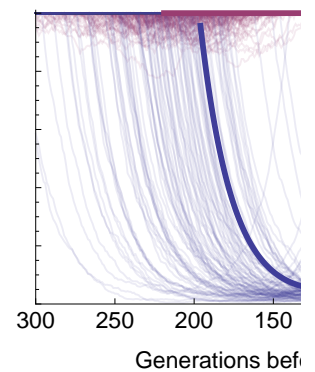
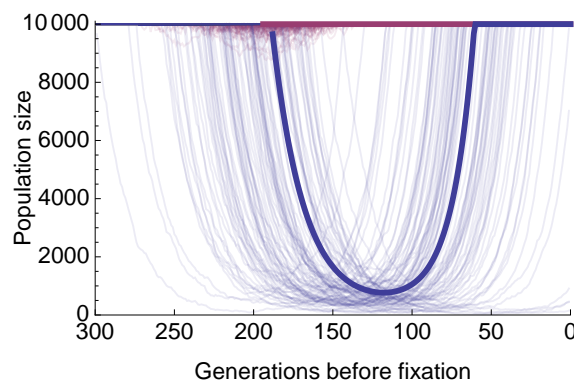
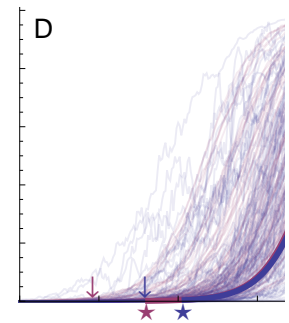
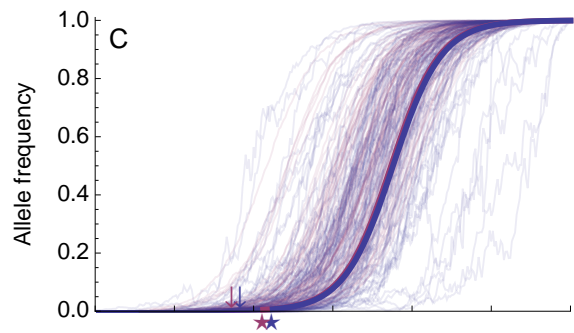
```
In[325]:= q0rescueSGV;
```

Compare the predictions to simulations

```
In[326]:= letters = {"A", "B"}, {"C", "D"};
ks = {100, 10};
ss = {0.2, 0.13};
GraphicsGrid[
  Table[
    Table[
      plotDynamicsBackwards[104, 0.05, ss[[i]],
        0.5, ks[[j]], 0, 0, 300, letters[[i, j]], 1, 100, Mean],
      {j, Length[ks]}],
    {i, Length[ss]}],
  Spacings -> {0, 0}
]
```



Out[329]=



## Compare with a true branching process

Next imagine a clonally-reproducing population, which is a true branching process (when a population is

below carrying capacity). The variance in reproductive success is

```
In[330]:= Expectation[X B, {X ∈ BernoulliDistribution[V]}] /. V → w / B;
Expectation[(X B)^2, {X ∈ BernoulliDistribution[V]}] - %^2 /. V → w / B // Simplify

Out[331]= (B - w) w
```

which is lower than in the sexual model because we no longer have stochasticity in Mendelian transmission.

To compare with the sexual model, we want to keep the probability of establishment and initial frequency the same (so conditioning has the same strength). The probability of fixation is higher in the clonal model, all else being equal, because the variance is lower due to a lack of Mendelian transmission. However, starting with only half as many copies (so that the initial genotype frequency in the clonal model is the same as the initial allele frequency in the sexual model, since there are twice as many alleles as genotypes) makes the probability of fixation roughly the same for both rescue and constant populations. For example,

```
In[332]:= PSGV /. ρ → pest /. v → w (3 + 4 B - 4 w) / 4 /. w → (1 + s h) (1 - d) /. d → 0 /. h → 0.5 /.
s → 0.13 /. B → 2 /. κ → 10
PSGV /. ρ → pest /. v → (B - w) w /. w → (1 + s h) (1 - d) /. d → 0 /. h → 0.5 /. s → 0.13 /.
B → 2 /. κ → 5

Out[332]= 0.515397

Out[333]= 0.479392
```

Note that the genotype frequency dynamics in the clonal model are the same as the allele frequency dynamics in the diploid additive model under weak selection (when the advantage in the haploid model is  $s h = s/2$ )

```
In[334]:= q[t + 1] == 
$$\frac{q[t] (1 + s h) (1 - d)}{q[t] (1 + s h) (1 - d) + (1 - q[t]) (1 - d)}$$
;
Series[
$$\frac{q[t] (1 + s h) (1 - d)}{q[t] (1 + s h) (1 - d) + (1 - q[t]) (1 - d)} - q[t], \{s, 0, 1\}] // Normal;
DSolve[{D[q[t], t] == %, q[0] == q0}, q[t], t] // Simplify // Flatten;
qtadditive == % /. h → 1 / 2$$

```

Solve::ifun : Inverse functions are being used by Solve, so  
some solutions may not be found; use Reduce for complete solution information. >>

```
Out[337]= True
```

The population size dynamics in the clonal model are slightly different, as the lack of homozygote mutants make for a slower recovery

```
In[338]:= ntadditiveClonal =
DSolve[{D[n[t], t] == n[t] (q[t] (1 + s h) (1 - d) + (1 - q[t]) (1 - d)) - n[t] /.
qtadditive /. h → 1 / 2, n[0] == N0}, n[t], t] // Simplify

Out[338]= 
$$\left\{ \left\{ n[t] \rightarrow e^{-d t} N_0 \left( 1 + \left( -1 + e^{\frac{s t}{2}} \right) q_0 \right)^{1-d} \right\} \right\}$$

```

Backwards in time this is

```
In[339]:= n[t] /. ntadditiveClonal /. t -> tfixadditive -  $\tau$  // Simplify;

ntadditivebackClonal[ $\tau$ _] :=  $e^{d \tau} N_0 \left( \frac{(1 - q_0) q_f}{q_0 (1 - q_f)} \right)^{-\frac{2d}{s}} \left( (1 - q_0) \left( 1 + \frac{q_f}{1 - q_f} e^{-\frac{s \tau}{2}} \right) \right)^{(1-d)}$ 

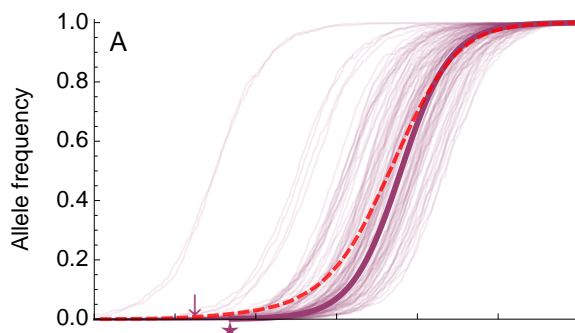
FullSimplify[ntadditivebackClonal[ $\tau$ ] == %%, {d < 1, 0 < q0 < 1}]

Out[341]= True
```

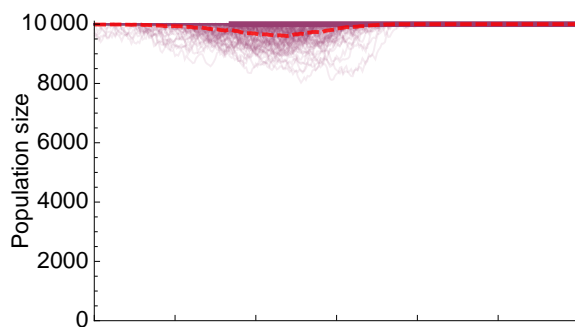
But let's just look at the constant N case (so that the N dynamics are the same too).

We then see that we get the beginning of the dynamics correct (as we do for the sexual model), but tend to underestimate the mean allele frequency later in the process (again, as we do for the sexual model)

```
In[342]:= letters = {"A", "B"}, {"C", "D"};
ks = {5};
ss = {0.13};
GraphicsGrid[
  Table[
    Table[
      plotDynamicsBackwardsClonal[104, 0, ss[[i]],
        0.5, ks[[j]], 0, 0, 300, letters[[i, j]], 0, 100, Mean],
      {j, Length[ks]}],
    {i, Length[ss]}],
  Spacings -> {0, 0}
]
```



Out[345]=



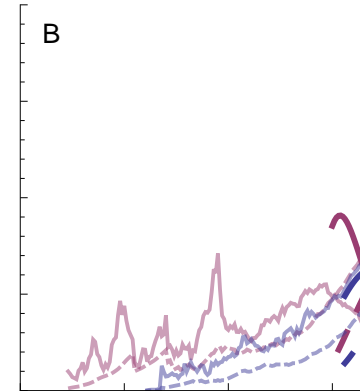
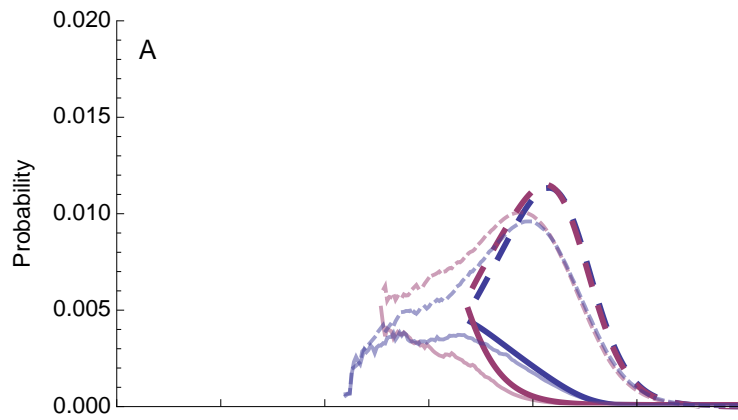
## The coalescent

```

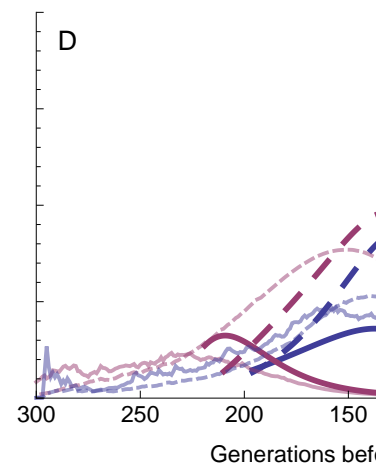
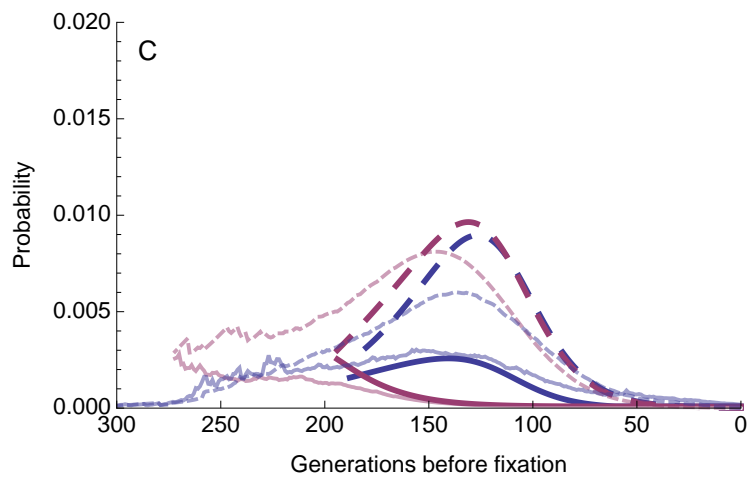
In[346]:= letters = {"A", "B"}, {"C", "D"};
ks = {100, 10};
ss = {0.2, 0.13};
GraphicsGrid[
  Table[
    Table[
      plotCoalescentSims[104, 0.05, ss[[i]], 0.5,
        ks[[j]], 0, 0, 2, 0.01, 300, 0.02, letters[[i, j]], 1, 100],
      {j, 2}],
    {i, 2}],
  Spacings -> {0, 0}
]

```





Out[349]=



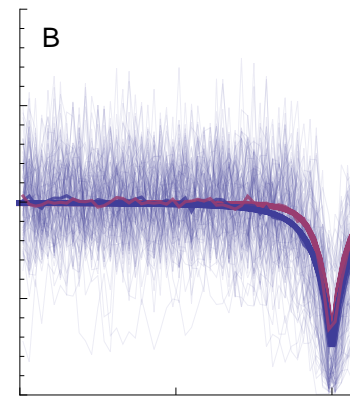
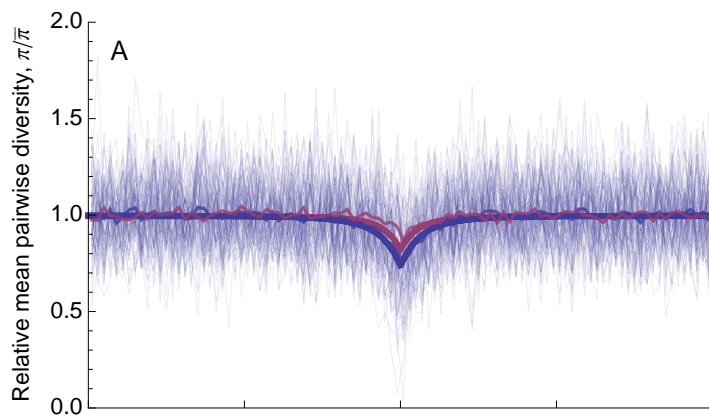
## Pairwise diversity

Relative pairwise diversity

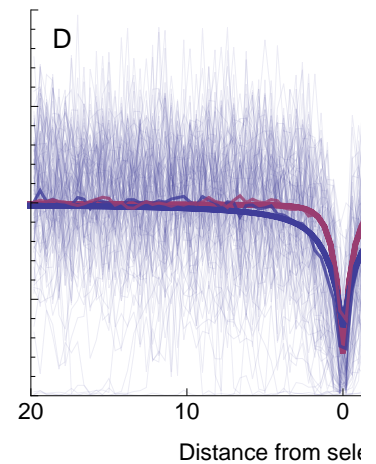
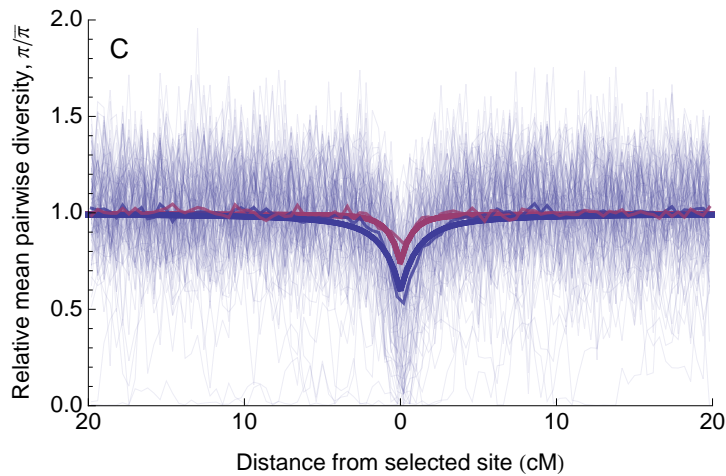
```

In[350]:= letters = {"A", "B"}, {"C", "D"};
ks = {100, 10};
ss = {0.2, 0.13};
GraphicsGrid[
  Table[
    Table[
      plotDiversityRelative[10^4, 0.05,
        ss[[i]], 0.5, ks[[j]], 0, 0, 100, letters[[i, j]], 1, 101],
      {j, 2}],
    {i, 2}],
  Spacings -> {0, 0}
]

```



Out[353]=

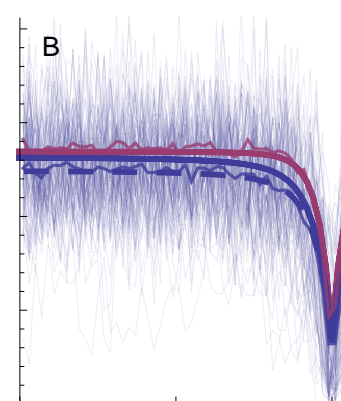
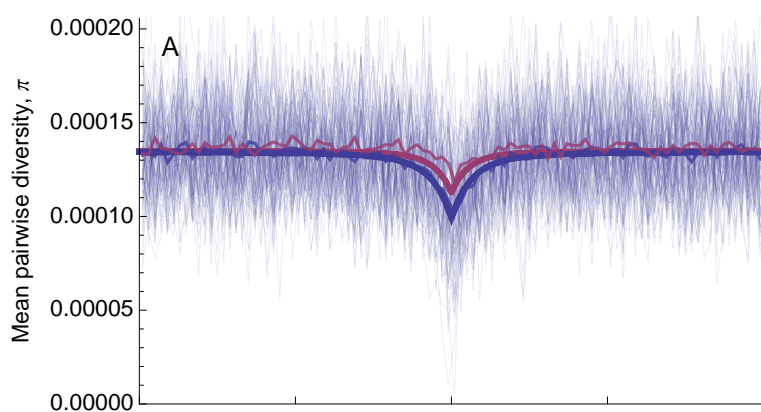


Absolute pairwise diversity

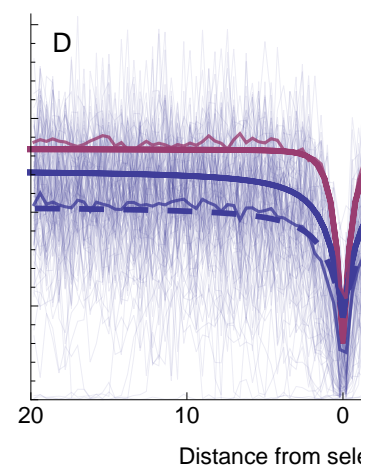
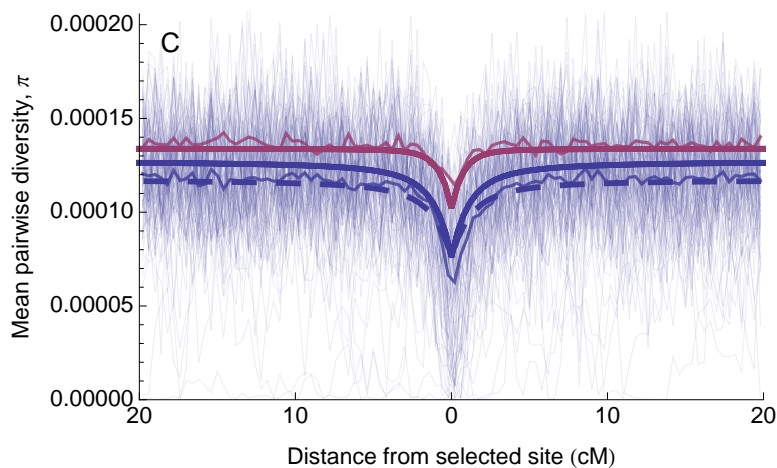
```

In[354]:= letters = {"A", "B"}, {"C", "D"};
ks = {100, 10};
ss = {0.2, 0.13};
GraphicsGrid[
  Table[
    Table[
      plotDiversity[10^4, 0.05, ss[[i]],
        0.5, ks[[j]], 0, 0, 100, letters[[i, j]], 1, 101],
      {j, 2}],
    {i, 2}],
  Spacings -> {0, 0}
]

```

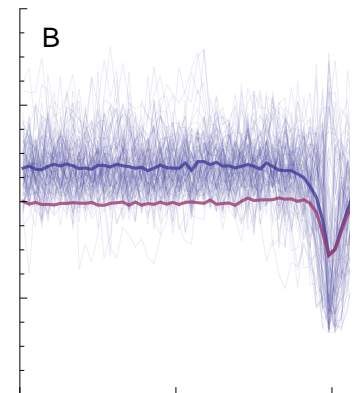
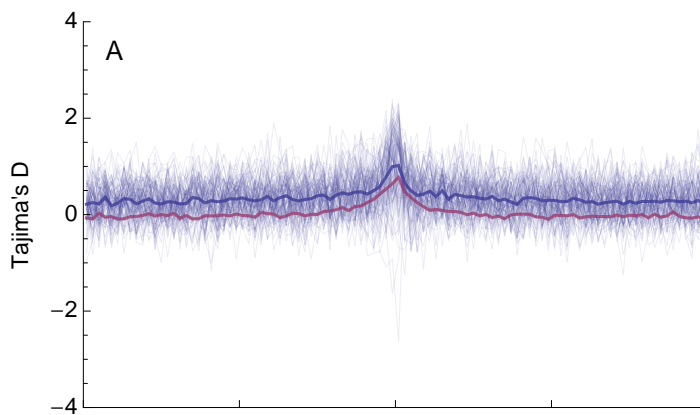


Out[357]=

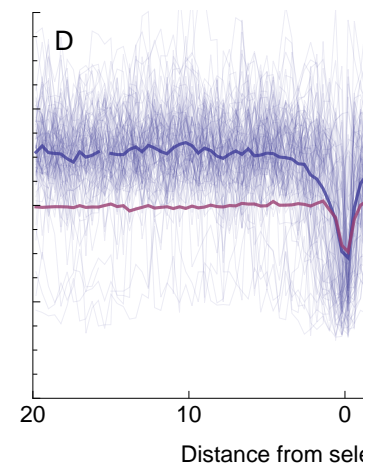
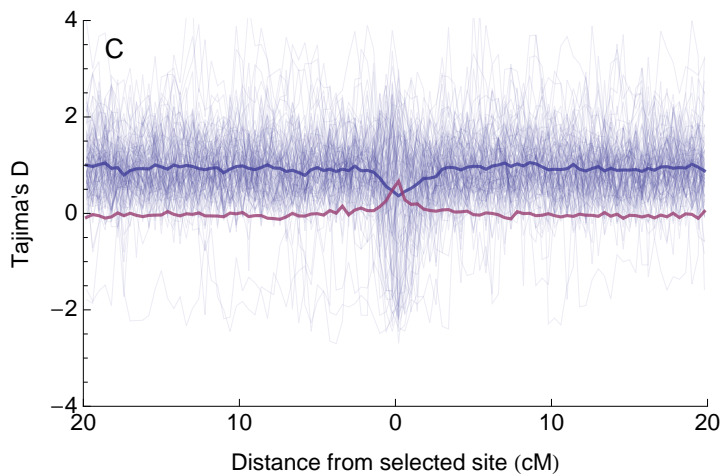


## Tajima's D

```
In[358]:= letters = {"A", "B"}, {"C", "D"};
ks = {100, 10};
ss = {0.2, 0.13};
GraphicsGrid[
  Table[
    Table[
      plotTajimasD[104, 0.05, ss[[i]], 0.5, ks[[j]], 0, 0, 100, letters[[i, j]], 1],
      {j, 2}],
    {i, 2}],
  Spacings -> {0, 0}
]
```



Out[361]=



## Sweep + Bottleneck (using expected bottleneck size)

We'd like to also compare the case where there is a sweep and a bottleneck, but the bottleneck is independent of the sweep. There are many ways to do it, but we'd like to concoct the bottleneck scenario such that it has the same effective population size during the sweep. Note that this does not

actually make the background diversity the same as the sweeps are of different lengths (since  $N$  affects  $q_0$  and  $q_f$ ).

Generically, the expected population size  $\tau$  generations before fixation is

```
In[362]:= Min[N0, ntadditiveback[τ]]
```

$$\text{Out[362]= } \text{Min}\left[N_0, e^{d\tau} N_0 \left(\frac{(1-q_0) q_f}{q_0 (1-q_f)}\right)^{-\frac{2d}{s}} \left((1-q_0) \left(1 + \frac{e^{-\frac{s\tau}{2}} q_f}{1-q_f}\right)\right)^{2(1-d)}\right]$$

The harmonic mean population size over  $t_{\text{fix}}$  generations is then

```
In[363]:= harmonicN[tfix_] := tfix / Sum[ $\frac{1}{\text{Min}[N_0, \text{ntadditiveback}[\tau]]}$ , {τ, 0, tfix}]
(*note this should actually have started at τ=1,
but the simulations have already been run and the error
is small when tfix is reasonably large, which it is*)
```

And we can calculate this numerically

```

In[364]:= params = {N0 -> 10^4, d -> 0.05, s -> 0.2, κ -> 100, h -> 0.5, B -> 2};
q0 = q0rescueSGV /. ρ -> pest /. v -> w (3 + 4 B - 4 w) / 4 /. w -> (1 + s h) (1 - d);
qf = 1 -  $\frac{1}{2 N_0 \rho}$  /. ρ -> pest /. v -> w  $\frac{(3 + 4 B - 4 W)}{4 Wbar^2}$  /. w -> ε + 1 /. ε -> 1 -  $\frac{W}{Wbar}$  /.
      Wbar -> (1 - d) (1 + s) /. W -> (1 - d) (1 + s h);
harmonicN[tfixadditive /. params] /. params

params = {N0 -> 10^4, d -> 0.05, s -> 0.2, κ -> 10, h -> 0.5, B -> 2};
q0 = q0rescueSGV /. ρ -> pest /. v -> w (3 + 4 B - 4 w) / 4 /. w -> (1 + s h) (1 - d);
qf = 1 -  $\frac{1}{2 N_0 \rho}$  /. ρ -> pest /. v -> w  $\frac{(3 + 4 B - 4 W)}{4 Wbar^2}$  /. w -> ε + 1 /. ε -> 1 -  $\frac{W}{Wbar}$  /.
      Wbar -> (1 - d) (1 + s) /. W -> (1 - d) (1 + s h);
harmonicN[tfixadditive /. params] /. params

params = {N0 -> 10^4, d -> 0.05, s -> 0.13, κ -> 100, h -> 0.5, B -> 2};
q0 = q0rescueSGV /. ρ -> pest /. v -> w (3 + 4 B - 4 w) / 4 /. w -> (1 + s h) (1 - d);
qf = 1 -  $\frac{1}{2 N_0 \rho}$  /. ρ -> pest /. v -> w  $\frac{(3 + 4 B - 4 W)}{4 Wbar^2}$  /. w -> ε + 1 /. ε -> 1 -  $\frac{W}{Wbar}$  /.
      Wbar -> (1 - d) (1 + s) /. W -> (1 - d) (1 + s h);
harmonicN[tfixadditive /. params] /. params

params = {N0 -> 10^4, d -> 0.05, s -> 0.13, κ -> 10, h -> 0.5, B -> 2};
q0 = q0rescueSGV /. ρ -> pest /. v -> w (3 + 4 B - 4 w) / 4 /. w -> (1 + s h) (1 - d);
qf = 1 -  $\frac{1}{2 N_0 \rho}$  /. ρ -> pest /. v -> w  $\frac{(3 + 4 B - 4 W)}{4 Wbar^2}$  /. w -> ε + 1 /. ε -> 1 -  $\frac{W}{Wbar}$  /.
      Wbar -> (1 - d) (1 + s) /. W -> (1 - d) (1 + s h);
harmonicN[tfixadditive /. params] /. params

Clear[q0, qf]

Out[367]= 4907.35

Out[371]= 2945.03

Out[375]= 2160.09

Out[379]= 1527.53

```

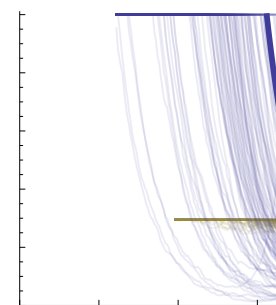
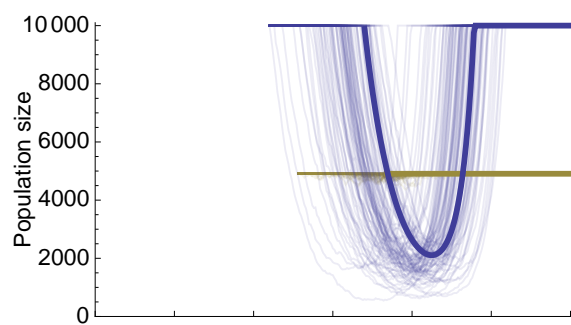
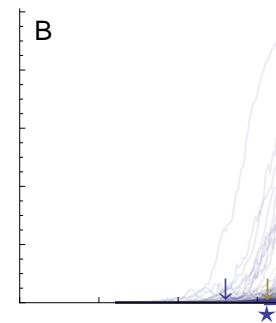
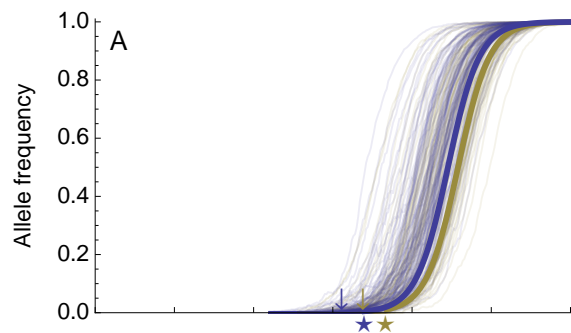
We can now run simulations where population size is approximately constant at each of these values during the sweep, to simulate a simultaneous but independent sweep and bottleneck with the same expected effective population size during the sweep.

Compare the dynamics

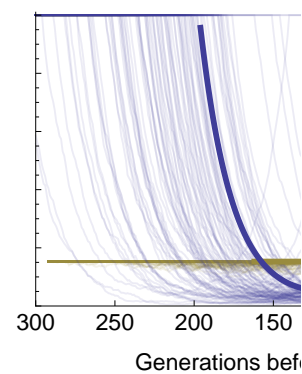
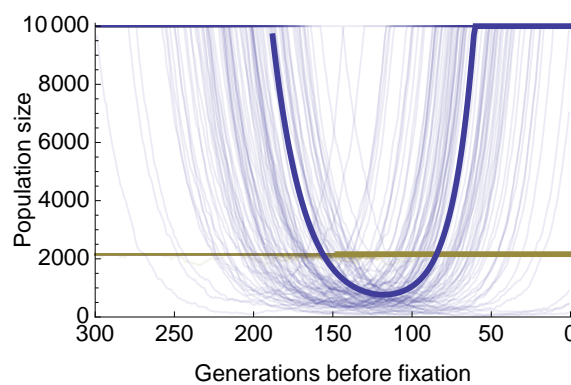
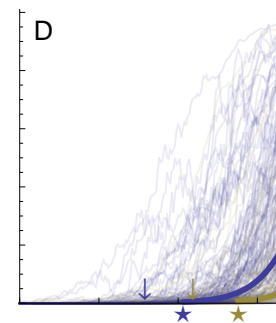
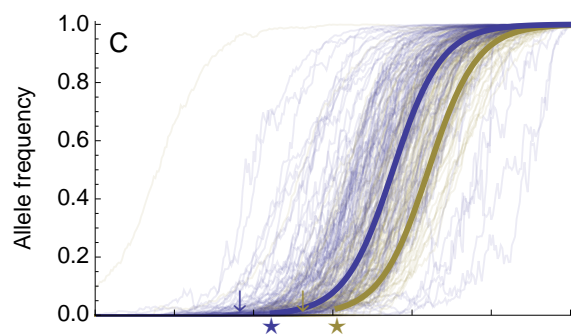
```

In[381]:= letters = {"A", "B"}, {"C", "D"};
ks = {100, 10};
ss = {0.2, 0.13};
Ns = {{4907, 2945}, {2160, 1528}};
GraphicsGrid[
  Table[
    Table[
      plotDynamicsBackwardsBottleneck[104, 0.05, ss[[i]], 0.5,
        ks[[j]], 0, 0, 300, letters[[i, j]], 0, 100, Mean, Ns[[i, j]]],
      {j, Length[ks]}],
    {i, Length[ss]}],
  Spacings → {0, 0}
]

```



Out[385]=



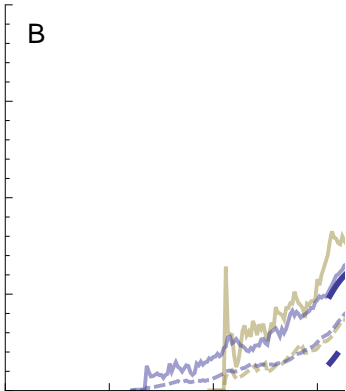
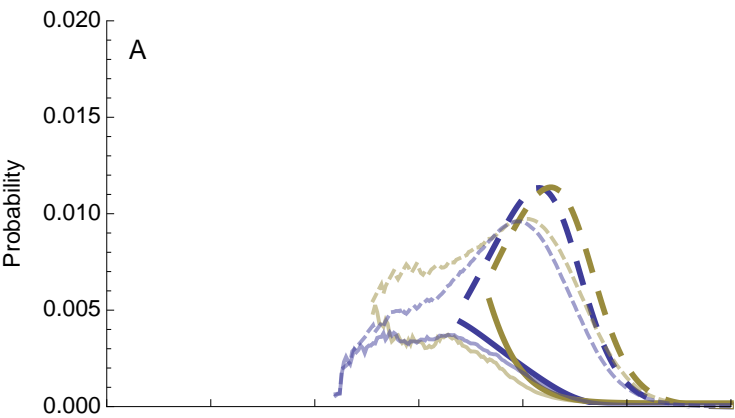
Compare the coalescent



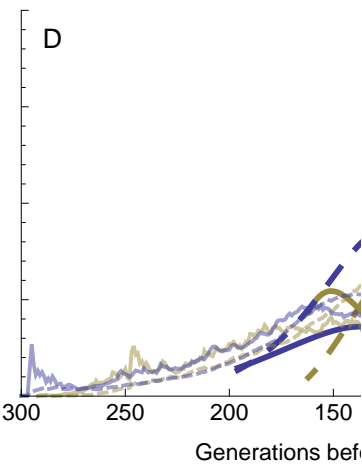
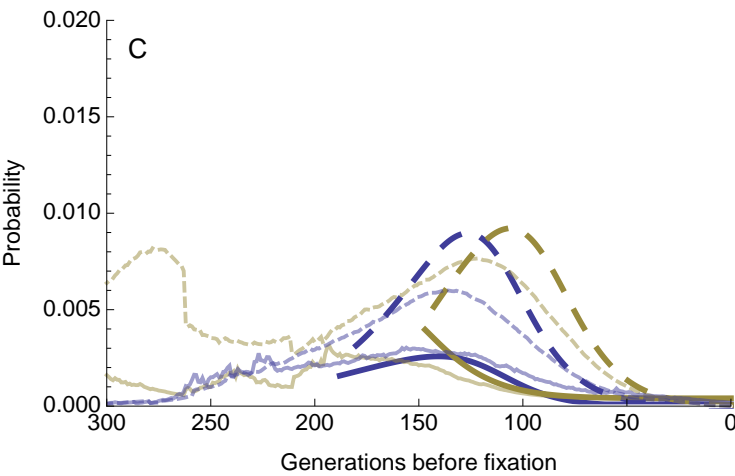
```

In[386]:= letters = {"A", "B"}, {"C", "D"};
ks = {100, 10};
ss = {0.2, 0.13};
Ns = {{4907, 2945}, {2160, 1528}};
GraphicsGrid[
  Table[
    Table[
      plotCoalescentSimsBottleneck[104, 0.05, ss[[i]], 0.5, ks[[j]],
        0, 0, 2, 0.01, 300, 0.02, letters[[i, j]], 0, 100, Ns[[i, j]],
        {j, 2}],
      {i, 2}],
    Spacings → {0, 0}
  ]

```



Out[390]=

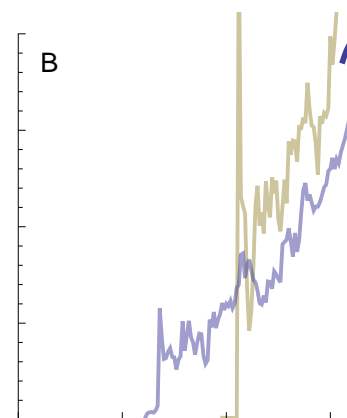
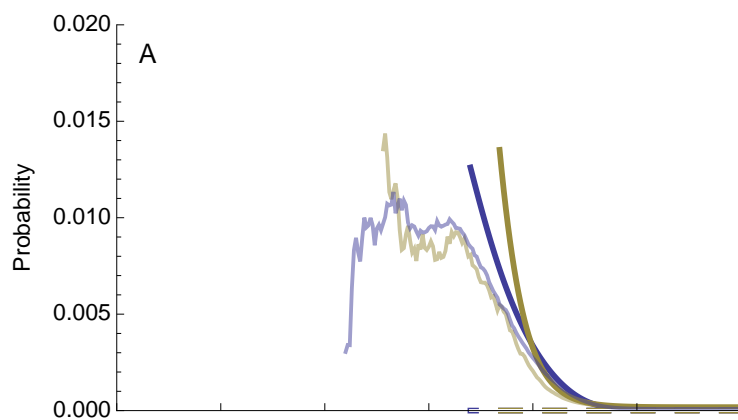


and right at the selected site

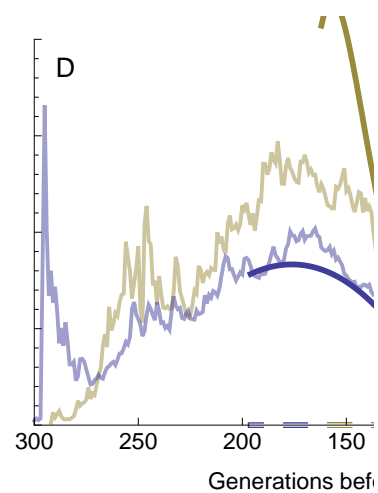
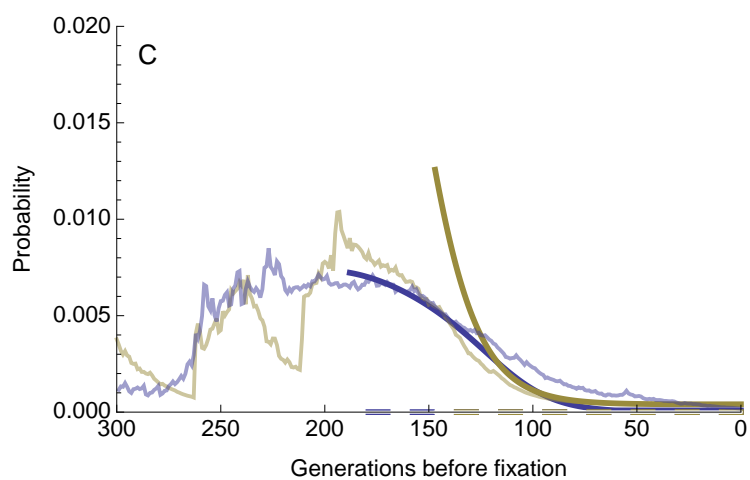
```

In[391]:= letters = {"A", "B"}, {"C", "D"};
ks = {100, 10};
ss = {0.2, 0.13};
Ns = {{4907, 2945}, {2160, 1528}};
GraphicsGrid[
  Table[
    Table[
      plotCoalescentSimsBottleneck[104, 0.05, ss[[i]], 0.5,
        ks[[j]], 0, 0, 2, 0, 300, 0.02, letters[[i, j]], 0, 100, Ns[[i, j]]],
      {j, 2}],
    {i, 2}],
  Spacings → {0, 0}
]

```



Out[395]=

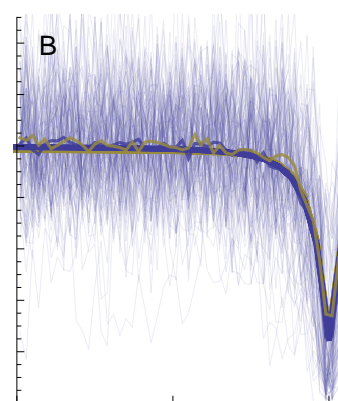
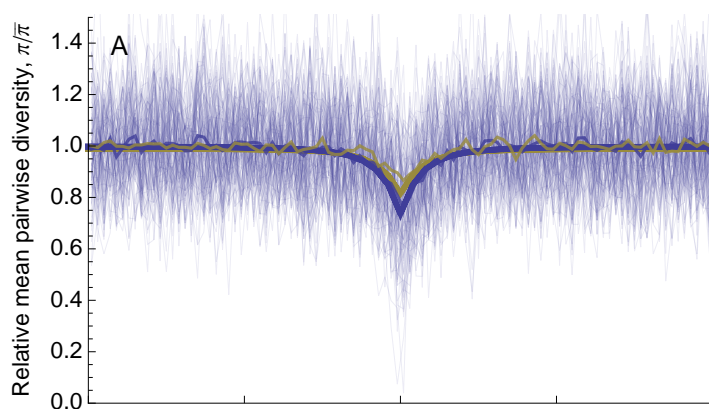


Compare relative pairwise diversity

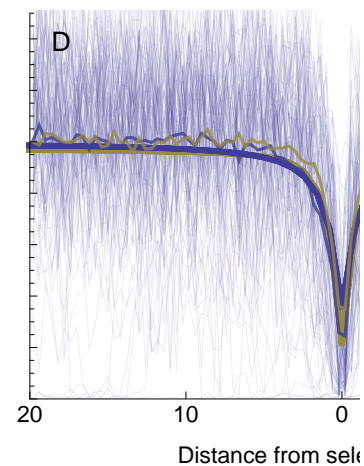
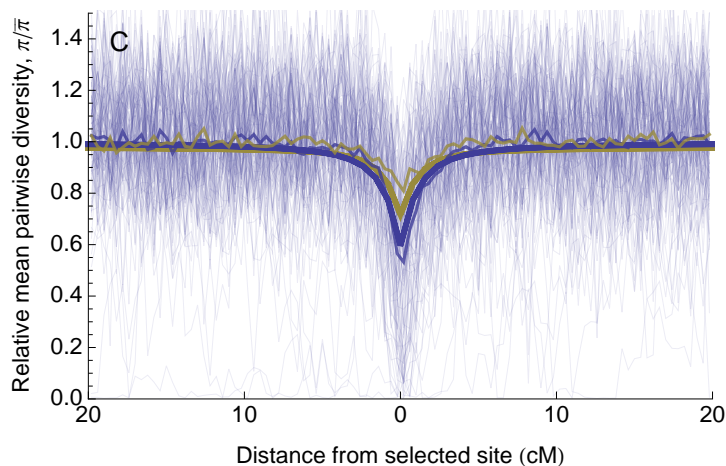
```

In[396]:= letters = {"A", "B"}, {"C", "D"};
ks = {100, 10};
ss = {0.2, 0.13};
Ns = {{4907, 2945}, {2160, 1528}};
GraphicsGrid[
  Table[
    Table[
      plotDiversityRelativeBottleneck[104, 0.05, ss[[i]],
        0.5, ks[[j]], 0, 0, 100, letters[[i, j]], 0, 51, Ns[[i, j]]],
      {j, 2}],
    {i, 2}],
  Spacings → {0, 0}
]

```



Out[400]=

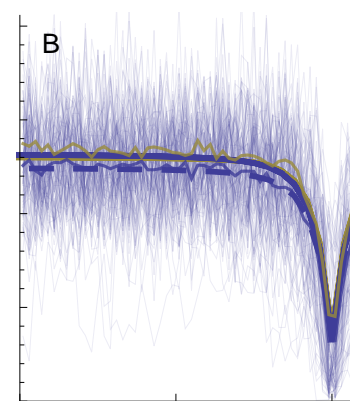
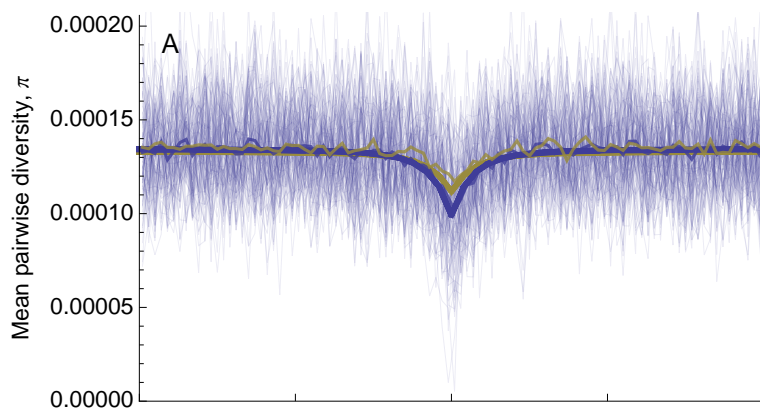


and absolute diversity (these differences mainly arise due to our overestimate of population size in the rescue scenario, which has caused us to simulate bottlenecked populations with larger effective population sizes)

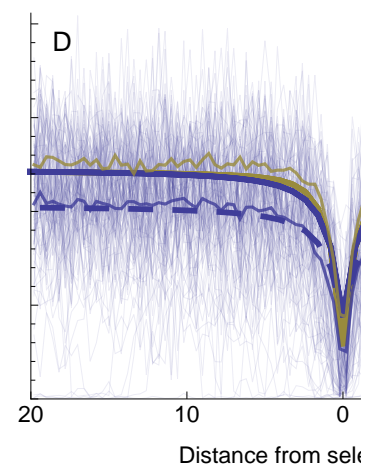
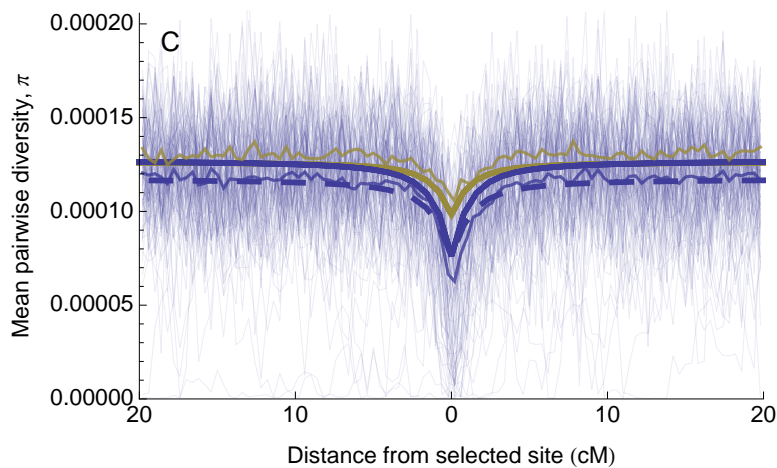
```

In[401]:= letters = {"A", "B"}, {"C", "D"};
ks = {100, 10};
ss = {0.2, 0.13};
Ns = {{4907, 2945}, {2160, 1528}};
GraphicsGrid[
  Table[
    Table[
      plotDiversityBottleneck[104, 0.05, ss[[i]],
        0.5, ks[[j]], 0, 0, 100, letters[[i, j]], 0, 51, Ns[[i, j]]],
      {j, 2}],
    {i, 2}],
  Spacings -> {0, 0}
]

```



Out[405]=

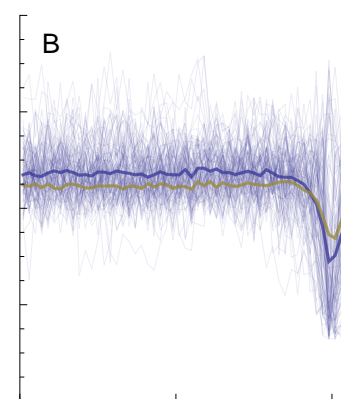
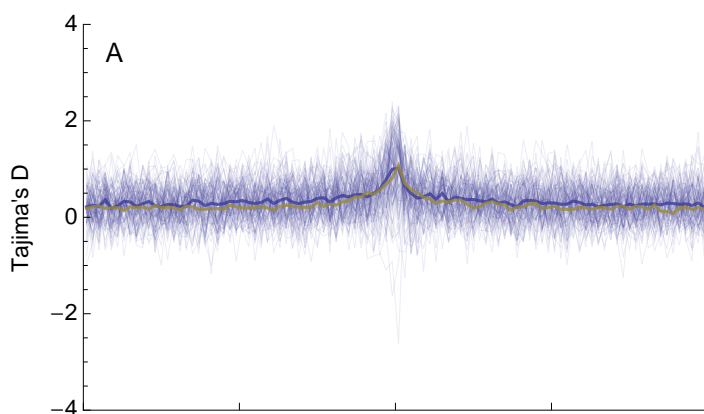


and Tajima's D

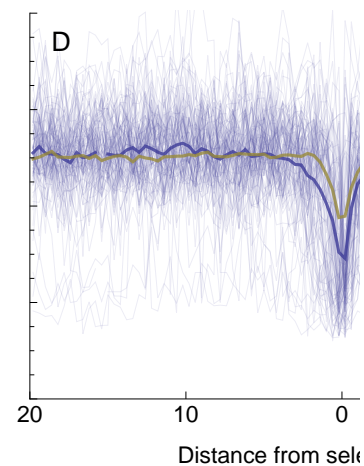
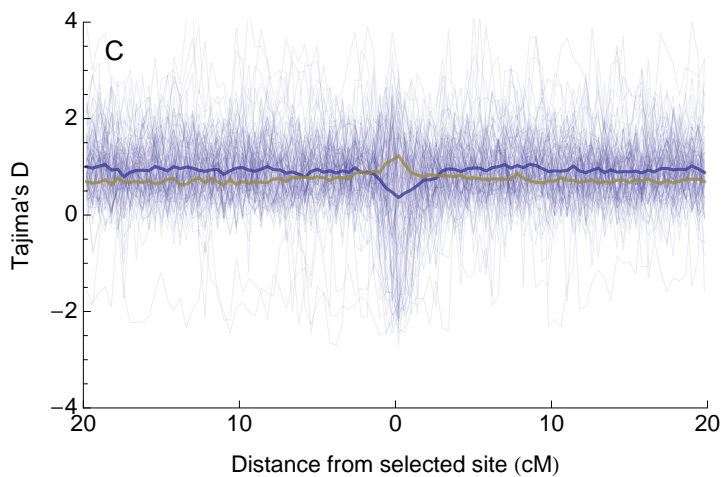
```

In[406]:= letters = {"A", "B"}, {"C", "D"};
ks = {100, 10};
ss = {0.2, 0.13};
Ns = {{4907, 2945}, {2160, 1528}};
GraphicsGrid[
  Table[
    Table[
      plotTajimasDBottleneck[104, 0.05, ss[[i]],
        0.5, ks[[j]], 0, 0, 100, letters[[i, j]], 0, Ns[[i, j]]],
      {j, 2}],
    {i, 2}],
  Spacings → {0, 0}
]

```



Out[410]=



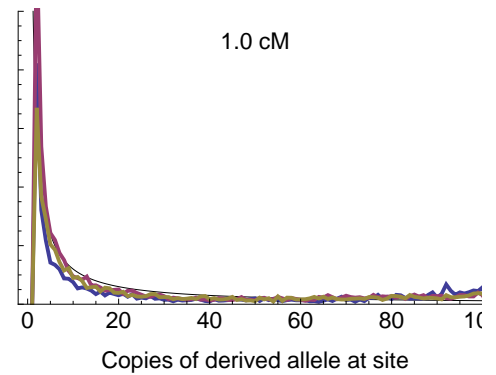
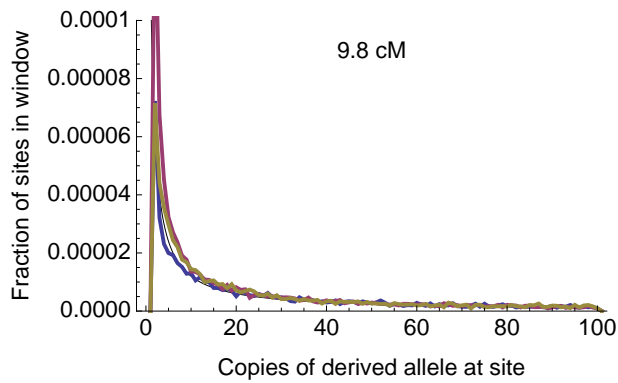
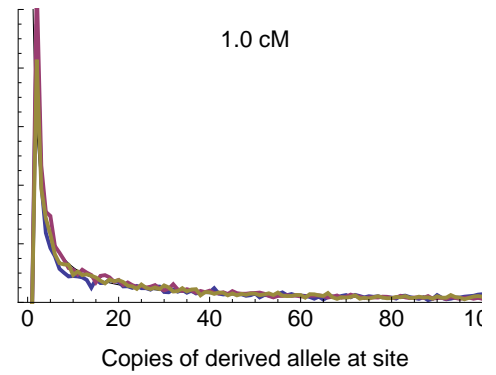
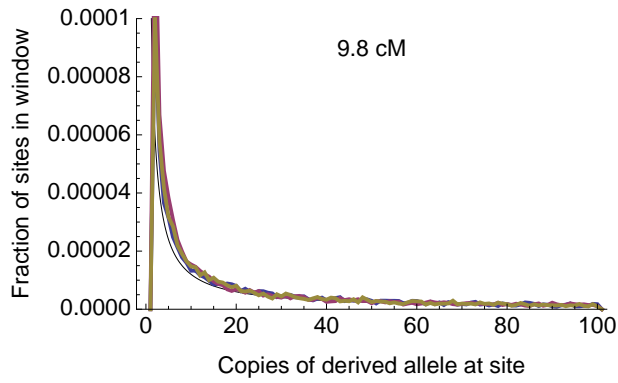
And finally lets just look at the full site frequency spectrum

```

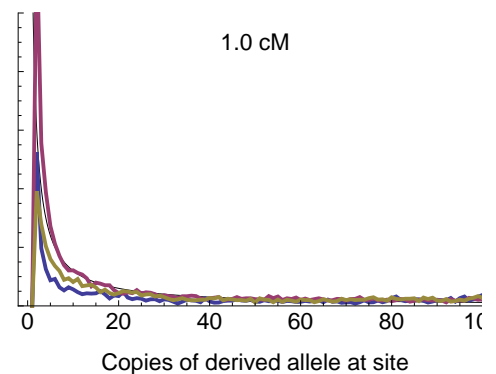
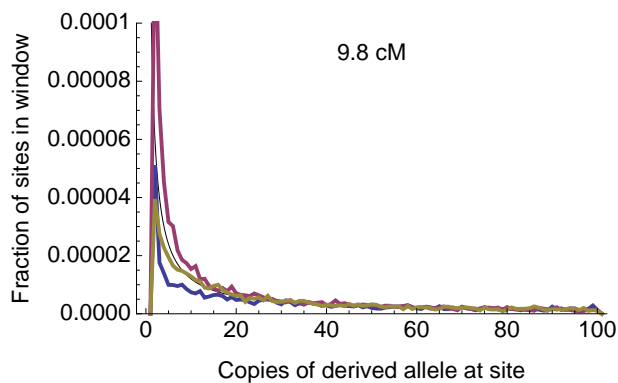
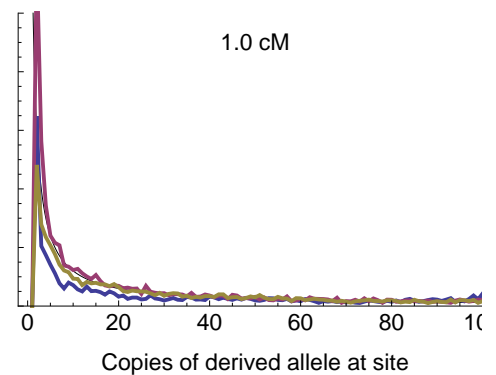
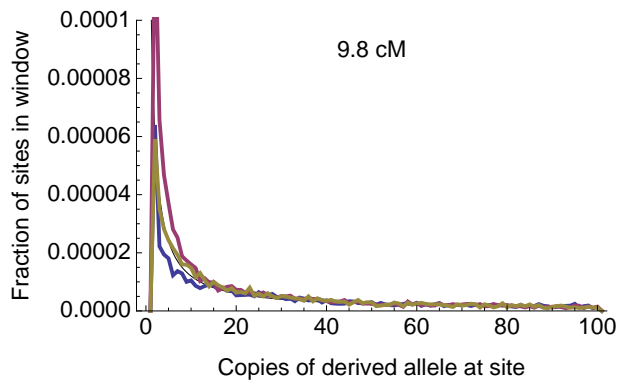
In[411]:= letters = {"A", "B"}, {"C", "D"};
ks = {100, 10};
ss = {0.2, 0.13};
Ns = {{4907, 2945}, {2160, 1528}};
GraphicsGrid[
  Flatten[
    Table[
      Table[
        plotsSFS[104, 0.05, ss[[i]], ks[[j]], 0, 0, 100, Ns[[i, j]], 6 * 10-9,
          {j, 2}],
        {i, 2}],
      {1, 2}],
    ImageSize → 1000
  ]

```





Out[415]=

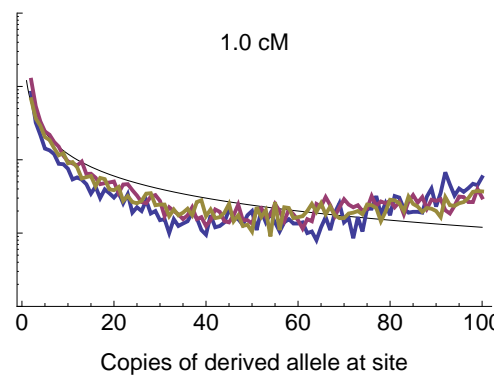
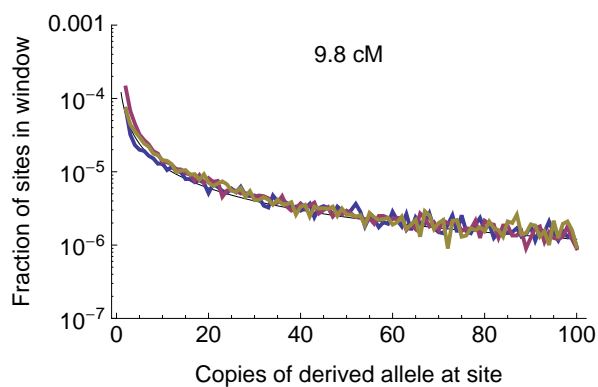
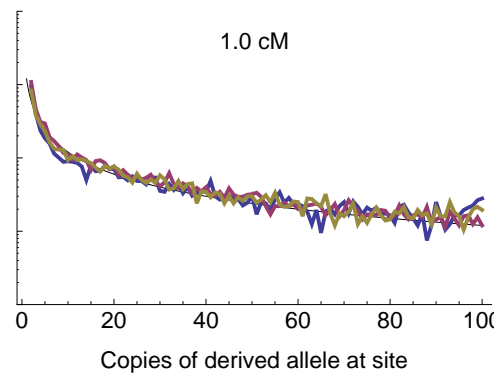
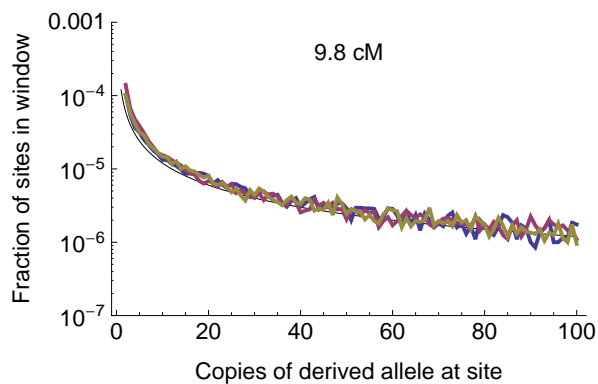


and perhaps also on a log scale

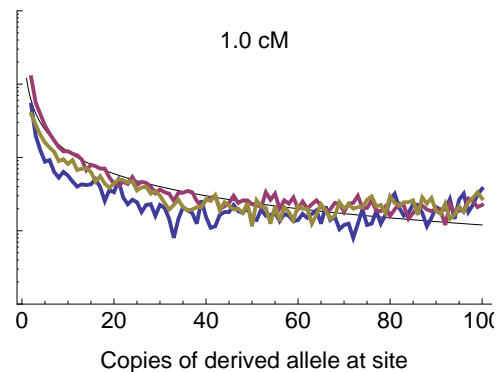
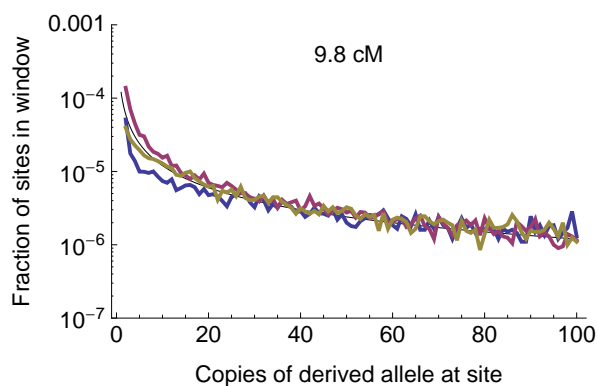
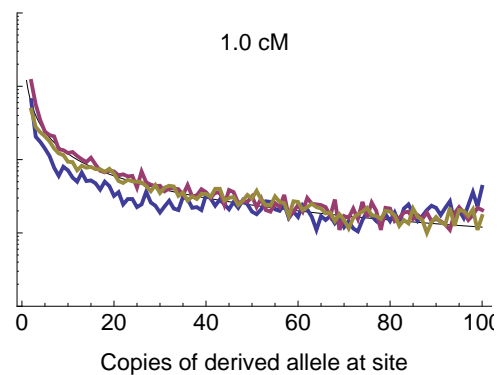
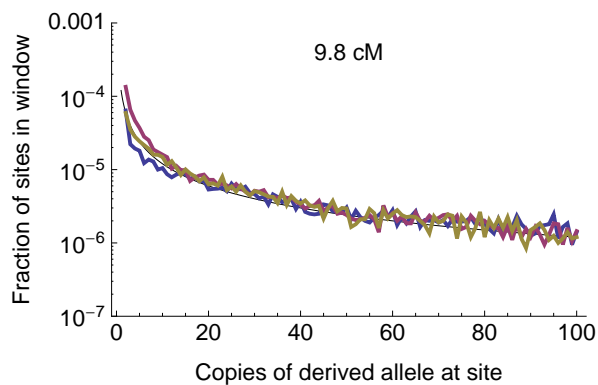
```

In[416]:= letters = {"A", "B"}, {"C", "D"};
ks = {100, 10};
ss = {0.2, 0.13};
Ns = {{4907, 2945}, {2160, 1528}};
GraphicsGrid[
  Flatten[
    Table[
      Table[
        plotsSFSlog[104, 0.05, ss[[i]], ks[[j]], 0, 0, 100, Ns[[i, j]], 6 * 10-9, 12],
        {j, 2}],
      {i, 2}],
    {1, 2}],
  ImageSize -> 1000
]

```



Out[420]=

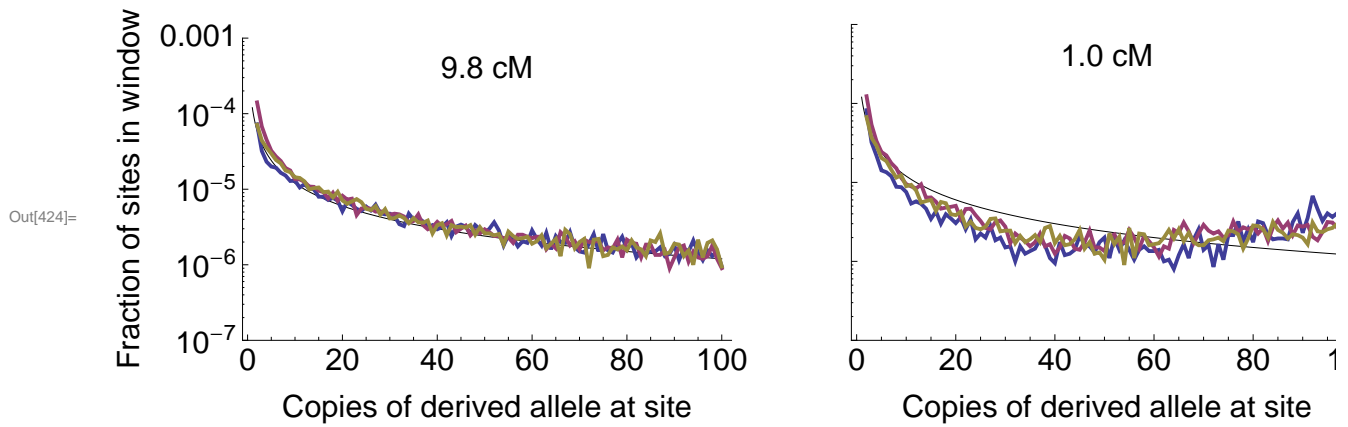


Save one for the manuscript

```

In[421]:= ks = 10;
ss = 0.2;
Ns = 2945;
GraphicsRow[
  plotsSFSlog[10^4, 0.05, ss, ks, 0, 0, 100, Ns, 6 * 10^-9, 16],
  ImageSize -> 1000,
  Spacings -> -30
]
Export[
  imagedir <> ToString[StringForm["SFS_rescue_s`_k`_bottle.pdf", ss, ks]], %];

```

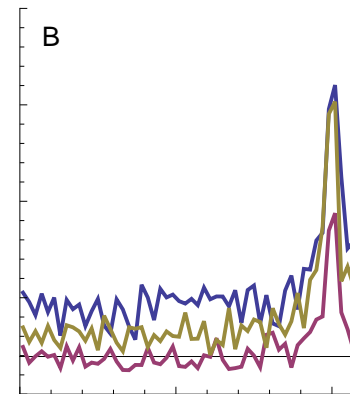
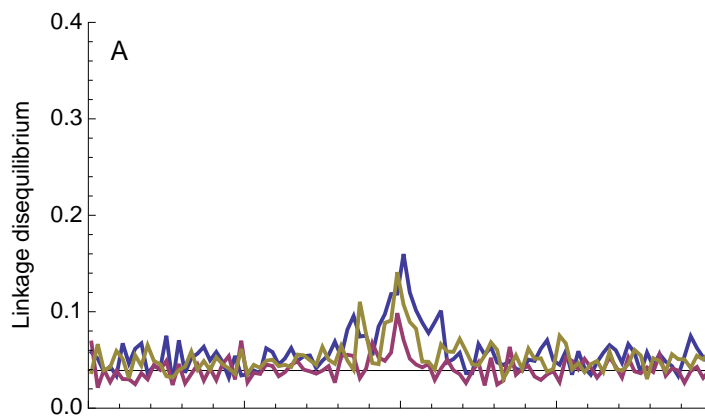


And finally, let's also take a look at linkage disequilibrium

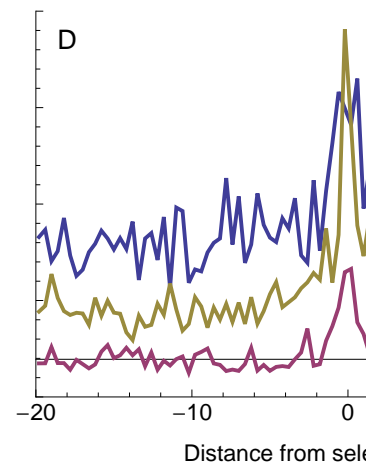
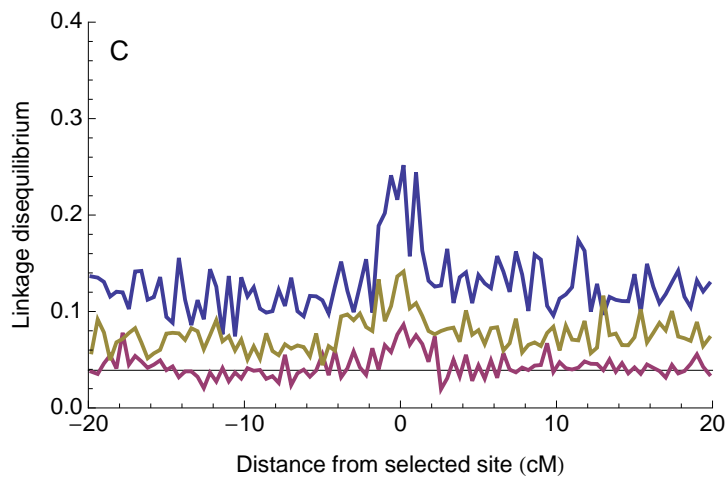
```

In[426]:= letters = {"A", "B"}, {"C", "D"};
ss = {0.2, 0.13};
ks = {100, 10};
ns = {{4907, 2945}, {2160, 1528}};
GraphicsGrid[
  Table[
    Table[
      plotLD[104, 0.05, ss[[i]],
        ks[[j]], 0, 0, 100, ns[[i, j]], 1, 0.4, letters[[i, j]]],
      {j, Length[ks]}],
    {i, Length[ss]}],
  Spacings -> {0, 0}
]

```



Out[430]=



## Sweep + Bottleneck (using observed bottleneck size)

Note that the mean harmonic mean population sizes are substantially lower than expected in rescue

```
In[431]:= getSimNe[104, 0.05, 0.2, 100, 0, 0, 10]
          getSimNe[104, 0.05, 0.2, 10, 0, 0, 10]
          getSimNe[104, 0.05, 0.13, 100, 0, 0, 10]
          getSimNe[104, 0.05, 0.13, 10, 0, 0, 10]
```

```
Out[431]= 3813.75
```

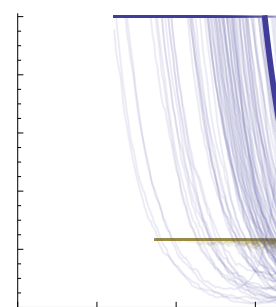
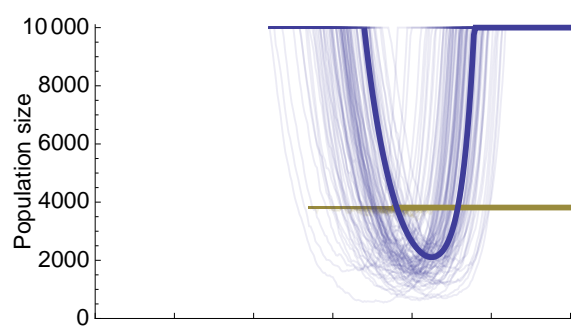
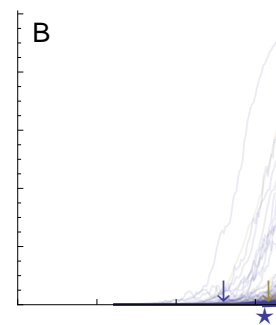
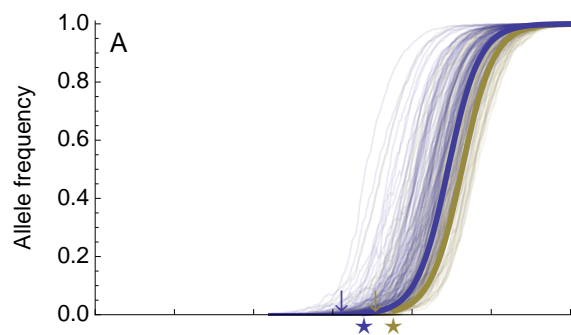
```
Out[432]= 2327.15
```

```
Out[433]= 1883.05
```

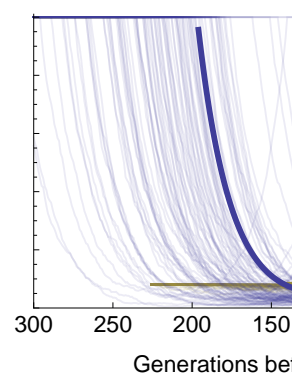
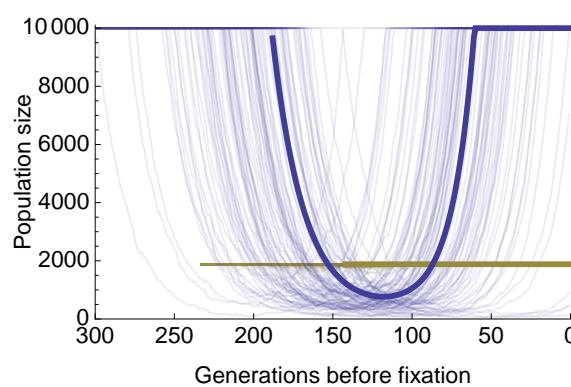
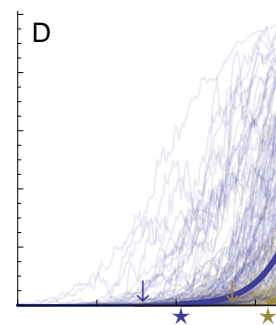
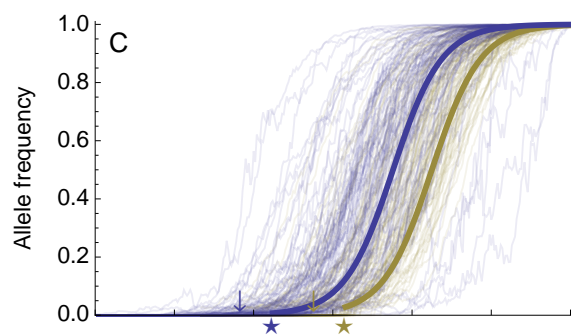
```
Out[434]= 799.321
```

We therefore also simulate under these bottleneck sizes to have a more fair comparison of simulated data.

```
In[435]:= letters = {"A", "B"}, {"C", "D"};
ks = {100, 10};
ss = {0.2, 0.13};
Ns = {{3814, 2327}, {1883, 799}};
GraphicsGrid[
  Table[
    Table[
      plotDynamicsBackwardsBottleneck[104, 0.05, ss[[i]], 0.5,
        ks[[j]], 0, 0, 300, letters[[i, j]], 0, 100, Mean, Ns[[i, j]]],
      {j, Length[ks]}],
    {i, Length[ss]}],
  Spacings → {0, 0}
]
```



Out[439]=

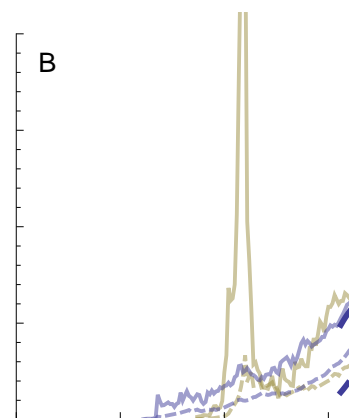
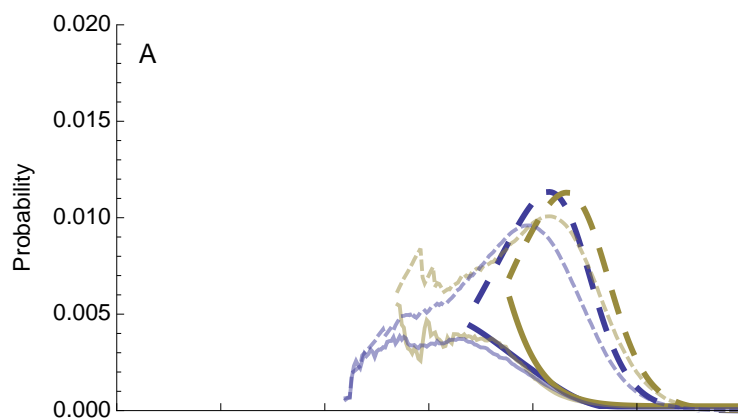


```

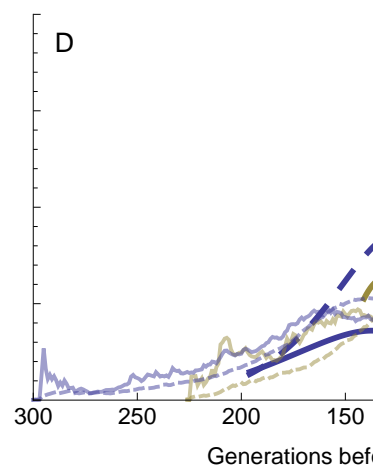
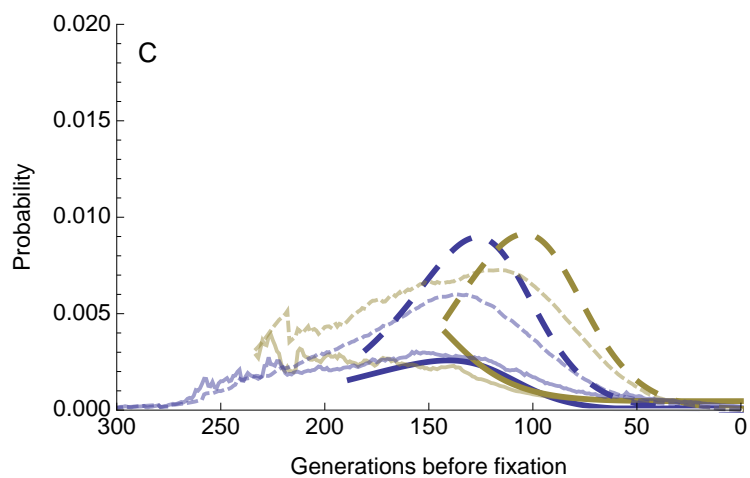
In[440]:= letters = {"A", "B"}, {"C", "D"};
ks = {100, 10};
ss = {0.2, 0.13};
Ns = {{3814, 2327}, {1883, 799}};
GraphicsGrid[
  Table[
    Table[
      plotCoalescentSimsBottleneck[104, 0.05, ss[[i]], 0.5, ks[[j]],
        0, 0, 2, 0.01, 300, 0.02, letters[[i, j]], 0, 100, Ns[[i, j]],
        {j, 2}],
      {i, 2}],
    Spacings → {0, 0}
  ]

```





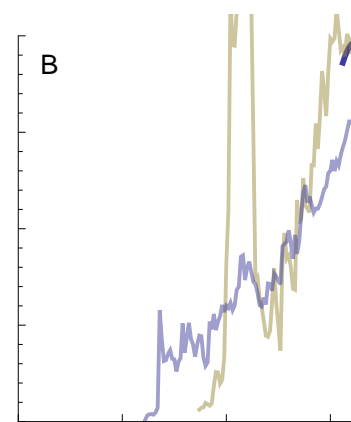
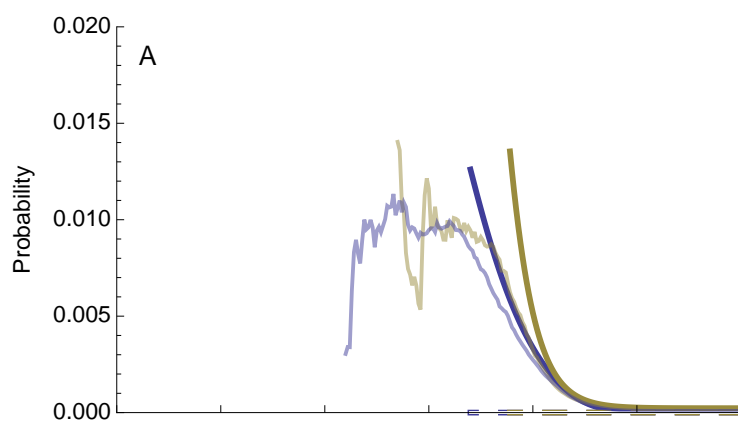
Out[444]=



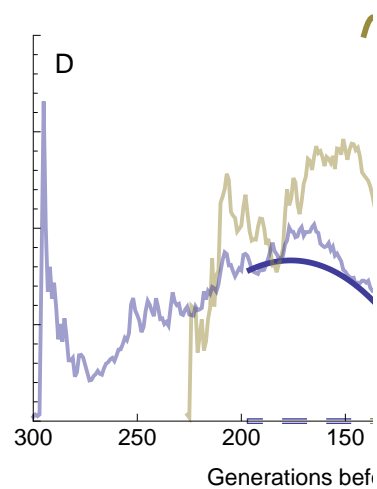
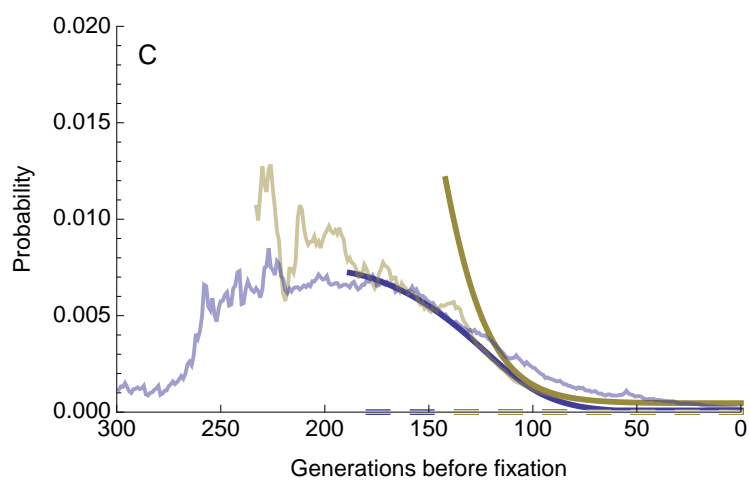
```

In[445]:= letters = {"A", "B"}, {"C", "D"};
ks = {100, 10};
ss = {0.2, 0.13};
Ns = {{3814, 2327}, {1883, 799}};
GraphicsGrid[
  Table[
    Table[
      plotCoalescentSimsBottleneck[104, 0.05, ss[[i]], 0.5,
        ks[[j]], 0, 0, 2, 0, 300, 0.02, letters[[i, j]], 0, 100, Ns[[i, j]]],
      {j, 2}],
    {i, 2}],
  Spacings → {0, 0}
]

```



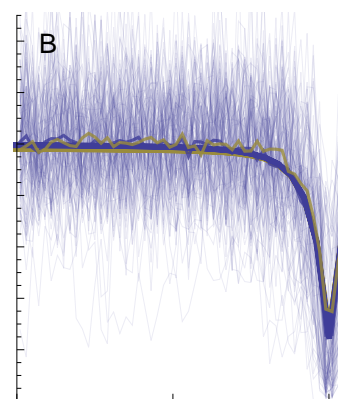
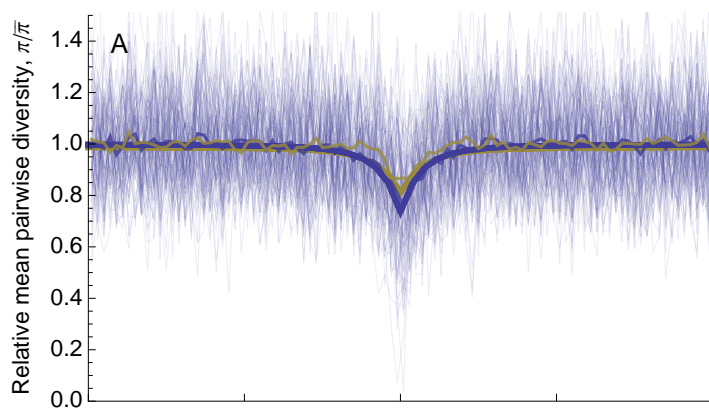
Out[449]=



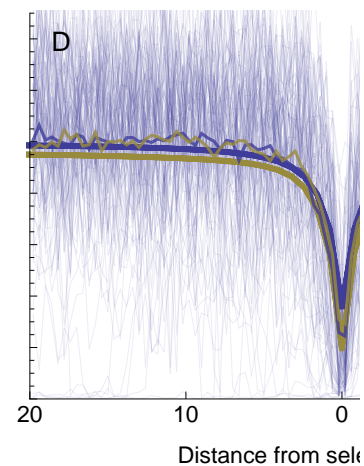
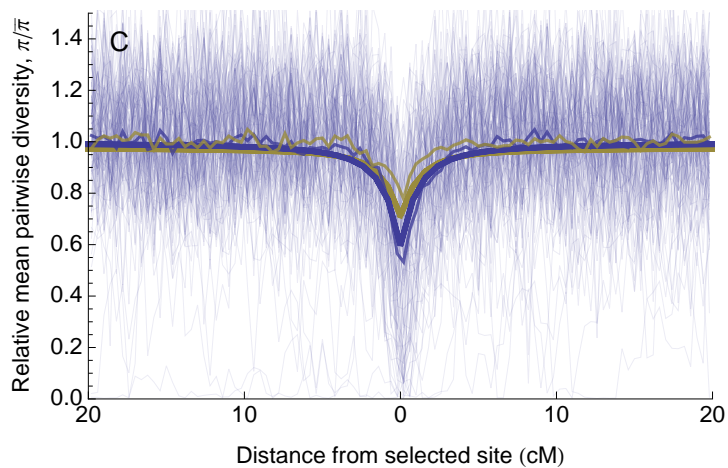
```

In[450]:= letters = {"A", "B"}, {"C", "D"};
ks = {100, 10};
ss = {0.2, 0.13};
Ns = {{3814, 2327}, {1883, 799}};
GraphicsGrid[
  Table[
    Table[
      plotDiversityRelativeBottleneck[104, 0.05, ss[[i]],
        0.5, ks[[j]], 0, 0, 100, letters[[i, j]], 0, 51, Ns[[i, j]]],
      {j, 2}],
    {i, 2}],
  Spacings → {0, 0}
]

```



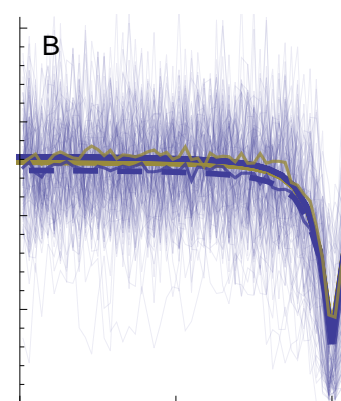
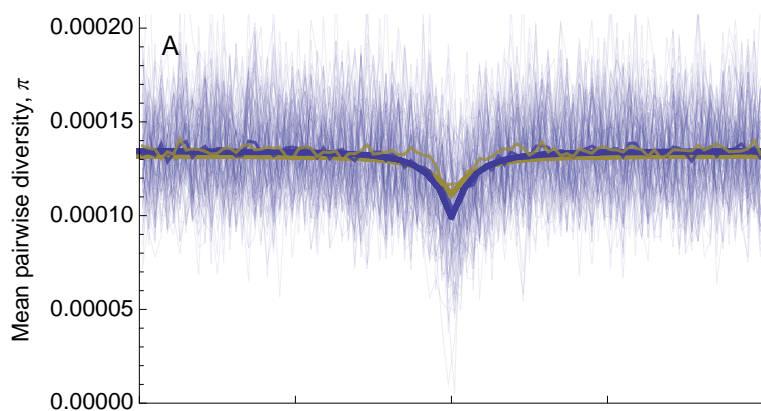
Out[454]=



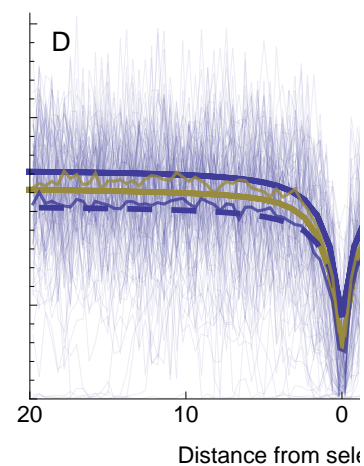
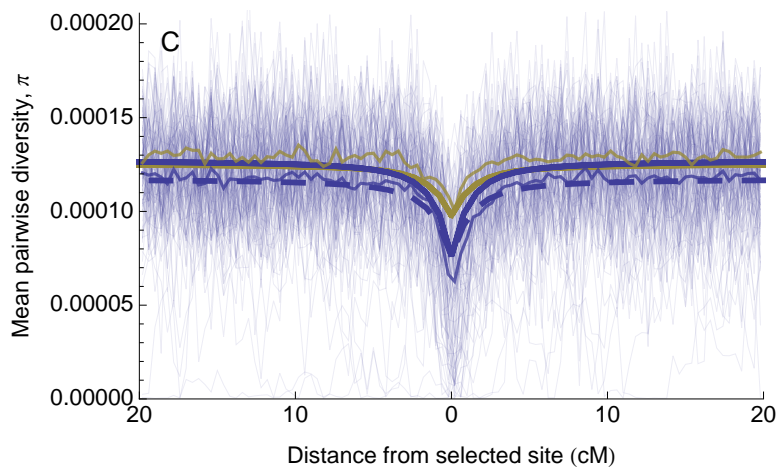
```

In[455]:= letters = {"A", "B"}, {"C", "D"};
ks = {100, 10};
ss = {0.2, 0.13};
Ns = {{3814, 2327}, {1883, 799}};
GraphicsGrid[
  Table[
    Table[
      plotDiversityBottleneck[104, 0.05, ss[[i]],
        0.5, ks[[j]], 0, 0, 100, letters[[i, j]], 0, 51, Ns[[i, j]]],
      {j, 2}],
    {i, 2}],
  Spacings → {0, 0}
]

```



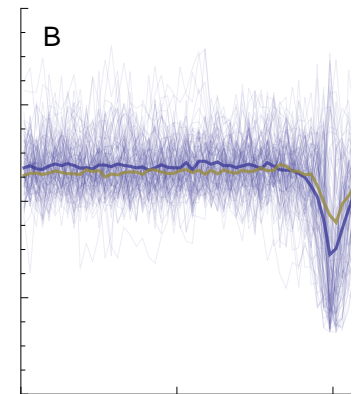
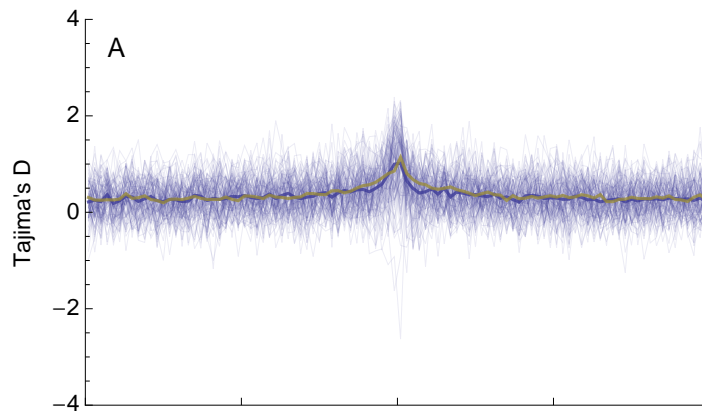
Out[459]=



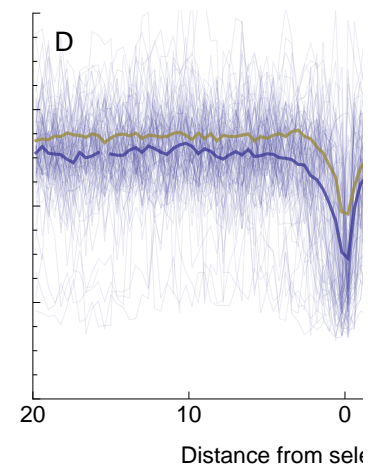
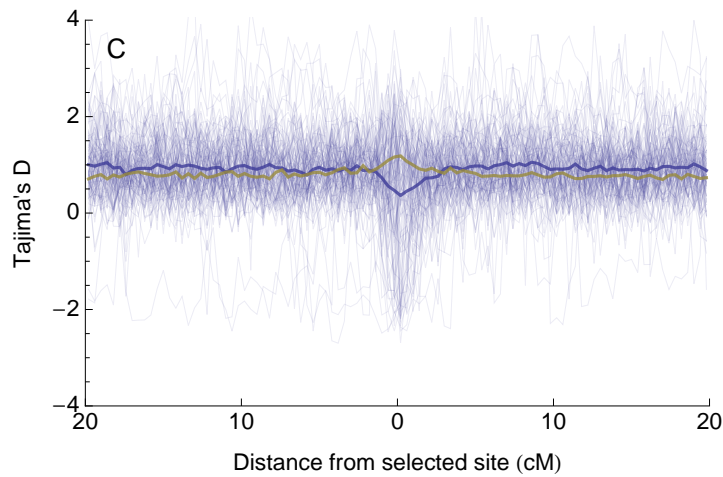
```

In[460]:= letters = {"A", "B"}, {"C", "D"};
ks = {100, 10};
ss = {0.2, 0.13};
Ns = {{3814, 2327}, {1883, 799}};
GraphicsGrid[
  Table[
    Table[
      plotTajimasDBottleneck[104, 0.05, ss[[i]],
        0.5, ks[[j]], 0, 0, 100, letters[[i, j]], 0, Ns[[i, j]]],
      {j, 2}],
    {i, 2}],
  Spacings → {0, 0}
]

```



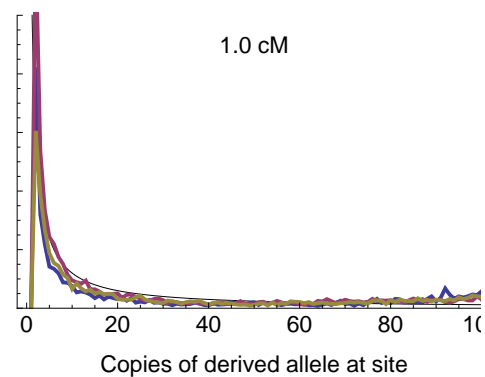
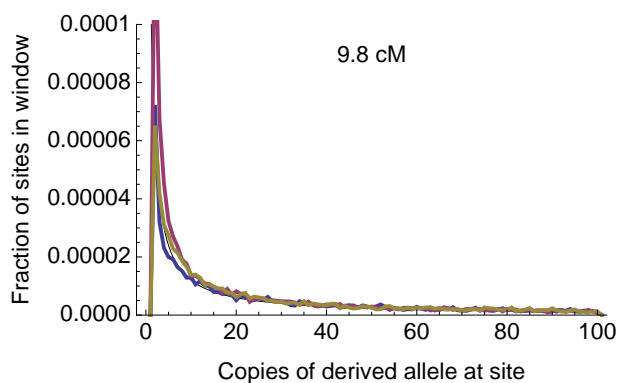
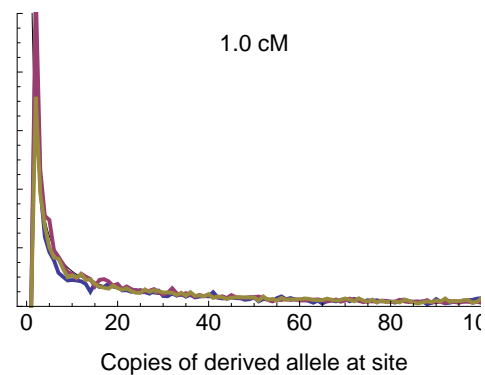
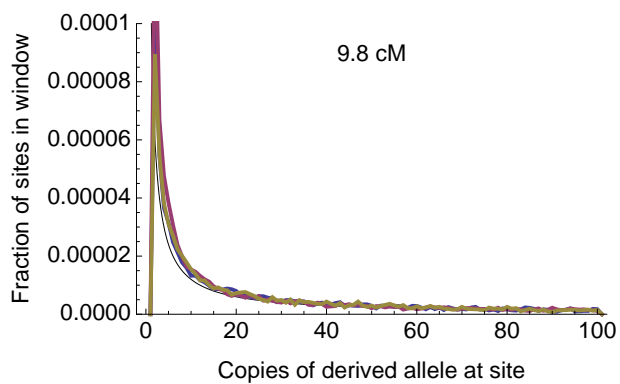
Out[464]=



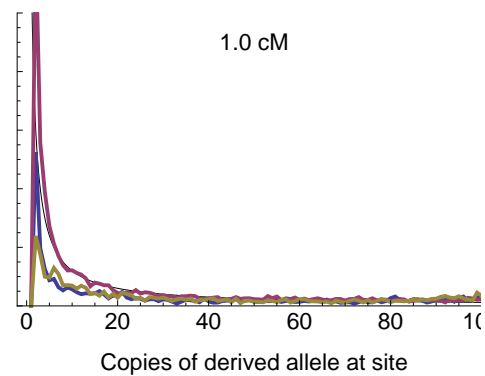
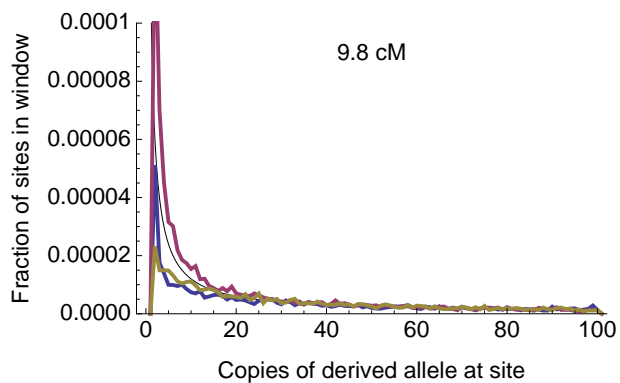
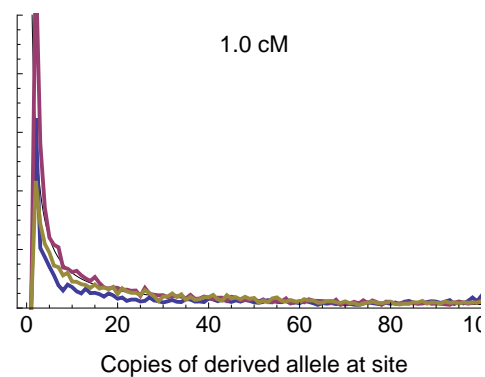
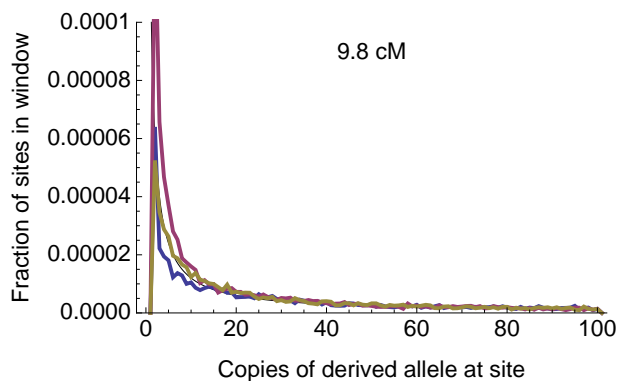
```

In[465]:= letters = {"A", "B"}, {"C", "D"};
ks = {100, 10};
ss = {0.2, 0.13};
Ns = {{3814, 2327}, {1883, 799}};
GraphicsGrid[
  Flatten[
    Table[
      Table[
        plotsSFS[104, 0.05, ss[[i]], ks[[j]], 0, 0, 100, Ns[[i, j]], 6 * 10-9],
        {j, 2}],
      {i, 2}],
    {1, 2}],
  ImageSize → 1000
]

```



Out[469]=

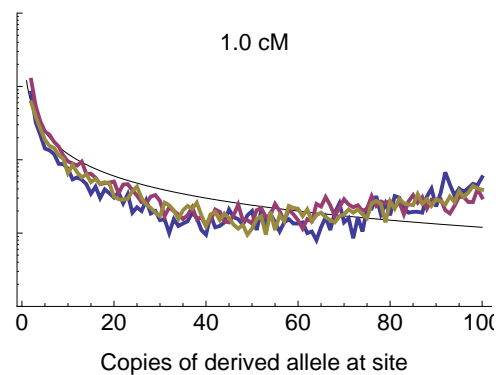
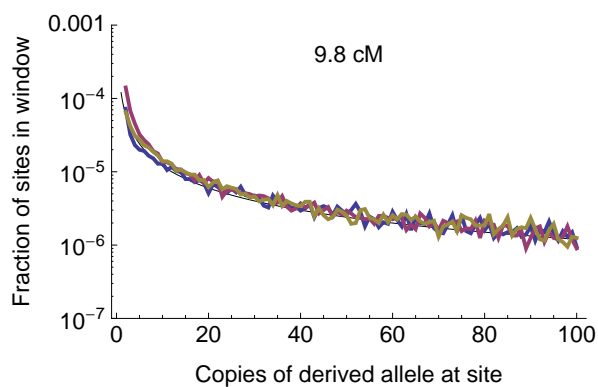
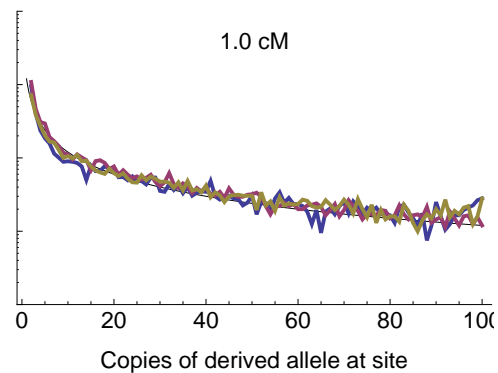
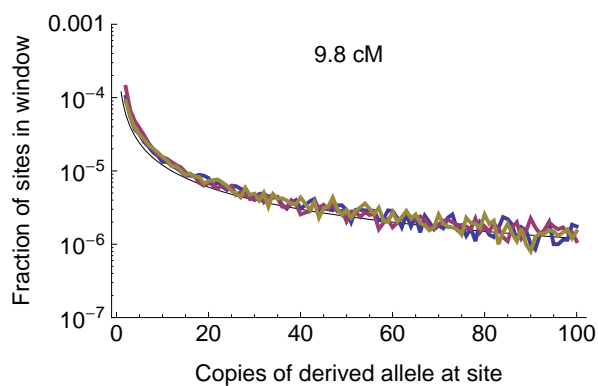




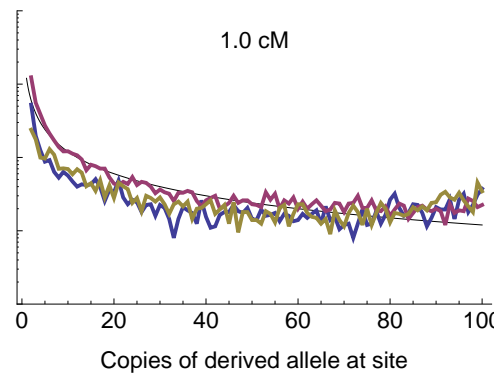
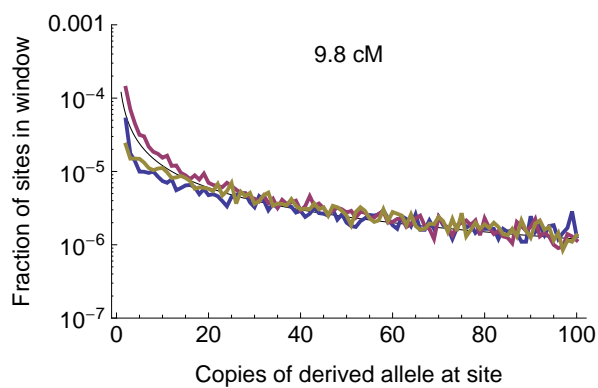
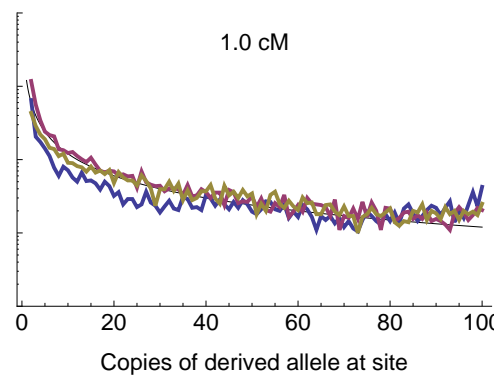
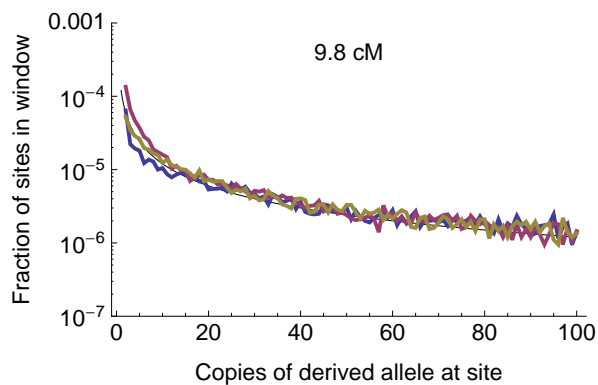
```

In[470]:= letters = {"A", "B"}, {"C", "D"};
ks = {100, 10};
ss = {0.2, 0.13};
Ns = {{3814, 2327}, {1883, 799}};
GraphicsGrid[
  Flatten[
    Table[
      Table[
        plotsFSlog[104, 0.05, ss[[i]], ks[[j]], 0, 0, 100, Ns[[i, j]], 6 * 10-9, 12],
        {j, 2}],
      {i, 2}],
    {1, 2}],
  ImageSize → 1000
]

```



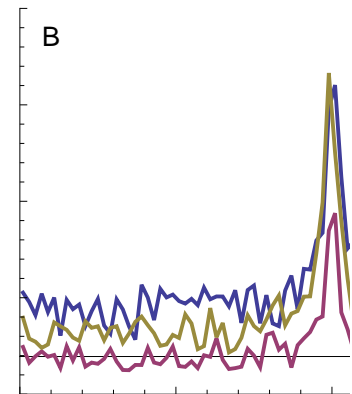
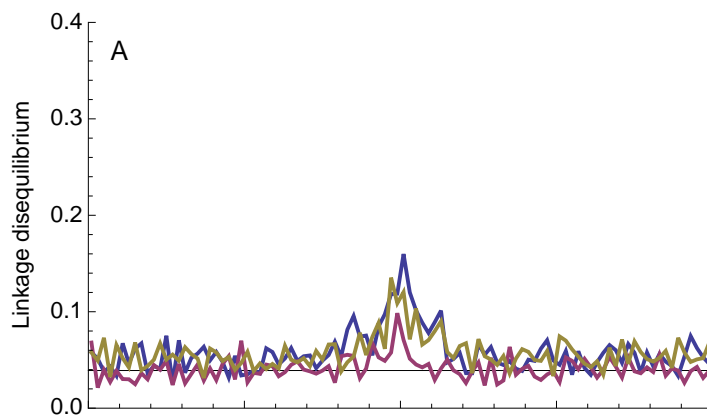
Out[474]=



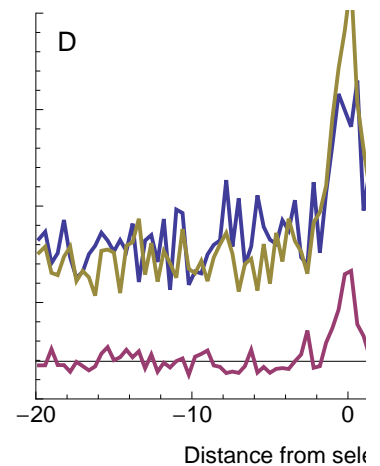
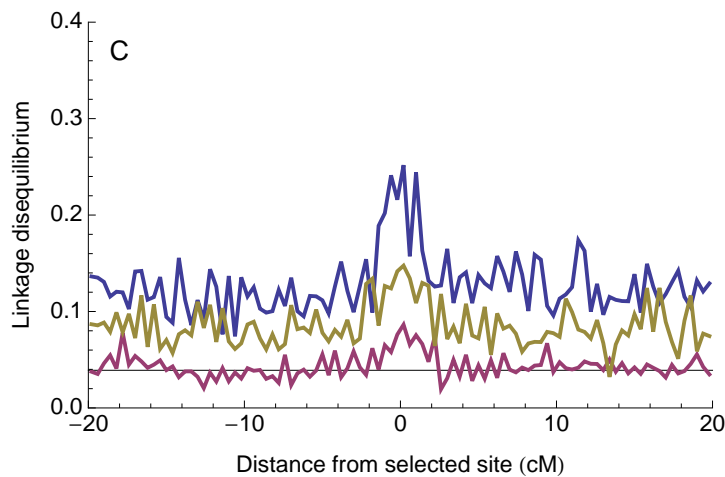
```

In[475]:= letters = {"A", "B"}, {"C", "D"};
ss = {0.2, 0.13};
ks = {100, 10};
Ns = {{3814, 2327}, {1883, 799}};
GraphicsGrid[
  Table[
    Table[
      plotLD[104, 0.05, ss[[i]],
        ks[[j]], 0, 0, 100, Ns[[i, j]], 0, 0.4, letters[[i, j]]],
      {j, Length[ks]}],
    {i, Length[ss]}],
  Spacings → {0, 0}
]

```



Out[479]=



## Rescue from de novo mutation

### The probability of rescue and soft selective sweeps

The first successful rescue mutation arrives according to a time-inhomogeneous Poisson process with

rate  $\lambda(t) = 2 N(t) u \rho$ , so that the probability of at it arriving by time  $T$  is  $1 - \exp\left(-\int_0^T \lambda(t) dt\right)$ . Therefore the probability of rescue is

```
In[480]:= Integrate[2 N0 Exp[-d t] u ρ, {t, 0, T}] // Simplify;
PDNM = Limit[1 - Exp[-%], T → ∞, Assumptions → d > 0]
```

```
Out[481]= 1 - e- $\frac{2 N0 u \rho}{d}$ 
```

Let's check this for a few values of  $s$  and  $u$

```
In[482]:= predictPrescue[sval_, dval_, hval_, Nval_] :=
(
  params = {s → sval, d → dval, h → hval, B → 2, N0 → Nval};
  p = PDNM /. ρ → pest /. v → w (3 + 4 B - 4 w) / 4 /. w → (1 + s h) (1 - d) /. params;
  p
)

getPrescue[dval_, sval_, us_] :=
(
  folder = StringForm["nestablish_DNM_d`_s`"],
  NumberForm[dval, {3, 2}], NumberForm[sval, {3, 2}]]];
  Table[data = Import[datadir <> ToString[folder] <> "/data/nestablish_u" <>
    Which[u = 10-6, "0.000001", u = 10-5.5, "0.000003", u ≥ 10-5,
      ToString[NumberForm[u // N, {7, 6}]]] <> ".txt", "Table"] // Flatten;
    {u, 100 / Length[data] // N},
    {u, us}
  ]
)

plotPrescueDNM[svals_, dval_, Nval_, hval_, us_] :=
(
  (*rescue theory*)
  theory = predictPrescue[svals[[1]], dval, hval, Nval];
  (*rescue theory*)
  theory2 = predictPrescue[svals[[2]], dval, hval, Nval];

  (*rescue simulations*)
  Prescue = getPrescue[dval, svals[[1]], us];
  (*rescue simulations*)
  Prescue2 = getPrescue[dval, svals[[2]], us];

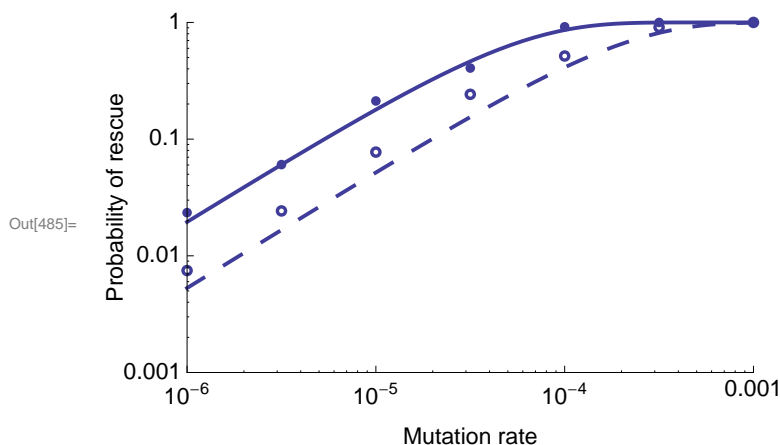
  Show[
    ListLogLogPlot[Prescue,
      PlotStyle → AbsolutePointSize[5], PlotRange → {{10-7, 10-2}, {10-4, 1}}],
    LogLogPlot[theory, {u, us[[1]], us[[-1]]}, PlotStyle → Thick],
    ListLogLogPlot[Prescue2,
      PlotMarkers → Graphics[{Thickness[0.4], Circle[]}, ImageSize → 6]],
    LogLogPlot[theory2, {u, us[[1]], us[[-1]]},
      PlotStyle → Directive[Thick, Dashing[Large]]],
    PlotRange → {Log@{10-6, 10-3}, Log@{10-3, 1}},
    Frame → {True, True, False, False},
```

```

FrameLabel → {"Mutation rate"},
LabelStyle → labelstyle,
ImagePadding → {{50, 15}, {40, 10}},
Epilog → {
  Rotate[Text[
    Style["Probability of rescue", labelstyle], Scaled@{-0.125, 0.5}],  $\pi/2$ ]
],
PlotRangeClipping → False,
PlotRangePadding → None
]
)

plotPrescueDNM[{0.2, 0.13}, 0.05, 104, 0.5, Table[10i/4, {i, -24, -11, 2}]]

```



and we see that it does a pretty good job.

Establishing A alleles arrive at rate  $2 N[t] (1 - q[t]) u \rho$ . For the first successful mutation, the population is overwhelming composed of a alleles and hence  $q$  is essentially zero at all previous times and  $N[t]$  declines exponentially at rate  $d$ . This allows us to derive the relatively simple results above. However, once A has established, the arrival of future successful mutations is complicated by the fact that  $q$  is no longer a constant and  $N[t]$  is no longer declining as a simple exponential. Now the mutation rate can be increased by an upswing in  $N$ , but this is reduced by declining numbers of a alleles,  $1 - q$ .

We can use our recursions above to write the arrival rate,  $\lambda[t]$ , in the next generation as a function of the arrival rate in the current generation

```

In[486]:= 2 (wbar n[t]) (1 - qnew) u  $\rho$  // Simplify;
% /. u →  $\frac{\lambda[t]}{2 n[t] (1 - q[t]) \rho}$  // Simplify
Out[487]= - (-1 + d) (1 + h s q[t])  $\lambda[t]$ 
- (-1 + d) (1 + h s q[t])  $\lambda[t]$ 

```

and we see that for  $h \neq 0$  this also depends on allele frequency because the A alleles prolong the persistence time of the a alleles.

When  $h \neq 0$  we therefore see a non-exponential decline of a alleles due to selection in heterozygotes

```

In[488]:= params = {s → 0.2, h → 0.5, N0 → 104, d → 0.05, k → 0, u → 10.-5, m → 0};
maxrep = 20;
tmax = 200;
legendplace = {6.5 / 8, 1 / 4};

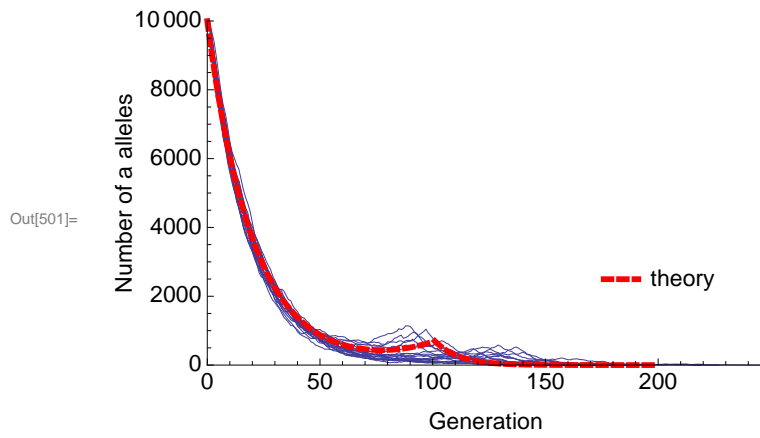
(*rescue: simulations*)
SetDirectory[NotebookDirectory[]];
folder = StringForm["K`_d`_s`_k`_u`_m`", N0 /. params,
  NumberForm[d /. params, {3, 2}], NumberForm[s /. params, {3, 2}],
  k /. params, NumberForm[u /. params, {6, 5}], m /. params];
Clear[data]
Table[data[i] = Import[datadir <> ToString[folder] <> "/data/dynamics_" <>
  ToString[i - 1] <> ".txt", "Table", "FieldSeparators" → " "], {i, maxrep}];
na = Table[(1 - data[i][[2 ;; 3]]) data[i][[2 ;; 2]], {i, maxrep}];

(*rescue: theory*)
q0 = q0rescueDNM /.  $\epsilon \rightarrow w - 1$  /.  $\rho \rightarrow \text{pest}$  /.  $v \rightarrow \frac{1}{4} (3 + 4 B - 4 w) w$  /.
  w → (1 + s h) (1 - d) /. B → 2;
Clear[n];
theoryq = q[t] /. qtadditive /. params;
theoryn = Min[Re[n[t]], N0] /. ntadditive /. params;

Show[
  ListPlot[
    na,
    Joined → True,
    PlotStyle → Directive[AbsoluteThickness[1 / 2], defaultcolors[[1]]],
    PlotRange → All
  ],
  Plot[
    (1 - theoryq) * theoryn, {t, 0, tmax},
    PlotStyle → Directive[AbsoluteThickness[3], Dashed, Red],
    PlotLegends →
      Placed[LineLegend[Style[#, labelstyle] & /@ {"theory"}], Scaled@legendplace],
    PlotRange → All
  ],
  Frame → {True, True, False, False},
  FrameStyle → labelstyle,
  FrameLabel → {"Generation"},
  PlotRangePadding → None,
  ImagePadding → padding,
  PlotRangeClipping → False,
  Epilog → {
    Rotate[
      Text[Style["Number of a alleles", labelstyle], Scaled@ylabelposition],  $\pi / 2$ 
    ],
    PlotRange → {0, N0 /. params}
  ]
]

Clear[q0]

```



but it ain't too far off from exponential for these parameters:

```

In[503]:= params = {s → 0.2, h → 0.5, N0 → 104, d → 0.05, k → 0, u → 10.-5, m → 0};
maxrep = 20;
tmax = 200;
legendplace = {6.5 / 8, 1 / 4};

(*rescue: simulations*)
SetDirectory[NotebookDirectory[]];
folder = StringForm["K`_d`_s`_k`_u`_m`", N0 /. params,
  NumberForm[d /. params, {3, 2}], NumberForm[s /. params, {3, 2}],
  k /. params, NumberForm[u /. params, {6, 5}], m /. params];
Clear[data]
Table[data[i] = Import[datadir <> ToString[folder] <> "/data/dynamics_" <>
  ToString[i - 1] <> ".txt", "Table", "FieldSeparators" → " "], {i, maxrep}];
na = Table[(1 - data[i][[2 ;; 3]]) data[i][[2 ;; 2]], {i, maxrep}];

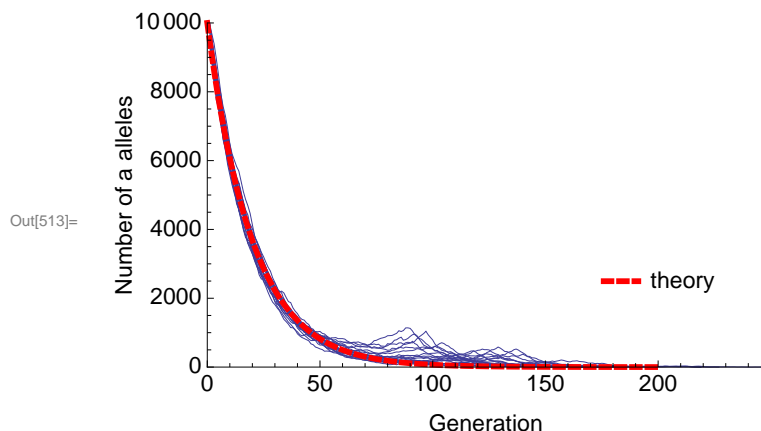
(*rescue: theory*)
theory = N0 Exp[-d t] /. params;

Show[
  ListPlot[
    na,
    Joined → True,
    PlotStyle → Directive[AbsoluteThickness[1 / 2], defaultcolors[[1]]],
    PlotRange → All
  ],
  Plot[
    theory, {t, 0, tmax},
    PlotStyle → Directive[AbsoluteThickness[3], Dashed, Red],
    PlotLegends →
      Placed[LineLegend[Style[#, labelstyle] & /@ {"theory"}], Scaled@legendplace],
    PlotRange → All
  ],
  Frame → {True, True, False, False},
  FrameStyle → labelstyle,
  FrameLabel → {"Generation"},
  PlotRangePadding → None,
  ImagePadding → padding,
  PlotRangeClipping → False,
  Epilog → {
    Rotate[
      Text[Style["Number of a alleles", labelstyle], Scaled@ylabelposition],  $\pi / 2$ 
    ],
  },
  PlotRange → {0, N0 /. params}
]

Clear[q0]

```





So we will use this exponential approximation, which will in general provide an underestimate when  $h > 0$ , and will get worse as  $h$  and  $s$  and  $u$  increase.

We therefore can just use the same rate of arrival of successful mutants that we used for the first mutant and the number of mutations that are expected to establish  $X$  follows a Poisson distribution

```
In[515]:= Integrate[2 N0 Exp[-d t] ρ u, {t, 0, ∞}, Assumptions → d > 0];
Simplify[PDF[PoissonDistribution[%], X], Assumptions → X ≥ 0]
```

Out[516]=

$$\frac{2^X e^{-\frac{2 N_0 u \rho}{d}} \left(\frac{N_0 u \rho}{d}\right)^X}{X!}$$

$$\frac{2^X e^{-\frac{2 N_0 u \rho}{d}} \left(\frac{N_0 u \rho}{d}\right)^X}{X!}$$

the distribution of  $X$  given rescue is then

```
In[517]:= Integrate[2 N0 Exp[-d t] ρ u, {t, 0, ∞}, Assumptions → d > 0];
Simplify[PDF[PoissonDistribution[%], X], Assumptions → X ≥ 0];
% / PDNM
```

Out[519]=

$$\frac{2^X e^{-\frac{2 N_0 u \rho}{d}} \left(\frac{N_0 u \rho}{d}\right)^X}{\left(1 - e^{-\frac{2 N_0 u \rho}{d}}\right) X!}$$

$$\frac{2^X e^{-\frac{2 N_0 u \rho}{d}} \left(\frac{N_0 u \rho}{d}\right)^X}{\left(1 - e^{-\frac{2 N_0 u \rho}{d}}\right) X!}$$

which has expectation

```
In[520]:= Integrate[2 N0 Exp[-d t] ρ u, {t, 0, ∞}, Assumptions → d > 0];
Simplify[PDF[PoissonDistribution[%], X], Assumptions → X ≥ 0];
% / PDNM;
nDNM = Sum[% X, {X, 1, ∞}];
Simplify[nDNM ==  $\frac{-\text{Log}[1 - \text{PDNM}]}{\text{PDNM}}$ , {N0 > 0, u > 0, ρ > 0, d > 0}]
```

Out[524]= True

True

And we can also derive the probability of a soft sweep given rescue

```
In[525]:= Integrate[2 N0 Exp[-d t] ρ u, {t, 0, ∞}, Assumptions → d > 0];
Simplify[PDF[PoissonDistribution[%], X], Assumptions → X ≥ 0];
% / PDNM;
psoftDNM = 1 - (% /. X → 1) // Simplify;
FullSimplify[psoftDNM == 1 - nDNM (1 - PDNM), {N0 > 0, u > 0, ρ > 0, d > 0}]
```

Out[529]= True

True

which is essentially the results of Wilson et al 2017 Genetics, who modeled a haploid population, when we ignore density-dependence (just a slight change to the Poisson rate in their equation 7 due to diploidy and our life-cycle).

Let's look at the probability of a soft sweep given rescue across a range of mutation rates for these two selection coefficients

```
In[530]:= predictPsoft[sval_, dval_, hval_, Nval_] :=
(
  params = {s → sval, d → dval, h → hval, B → 2, N0 → Nval};
  p = psoftDNM /. ρ → pest /. v → w (3 + 4 B - 4 w) / 4 /. w → (1 + s h) (1 - d) /. params;
  p
)

getPsoft[dval_, sval_, us_] :=
(
  folder = StringForm["nestablish_DNM_d`_s`",
    NumberForm[dval, {3, 2}], NumberForm[sval, {3, 2}]];
  Table[data = Import[datadir <> ToString[folder] <> "/data/nestablish_u" <>
    Which[u == 10-6, "0.000001", u == 10-5.75, "0.000002",
      u == 10-5.5, "0.000003", u == 10-5.25, "0.000006", u ≥ 10-5,
      ToString[NumberForm[u // N, {7, 6}]]] <> ".txt", "Table"] // Flatten;
    data = Length[Select[data, # > 1 &]];
    {u, data / 100},
    {u, us}
  ]
)

plotPsoftDNM[svals_, dval_, Nval_, hval_, uss_] :=
(
```

```

(*rescue theory*)
theory = predictPsoft[svals[[1]], dval, hval, Nval];
(*rescue theory*)
theory2 = predictPsoft[svals[[2]], dval, hval, Nval];
(*constant theory*)
theorySweep =

  1 - Product[ $\frac{j}{j + \theta}$ , {j, 2 N0 - 1}] /.  $\theta \rightarrow 2 N0 \text{ NeN } u /. \text{NeN} \rightarrow 4 / 7 /. N0 \rightarrow Nval // N$ ;

(*rescue simulations*)
Prescue = getPsoft[dval, svals[[1]], uss[[1]]];
(*rescue simulations*)
Prescue2 = getPsoft[dval, svals[[2]], uss[[1]]];
(*constant simulations*)
Psweep = getPsoft[0, svals[[1]], uss[[1]]];
(*constant simulations*)
Psweep2 = getPsoft[0, svals[[2]], uss[[2]]];

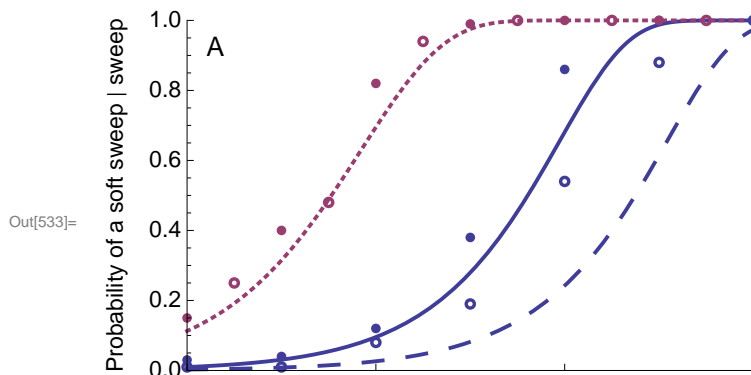
Show[
  LogLinearPlot[theory, {u, uss[[1, 1]], uss[[1, -1]]},
    PlotStyle → Thick, FrameTicks → {{Automatic, Automatic},
      {{{{10-6, "10-6"}, {10-5, "10-5"}, {10-4, "10-4"}, {10-3, "10-3"}}}, Automatic}}],
    ListLogLinearPlot[Prescue, PlotStyle → AbsolutePointSize[5],
      PlotRange → {0, 1}],
    ListLogLinearPlot[Psweep, PlotStyle →
      Directive[AbsolutePointSize[5], defaultcolors[[2]]],
    LogLinearPlot[{theorySweep}, {u, uss[[1, 1]], uss[[1, -1]]},
      PlotStyle → Directive[Thick, Dotted],
    ListLogLinearPlot[Prescue2, PlotMarkers →
      Graphics[{Thickness[0.4], Circle[]}, ImageSize → 6]],
    ListLogLinearPlot[Psweep2, PlotMarkers →
      Graphics[{defaultcolors[[2]], Thickness[0.4], Circle[]}, ImageSize → 6]],
    LogLinearPlot[theory2, {u, uss[[1, 1]], uss[[1, -1]]},
      PlotStyle → Directive[Thick, Dashing[Large]],
      PlotRange → {0, 1},
      Frame → {True, True, False, False},
      FrameTicksStyle → {FontColor → White, Automatic, Automatic, Automatic},
      LabelStyle → labelstyle,
      ImagePadding → {{50, 15}, {40, 10}},
      PlotRange → All,
      Epilog → {
        Text[Style["A", letterstyle], Scaled@letterposition],
        Rotate[Text[Style["Probability of a soft sweep | sweep", labelstyle],
          Scaled@{-0.125, 0.5}],  $\pi / 2$ ]
      },
      PlotRangeClipping → False,
      PlotRangePadding → None
  ]
)

```

```

plotPsoftDNM[{0.2, 0.13}, 0.05, 104, 0.5,
  {Table[10i/4, {i, -24, -11, 2}], Table[10i/4, {i, -23, -12, 2}]}]
Export[imagedir <> "PsoftDNM.pdf", %];

```



and the expected number of copies of the beneficial allele that establish given rescue

```

In[535]:= predictNest[sval_, dval_, hval_, Nval_] :=
(
  params = {s → sval, d → dval, h → hval, B → 2, N0 → Nval};
  n = nDNM /. ρ → pest /. v → w (3 + 4 B - 4 w) / 4 /. w → (1 + s h) (1 - d) /. params;
  n
)

getNest[dval_, sval_, us_] :=
(
  folder = StringForm["nestablish_DNM_d`_s`",
    NumberForm[dval, {3, 2}], NumberForm[sval, {3, 2}]];
  Table[data = Import[datadir <> ToString[folder] <> "/data/nestablish_u" <>
    Which[u == 10-6, "0.000001", u == 10-5.75, "0.000002",
      u == 10-5.5, "0.000003", u == 10-5.25, "0.000006", u ≥ 10-5,
      ToString[NumberForm[u // N, {7, 6}]]] <> ".txt", "Table"] // Flatten;
    data = Select[data, # > 0 &];
    {u, Mean[data], StandardDeviation[data] / Sqrt[Length[data]}},
    {u, us}
  ]
)

plotNestablishDNMrescue[svals_, dval_, Nval_, hval_, uss_] :=
(
  (*rescue theory*)
  theory = predictNest[svals[[1]], dval, hval, Nval];

```

```

(*rescue theory*)
theory2 = predictNest[svals[[2]], dval, hval, Nval];
(*constant theory*)

theorySweep = Sum[ $\frac{\theta}{j-1+\theta}$ , {j, 2 N0}] /.  $\theta \rightarrow 2 N0 \text{ NeN } u$  /.  $N0 \rightarrow Nval$  /.  $\text{NeN} \rightarrow 4 / 7$ ;

(*rescue simulations*)
Prescue = getNest[dval, svals[[1]], uss[[1]]];
(*rescue simulations*)
Prescue2 = getNest[dval, svals[[2]], uss[[1]]];
(*constant simulations*)
Psweep = getNest[0, svals[[1]], uss[[1]]];
(*constant simulations*)
Psweep2 = getNest[0, svals[[2]], uss[[2]]];

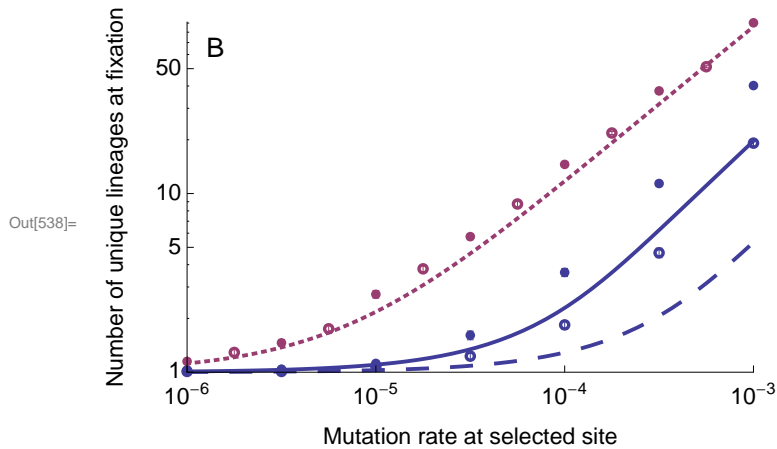
Show[
  LogLogPlot[theory, {u, uss[[1, 1]], uss[[1, -1]]}, PlotStyle -> Thick,
    PlotRange -> {{10-6, 10-3}, {1, 1000}}, FrameTicks -> {{Automatic, Automatic},
      {{{10-6, "10-6"}, {10-5, "10-5"}, {10-4, "10-4"}, {10-3, "10-3"}}}, Automatic]]],
  ErrorListPlot[Prescue, PlotStyle -> AbsolutePointSize[5]] /.
    {x_Real, y_Real} -> {Log@x, Log@y},
  ErrorListPlot[Psweep, PlotStyle -> Directive[AbsolutePointSize[5],
    defaultcolors[[2]]]] /. {x_Real, y_Real} -> {Log@x, Log@y},
  ErrorListPlot[Prescue2, PlotMarkers -> Graphics[{Thickness[0.4], Circle[]},
    ImageSize -> 6]] /. {x_Real, y_Real} -> {Log@x, Log@y},
  ErrorListPlot[Psweep2, PlotStyle -> defaultcolors[[2]], PlotMarkers ->
    Graphics[{defaultcolors[[2]], Thickness[0.4], Circle[]}, ImageSize -> 6]] /.
    {x_Real, y_Real} -> {Log@x, Log@y},
  LogLogPlot[{theorySweep}, {u, uss[[1, 1]], uss[[1, -1]]},
    PlotStyle -> Directive[Thick, Dotted]],
  LogLogPlot[theory2, {u, uss[[1, 1]], uss[[1, -1]]},
    PlotStyle -> Directive[Thick, Dashing[Large]]],
  (*PlotRange->Log@{{10-7, 10-2}, {1, 100}}, *)
  Frame -> {True, True, False, False},
  FrameLabel -> {"Mutation rate at selected site"},
  LabelStyle -> labelstyle,
  ImagePadding -> {{50, 15}, {40, 10}},
  PlotRange -> All,
  Epilog -> {
    Text[Style["B", letterstyle], Scaled@letterposition],
    Rotate[Text[Style["Number of unique lineages at fixation", labelstyle],
      Scaled@{-0.125, 0.5}],  $\pi / 2$ ]
  },
  PlotRangeClipping -> False,
  PlotRangePadding -> None
]
)

```

```

plotNestablishDNMrescue[{0.2, 0.13}, 0.05, 10^4, 0.5,
  {Table[10^(i/4), {i, -24, -11, 2}], Table[10^(i/4), {i, -23, -12, 2}]}]
Export[imagedir <> "NumberEstDNM.pdf", %];

```



## Effective initial allele frequency and the backward-time dynamics

Taking the derivative of the approximate cumulative distribution function with respect to  $T$  and dividing by the probability of rescue gives the PDF of waiting times until the first rescue mutation

```

In[540]:= 1 - Exp[-Integrate[λ[t], {t, 0, T}]];
D[%, T] / PDNM /. λ[t_] → 2 N0 Exp[-d t] u ρ /. T → t

```

$$\text{Out[541]} = \frac{2 e^{-d t} - \frac{2 (1 - e^{-d t}) N0 u \rho}{d}}{1 - e^{-\frac{2 N0 u \rho}{d}}}$$

This has mean

```

In[542]:= 1 - Exp[-Integrate[λ[t], {t, 0, T}]];
D[%, T] / PDNM /. λ[t_] → 2 N0 Exp[-d t] u ρ /. T → t;
Integrate[t %, {t, 0, ∞}, Assumptions → d > 0]

```

$$\text{Out[544]} = \left( \text{EulerGamma} + \text{Gamma}\left[0, -\frac{2 N0 u \rho}{d}\right] + \text{Log}\left[-\frac{2 N0 u \rho}{d}\right] \right) / \left( d - d e^{\frac{2 N0 u \rho}{d}} \right)$$

and standard deviation

```

In[545]:= 1 - Exp[-Integrate[λ[t], {t, 0, T}]];
D[%, T] / PDNM /. λ[t_] → 2 N0 Exp[-d t] u ρ /. T → t;
(Integrate[t^2 %, {t, 0, ∞}, Assumptions → d > 0] -

```

$$\text{Out[547]} = \sqrt{\left( \frac{1}{d^3} 2 N0 u \rho \left( -1 + \text{Coth}\left[\frac{N0 u \rho}{d}\right] \right) \text{HypergeometricPFQ}\left[\{1, 1, 1\}, \{2, 2, 2\}, \frac{2 N0 u \rho}{d}\right] - \left( \text{EulerGamma} + \text{Gamma}\left[0, -\frac{2 N0 u \rho}{d}\right] + \text{Log}\left[-\frac{2 N0 u \rho}{d}\right] \right)^2 / \left( d - d e^{\frac{2 N0 u \rho}{d}} \right)^2 \right)}$$

If we consider a frequency of  $1/2N0\rho$  as the time of arrival (given that will quickly be reached or the mutation lost), we can compare these estimates with simulations

```

In[548]:= timeDNM[Nval_, dval_, sval_, hval_, kval_, uval_, mval_, nreps_] :=
(
  folder = StringForm["K``_d``_s``_k``_u``_m``",
    Nval, NumberForm[dval, {3, 2}], NumberForm[sval, {3, 2}], kval,
    NumberForm[uval, {6, 5}], If[0 < mval < 1, NumberForm[mval, {3, 2}], mval]];
  Clear[data];
  Table[data[i] = Import[datadir <> ToString[folder] <> "/data/dynamics_" <>
    ToString[i - 1] <> ".txt", "Table", "FieldSeparators" → " ", {i, nreps}];
  allp = Table[data[i][[2 ;; 3]], {i, nreps}];
  threshold =
    1 / (2 N0 ρ) /. ρ → pest /. v → (3 + 4 B - 4 w) w / 4 /. w → (1 + s h) (1 - d) /. B → 2 /.
      N0 → Nval /. d → dval /. s → sval /. h → hval /. u → uval;
  data = Table[Position[allp[[i]], x_ /; x > threshold][[1]] // First,
    {i, Length[allp]}];
  theory = 
$$\frac{2 e^{-d t - \frac{2 (1 - e^{-d t}) N_0 u \rho}{d}} N_0 u \rho}{1 - e^{-\frac{2 N_0 u \rho}{d}}} /. \rho \rightarrow \text{pest} /. v \rightarrow \frac{1}{4} (3 + 4 B - 4 w) w /. w \rightarrow (1 + s h) (1 - d) /. B \rightarrow 2 /. \\ (1 - d) /. B \rightarrow 2 /. h \rightarrow hval /. N_0 \rightarrow Nval /. d \rightarrow dval /. s \rightarrow sval /. u \rightarrow uval;$$

  mean = 
$$\left( \text{EulerGamma} + \text{Gamma}\left[0, -\frac{2 N_0 u \rho}{d}\right] + \text{Log}\left[-\frac{2 N_0 u \rho}{d}\right] \right) / \left( d - d e^{\frac{2 N_0 u \rho}{d}} \right) /. \rho \rightarrow \text{pest} /. \\ v \rightarrow \frac{1}{4} (3 + 4 B - 4 w) w /. w \rightarrow (1 + s h) (1 - d) /. B \rightarrow 2 /. \\ h \rightarrow hval /. N_0 \rightarrow Nval /. d \rightarrow dval /. s \rightarrow sval /. u \rightarrow uval;$$

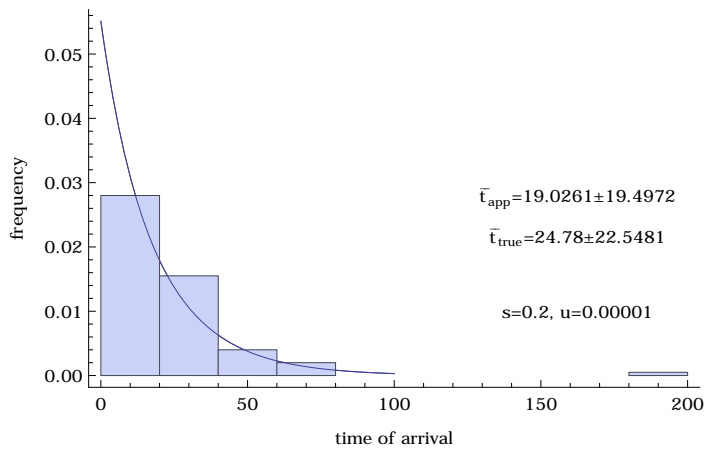
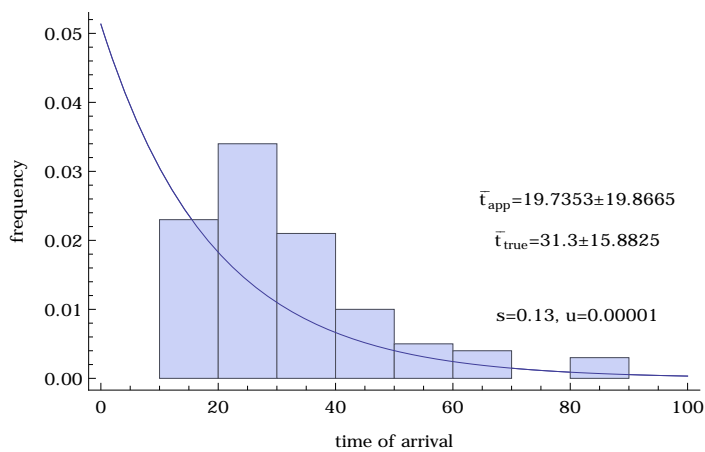
  sd = 
$$\sqrt{\left( \frac{1}{d^3} 2 N_0 u \rho \left( -1 + \text{Coth}\left[\frac{N_0 u \rho}{d}\right] \right) \text{HypergeometricPFQ}\left[\{1, 1, 1\}, \right. \right. \\ \left. \left. \{2, 2, 2\}, \frac{2 N_0 u \rho}{d} \right] - \left( \text{EulerGamma} + \text{Gamma}\left[0, -\frac{2 N_0 u \rho}{d}\right] + \right. \right. \\ \left. \left. \text{Log}\left[-\frac{2 N_0 u \rho}{d}\right] \right)^2 / \left( d - d e^{\frac{2 N_0 u \rho}{d}} \right)^2 \right) /. \rho \rightarrow \text{pest} /. \\ v \rightarrow \frac{1}{4} (3 + 4 B - 4 w) w /. w \rightarrow (1 + s h) (1 - d) /. B \rightarrow 2 /. h \rightarrow hval /. \\ N_0 \rightarrow Nval /. d \rightarrow dval /. s \rightarrow sval /. u \rightarrow uval;$$

  Show[
    Plot[theory, {t, 0, 100}, PlotRange → All],
    Histogram[data, Automatic, "PDF", PlotRange → All],
    Plot[theory, {t, 0, 100}],
    Frame → {True, True, False, False},
    FrameLabel → {"time of arrival", "frequency"},
    Epilog → {
      Text[StringForm["s=`, u=``", sval, uval], Scaled@{0.8, 0.2}],
      Text[StringForm["ttrue=``±``, Mean[data] // N, \\ StandardDeviation[data] // N", Scaled@{0.8, 0.4}],
      Text[StringForm["tapp=``±``, Re[mean] // N, Re[sd] // N", Scaled@{0.8, 0.5}]
    ]
  ]
)

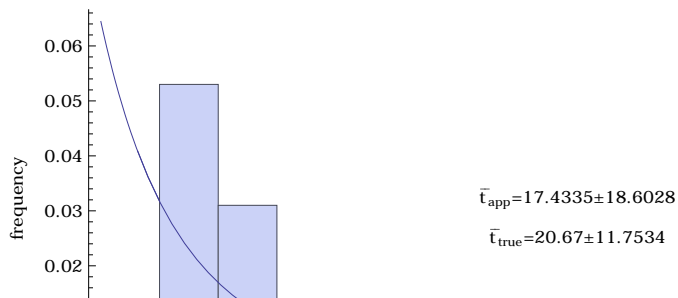
```

)

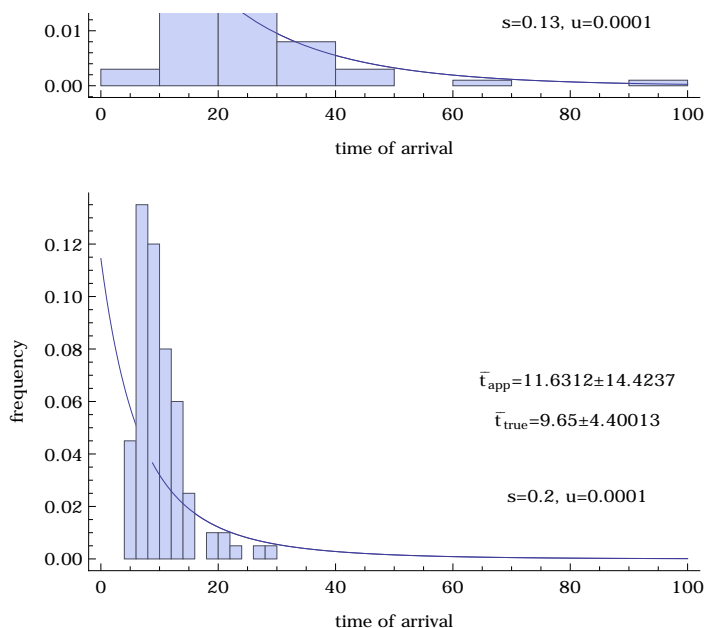
```
GraphicsColumn[{
  timeDNM[104, 0.05, 0.13, 0.5, 0, 10-5 // N, 0, 100],
  timeDNM[104, 0.05, 0.2, 0.5, 0, 10-5 // N, 0, 100],
  timeDNM[104, 0.05, 0.13, 0.5, 0, 10-4 // N, 0, 100],
  timeDNM[104, 0.05, 0.2, 0.5, 0, 10-4 // N, 0, 100]
},
ImageSize -> 400
]
```



Out[549]=







We see a decent correspondance with the means, except maybe in the first case where the realized mean arrival time is a bit later than expected.

Let's call  $c$  the expected number of mutations, then the pdf can be written

$$\begin{aligned} \text{In[550]} &:= 1 - \text{Exp}[-\text{Integrate}[\lambda[t], \{t, 0, T\}]]; \\ &D[\%, T] / \text{PDNM} /. \lambda[t_] \rightarrow 2 N0 \text{Exp}[-d t] u \rho /. T \rightarrow t; \\ &\% /. u \rightarrow \frac{c}{2 N0 \rho / d} \\ \text{Out[552]} &:= \frac{c d e^{-c} (1 - e^{-d t})^{-d t}}{1 - e^{-c}} \end{aligned}$$

The first establishing mutation, arriving at time  $\tau$ , will grow in number nearly exponentially, at rate  $\epsilon = w-1 = (1+s h)(1-d)-1$ , while it is rare. And given that it establishes will very quickly reach  $1/\rho$  copies. Integrating over the arrival times  $\tau$  then gives the expected number of copies  $t$  generations after environmental change

$$\begin{aligned} \text{In[553]} &:= 1 - \text{Exp}[-\text{Integrate}[\lambda[t], \{t, 0, T\}]]; \\ &D[\%, T] / \text{PDNM} /. \lambda[t_] \rightarrow 2 N0 \text{Exp}[-d t] u \rho /. T \rightarrow t; \\ &\% /. u \rightarrow \frac{c}{2 N0 \rho / d} /. t \rightarrow \tau; \\ &\text{Integrate}[(\text{Exp}[\epsilon (t - \tau)] / \rho) \%, \{\tau, 0, \infty\}, \text{Assumptions} \rightarrow \{d > 0, \rho > 0, \epsilon > 0\}] \\ \text{Out[556]} &:= \frac{(-c)^{-\frac{\epsilon}{d}} e^{t \epsilon} \left( -\text{Gamma}\left[\frac{d+\epsilon}{d}\right] + \text{Gamma}\left[\frac{d+\epsilon}{d}, -c\right] \right)}{(-1 + e^c) \rho} \end{aligned}$$

so that it is as if the initial frequency was

```

In[557]:= 1 - Exp[-Integrate[λ[t], {t, 0, T}]];
D[%, T] / PDNM /. λ[t_] → 2 N0 Exp[-d t] u ρ /. T → t;
% /. u →  $\frac{c}{2 N0 \rho / d}$  /. t → τ;
Integrate[(1 / ρ) Exp[ε (t - τ)] %, {τ, 0, ∞}, Assumptions → {d > 0, ρ > 0, ε > 0}];
 $\frac{1}{2 N0}$  % /. t → 0 // Simplify;
q0rescueDNM = % /. c → u 2 N0 ρ / d
Simplify[% ==  $\frac{1}{2 N0 \rho} ((1 - PDNM) (\text{Gamma}[1 + \epsilon / d, \text{Log}[1 - PDNM]] - \text{Gamma}[1 + \epsilon / d]) ) /$ 
 $(PDNM \text{Log}[1 - PDNM]^{\frac{\epsilon}{d}})$ , {N0 > 0, u > 0, ρ > 0, d > 0, ε > 0}]
Out[562]=  $\left( 2^{-1 - \frac{\epsilon}{d}} \left( -\frac{N0 u \rho}{d} \right)^{-\frac{\epsilon}{d}} \left( -\text{Gamma}\left[\frac{d + \epsilon}{d}\right] + \text{Gamma}\left[\frac{d + \epsilon}{d}, -\frac{2 N0 u \rho}{d}\right] \right) \right) / \left( \left( -1 + e^{\frac{2 N0 u \rho}{d}} \right) N0 \rho \right)$ 
Out[563]= True

```

When the expected number of rescue mutations,  $c$ , is very small this reduces to the Orr & Unckless 2014 result (where the waiting time is independent of the mutation rate, the last term in their equation 19)

```

In[564]:= 1 - Exp[-Integrate[λ[t], {t, 0, T}]];
D[%, T] / PDNM /. λ[t_] → 2 N0 Exp[-d t] u ρ /. T → t;
% /. u →  $\frac{c}{2 N0 \rho / d}$  /. t → τ;
Integrate[(Exp[ε (t - τ)] / ρ) %, {τ, 0, ∞}, Assumptions → {d > 0, ρ > 0, ε > 0}];
Simplify[Series[%, {c, 0, 0}] // Normal, {c > 0, ρ > 0, d > 0, ε > 0}]
% /. ε → s - d /. ρ → 2 (s - d) /. d → r // Simplify
Out[568]=  $\frac{d e^{t \epsilon}}{d \rho + \epsilon \rho}$ 
Out[569]=  $-\frac{e^{-r t + s t} r}{2 r s - 2 s^2}$ 

```

In a population of constant size the waiting time for an allele that fixes is exponential

```

In[570]:= Simplify[PDF[ExponentialDistribution[λ], τ], τ > 0] /. λ → 2 N0 u ρ
Out[570]=  $2 e^{-2 N0 u \rho \tau} N0 u \rho$ 

```

which has mean  $1/\lambda$  and standard deviation  $1/\lambda$ .

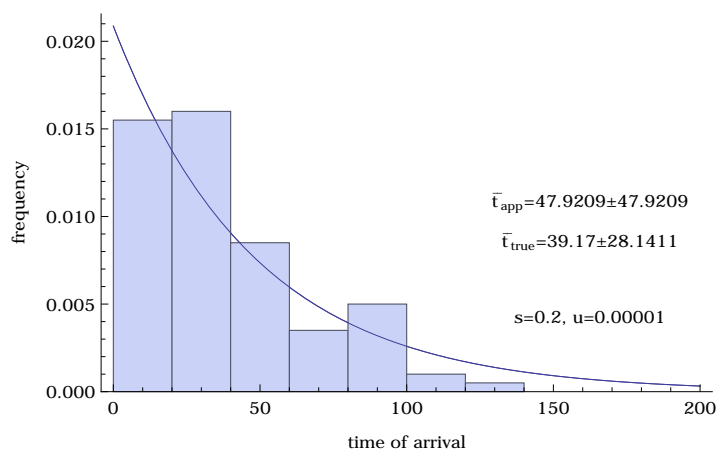
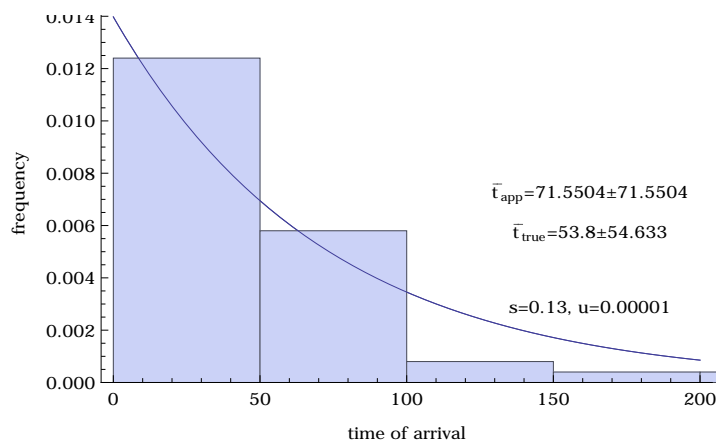
Using  $1/(2N0 \rho)$  as the initial frequency (as for the case of rescue) we can compare this to simulations

```

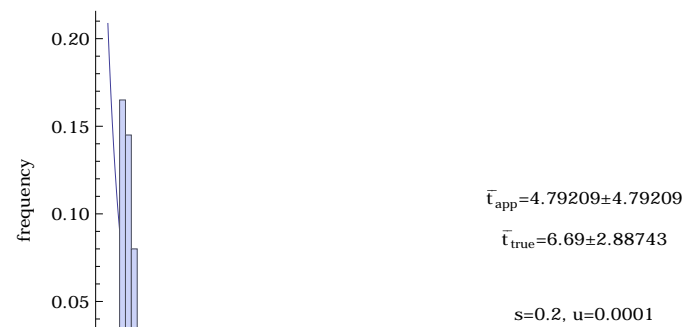
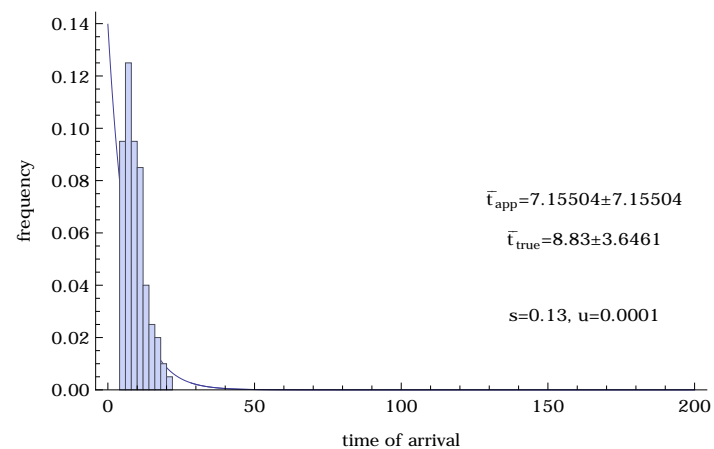
In[571]:= timeDNMconstant[Nval_, dval_, sval_, hval_, kval_, uval_, mval_, nreps_] :=
(
  folder = StringForm["K`_d`_s`_k`_u`_m`",
    Nval, NumberForm[dval, {3, 2}], NumberForm[sval, {3, 2}], kval,
    NumberForm[uval, {6, 5}], If[0 < mval < 1, NumberForm[mval, {3, 2}], mval]];
  Clear[data];
  Table[data[i] = Import[datadir <> ToString[folder] <> "/data/dynamics_" <>
    ToString[i - 1] <> ".txt", "Table", "FieldSeparators" -> " ", {i, nreps}];
  allp = Table[data[i][[2 ;; 3]], {i, nreps}];
  threshold =
    1 / (2 N0  $\rho$ ) /.  $\rho \rightarrow \text{pest}$  /.  $v \rightarrow (3 + 4 B - 4 w) w / 4$  /.  $w \rightarrow (1 + s h) (1 - d)$  /.  $B \rightarrow 2$  /.
    N0 -> Nval /.  $d \rightarrow dval$  /.  $s \rightarrow sval$  /.  $h \rightarrow hval$  /.  $u \rightarrow uval$ ;
  data = Table[Position[allp[[i]], x_ /; x > threshold][[1]] // First,
    {i, Length[allp]}];
  theory =  $e^{-\lambda t} \lambda$  /.  $\lambda \rightarrow 2 N0 u \rho$  /.  $\rho \rightarrow \text{pest}$  /.  $v \rightarrow \frac{1}{4} (3 + 4 B - 4 w) w$  /.  $w \rightarrow (1 + s h) (1 - d)$  /.
    B -> 2 /.  $h \rightarrow hval$  /. N0 -> Nval /.  $d \rightarrow dval$  /.  $s \rightarrow sval$  /.  $u \rightarrow uval$ ;
  mean =  $1 / \lambda$  /.  $\lambda \rightarrow 2 N0 u \rho$  /.  $\rho \rightarrow \text{pest}$  /.  $v \rightarrow \frac{1}{4} (3 + 4 B - 4 w) w$  /.  $w \rightarrow (1 + s h) (1 - d)$  /.
    B -> 2 /.  $h \rightarrow hval$  /. N0 -> Nval /.  $d \rightarrow dval$  /.  $s \rightarrow sval$  /.  $u \rightarrow uval$ ;
  sd =  $1 / \lambda$  /.  $\lambda \rightarrow 2 N0 u \rho$  /.  $\rho \rightarrow \text{pest}$  /.  $v \rightarrow \frac{1}{4} (3 + 4 B - 4 w) w$  /.  $w \rightarrow (1 + s h) (1 - d)$  /.
    B -> 2 /.  $h \rightarrow hval$  /. N0 -> Nval /.  $d \rightarrow dval$  /.  $s \rightarrow sval$  /.  $u \rightarrow uval$ ;
  Show[
    Plot[theory, {t, 0, 200}, PlotRange -> {0, All}],
    Histogram[data, Automatic, "PDF", PlotRange -> All],
    Plot[theory, {t, 0, 200}],
    Frame -> {True, True, False, False},
    FrameLabel -> {"time of arrival", "frequency"},
    Epilog -> {
      Text[StringForm["s=`", u="`", sval, uval], Scaled@{0.8, 0.2}],
      Text[StringForm[" $\bar{t}_{\text{true}} = \pm$ ", Mean[data] // N,
        StandardDeviation[data] // N], Scaled@{0.8, 0.4}],
      Text[StringForm[" $\bar{t}_{\text{app}} = \pm$ ", Re[mean] // N, Re[sd] // N], Scaled@{0.8, 0.5}]
    }
  ]
)

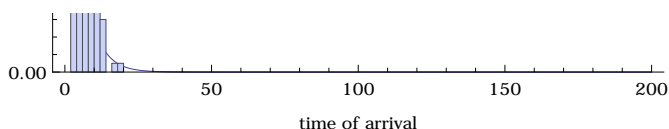
GraphicsColumn[{
  timeDNMconstant[104, 0.0, 0.13, 0.5, 0, 10-5 // N, 0, 100],
  timeDNMconstant[104, 0.0, 0.2, 0.5, 0, 10-5 // N, 0, 100],
  timeDNMconstant[104, 0.0, 0.13, 0.5, 0, 10-4 // N, 0, 100],
  timeDNMconstant[104, 0.0, 0.2, 0.5, 0, 10-4 // N, 0, 100]
},
ImageSize -> 400
]

```



Out[572]=





Given the successful allele arises at time  $\tau$  and fixes, it will very quickly increase to frequency  $1/(2N_0\rho)$ . The frequency will then increase at rate  $\epsilon$  while still rare, so that the frequency at time  $t$  is expected to be

```
In[573]:= Simplify[PDF[ExponentialDistribution[λ], τ], τ > 0];
Integrate[ $\left(\frac{1}{2 N_0 \rho}\right) \text{Exp}[\epsilon (t - \tau)]$  %, {τ, 0, ∞},
Assumptions → {λ > 0, ρ > 0, ε > 0}] /. λ → 2 N0 u ρ // Simplify
```

```
Out[574]= 
$$\frac{e^{t \epsilon} u}{\epsilon + 2 N_0 u \rho}$$

```

Such that it is as if the frequency at time 0 was

```
In[575]:= Simplify[PDF[ExponentialDistribution[λ], τ], τ > 0];
Integrate[ $\left(\frac{1}{2 N_0 \rho}\right) \text{Exp}[\epsilon (t - \tau)]$  %, {τ, 0, ∞},
Assumptions → {λ > 0, ρ > 0, ε > 0}] /. λ → 2 N0 u ρ // Simplify;
% /.
t →
0
```

```
Out[577]= 
$$\frac{u}{\epsilon + 2 N_0 u \rho}$$

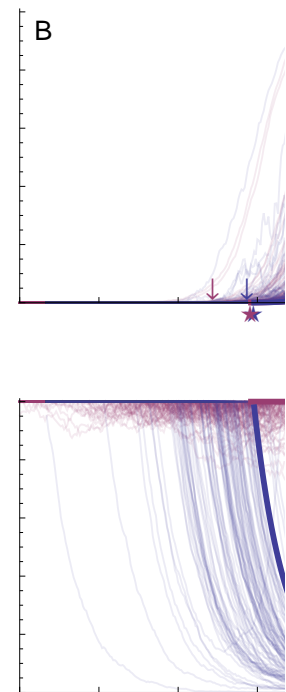
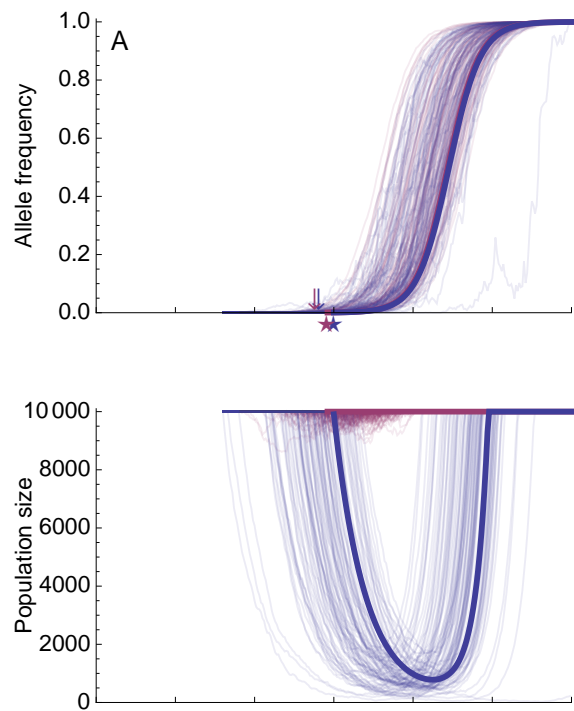
```

However, given that the population size is expected to be constant and the allele frequency dynamics backwards in time do not depend on the time the sweep starts, we need only concern ourselves with the effective initial frequency at the time of establishment

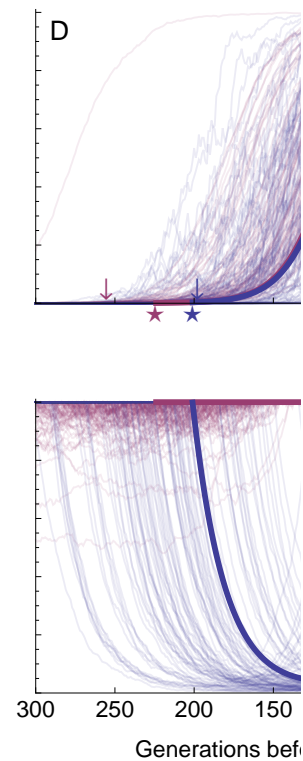
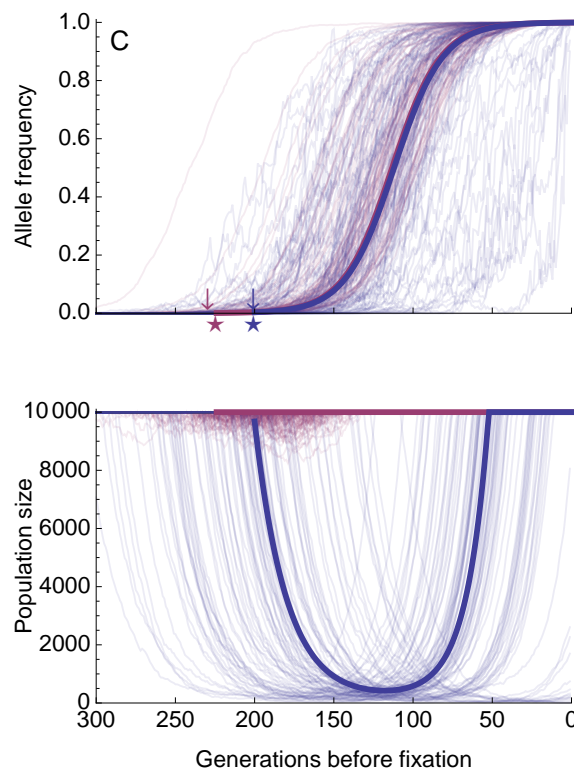
```
In[578]:= q0sweepDNM =  $\frac{1}{2 N_0 \rho}$ ;
```

Plot the backward-time dynamics

```
In[579]:= letters = {"A", "B"}, {"C", "D"};
us = {10.-4, 10.-5};
ss = {0.2, 0.13};
GraphicsGrid[
Table[
Table[
plotDynamicsBackwards[104, 0.05, ss[[i]],
0.5, 0, us[[j]], 0, 300, letters[[i, j]], 1, 100, Median],
{j, 2}],
{i, 2}],
Spacings → {0, 0}
]
```

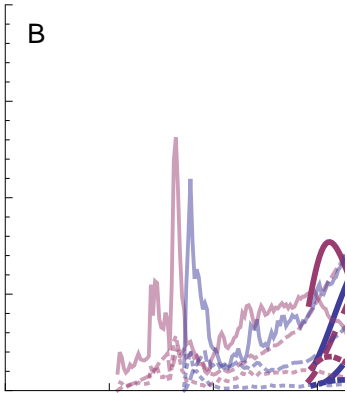
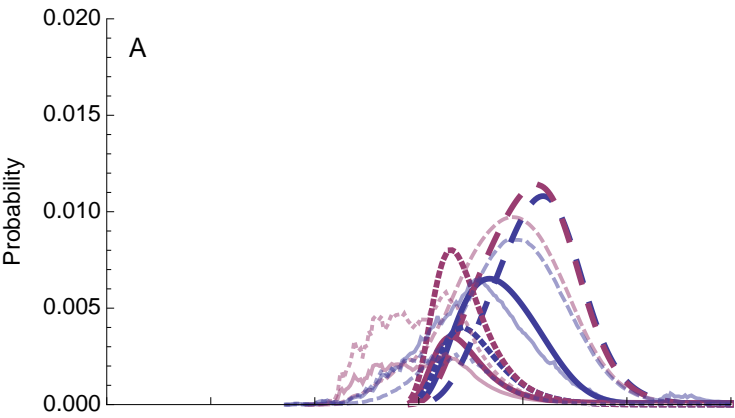


Out[582]=

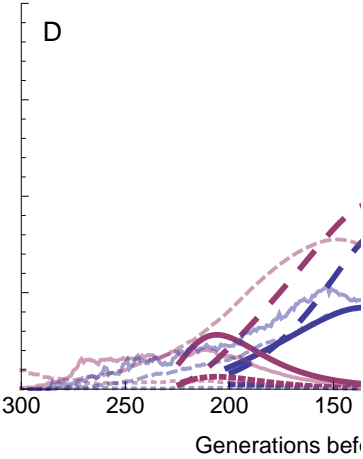
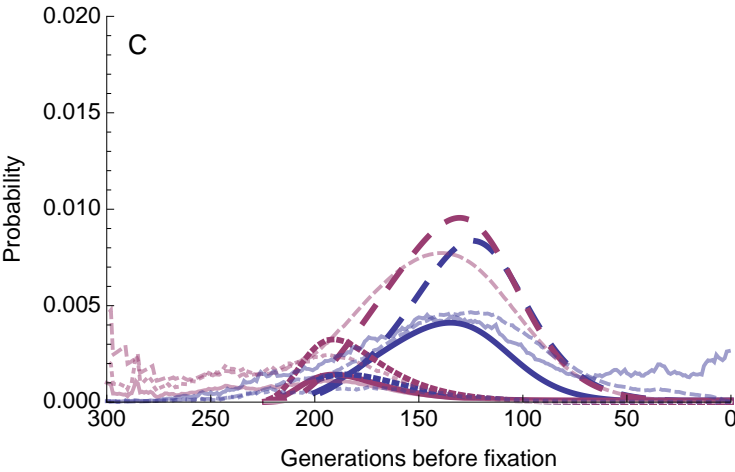


## The coalescent

```
In[583]:= letters = {"A", "B"}, {"C", "D"};
us = {10.-4, 10.-5};
ss = {0.2, 0.13};
GraphicsGrid[
  Table[
    Table[
      plotCoalescentSims[104, 0.05, ss[[i]], 0.5,
        0, us[[j]], 0, 2, 0.01, 300, 0.02, letters[[i, j]], 1, 100],
      {j, 2}],
    {i, 2}],
  Spacings → {0, 0}
]
```



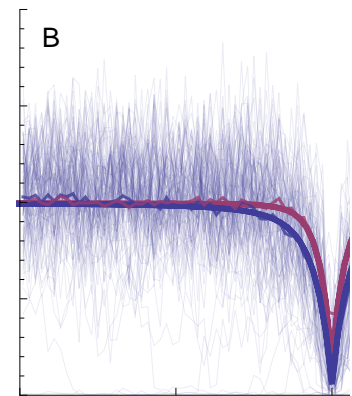
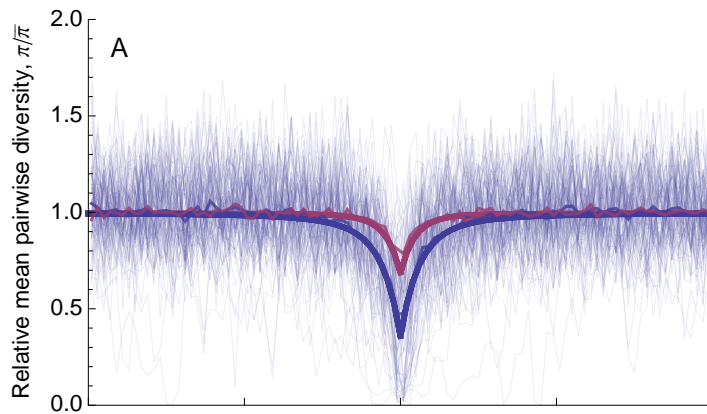
Out[586]=



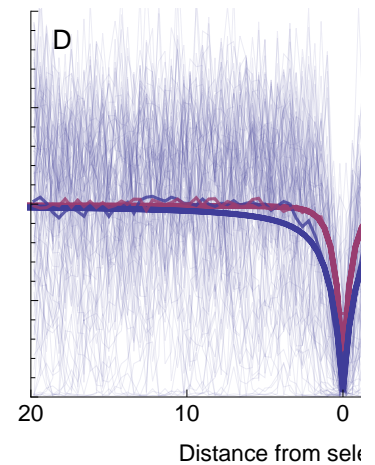
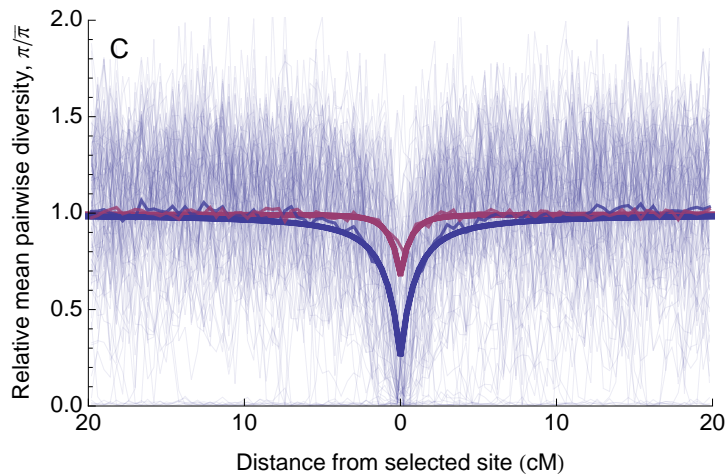


## Pairwise diversity

```
In[598]:= letters = {"A", "B"}, {"C", "D"};
us = {10.-4, 10.-5};
ss = {0.2, 0.13};
GraphicsGrid[
  Table[
    Table[
      plotDiversityRelative[104, 0.05,
        ss[[i]], 0.5, 0, us[[j]], 0, 100, letters[[i, j]], 1, 101],
      {j, 2}],
    {i, 2}],
  Spacings → {0, 0}
]
```



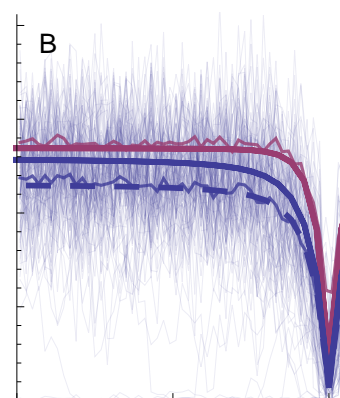
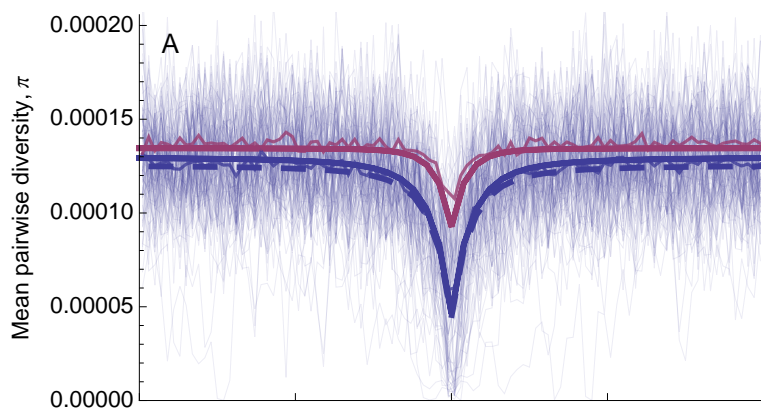
Out[601]=



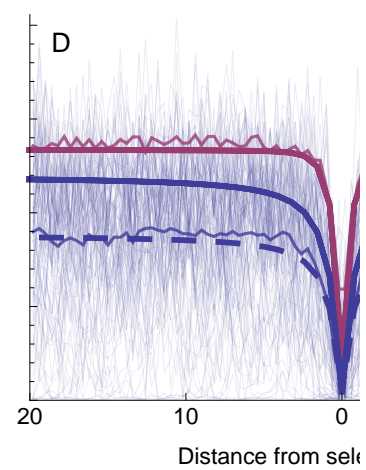
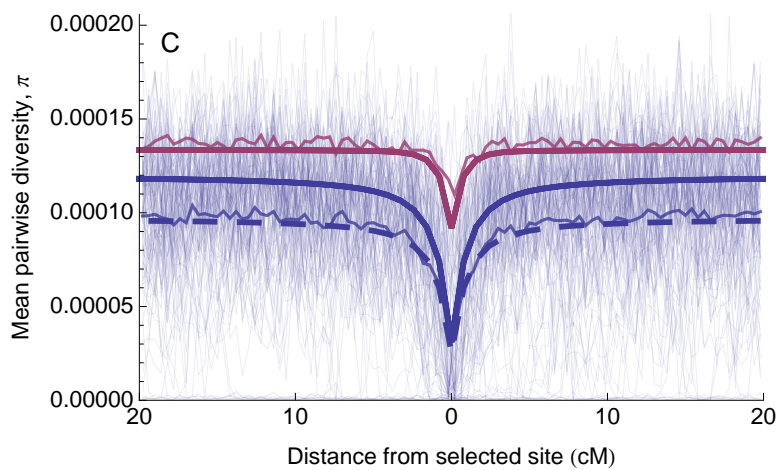
```

In[614]:= letters = {"A", "B"}, {"C", "D"};
us = {10.-4, 10.-5};
ss = {0.2, 0.13};
GraphicsGrid[
  Table[
    Table[
      plotDiversity[104, 0.05, ss[[i]], 0.5, 0, us[[j]], 0, 100, letters[[i, j]], 0, 51],
      {j, 2}],
    {i, 2}],
  Spacings → {0, 0}
]

```



Out[617]=

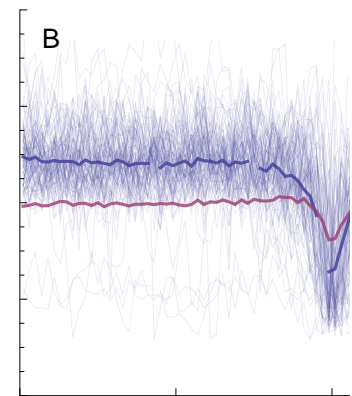
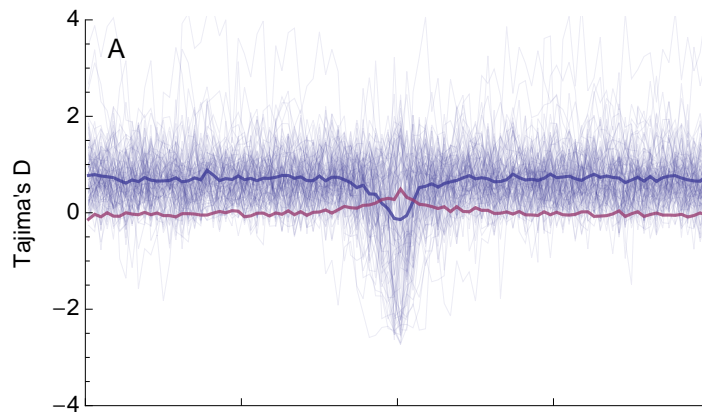


## Tajima's D

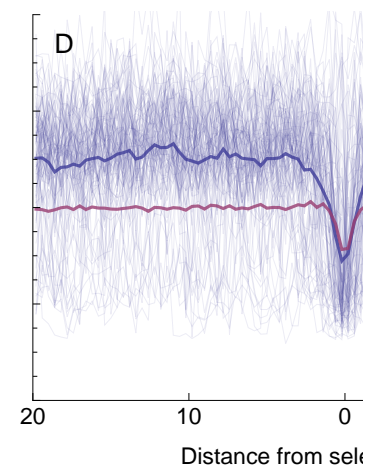
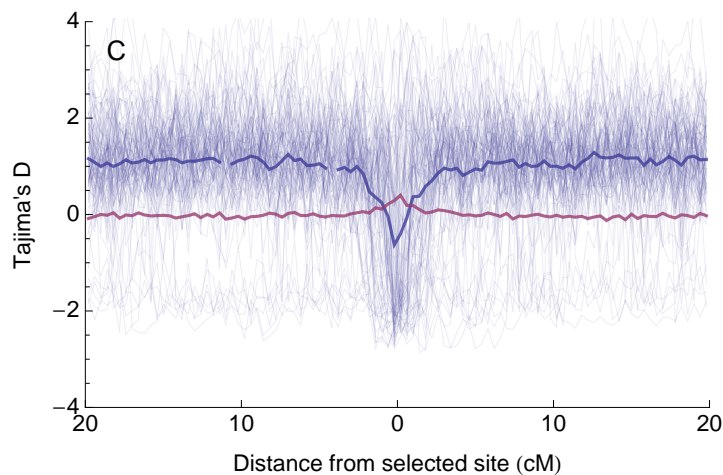
```

In[618]:= letters = {"A", "B"}, {"C", "D"};
us = {10.-4, 10.-5};
ss = {0.2, 0.13};
GraphicsGrid[
  Table[
    Table[
      plotTajimasD[104, 0.05, ss[[i]], 0.5, 0, us[[j]], 0, 100, letters[[i, j]], 1],
      {j, 2}],
    {i, 2}],
  Spacings → {0, 0}
]

```



Out[621]=



## Rescue from migration

### The probability of rescue and soft selective sweeps

If we assume beneficial alleles arrive at constant rate  $m$  then the probability a successful one has not

arrived by time  $t$  is

```
In[622]:= PDF[PoissonDistribution[m ρ t], X] /. X → 0
```

```
Out[622]:= e-m t ρ
```

The population is expected to go extinct in

```
In[623]:= Solve[1 == N0 Exp[-d t], t] /. C[1] → 0 // Flatten
```

```
Out[623]:= {t →  $\frac{\text{Log}[N0]}{d}$ }
```

generations and thus the probability of rescue by a migrant allele is roughly the probability at least one has established by then

```
In[624]:= Solve[1 == N0 Exp[-d t], t] /. C[1] → 0 // Flatten;
PMIG = 1 - PDF[PoissonDistribution[m ρ t], X] /. X → 0 /. %
```

```
Out[625]:= 1 - N0- $\frac{m \rho}{d}$ 
```

Check with simulations

```
In[626]:= predictPrescue[sval_, dval_, hval_, Nval_] :=
(
  params = {s → sval, d → dval, h → hval, B → 2, N0 → Nval};
  p = PMIG /. ρ → pest /. v → w (3 + 4 B - 4 w) / 4 /. w → (1 + s h) (1 - d) /. params;
  p
)
```

```
getPrescue[dval_, sval_, ms_] :=
(
  folder = StringForm["nestablish_MIG_d`_s`"],
  NumberForm[dval, {3, 2}], NumberForm[sval, {3, 2}]]];
  Table[data = Import[datadir <> ToString[folder] <> "/data/nestablish_m" <>
    ToString[NumberForm[m, {4, 3}]] <> ".txt", "Table"] // Flatten;
    {m, 100 / Length[data] // N},
    {m, ms}
  ]
)
```

```
plotPrescueMIG[svals_, dval_, Nval_, hval_, ms1_, ms2_] :=
(
```

```
  (*rescue theory*)
  theory = predictPrescue[svals[[1]], dval, hval, Nval];
  (*rescue theory*)
  theory2 = predictPrescue[svals[[2]], dval, hval, Nval];
```

```
  (*rescue simulations*)
  Prescue = getPrescue[dval, svals[[1]], ms1];
  (*rescue simulations*)
  Prescue2 = getPrescue[dval, svals[[2]], ms2];
```

```
  Show[
    ListLogLogPlot[Prescue,
```

```

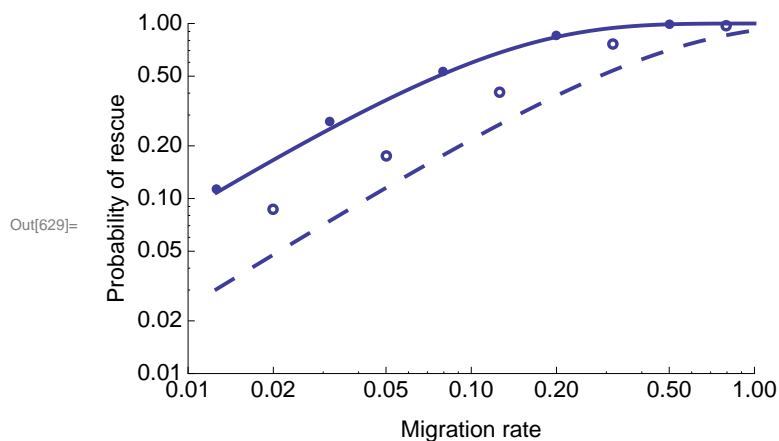
PlotStyle → AbsolutePointSize[5], PlotRange → {{10-2, 100}, {10-2, 1}},
LogLogPlot[theory, {m, ms1[[1]], 1}, PlotStyle → Thick],
ListLogLogPlot[Prescue2,
  PlotMarkers → Graphics[{Thickness[0.4], Circle[]}, ImageSize → 6]],
LogLogPlot[theory2, {m, ms1[[1]], 1},
  PlotStyle → Directive[Thick, Dashing[Large]]],
PlotRange → Log@{{10-2, 100}, {10-2, 1}},
Frame → {True, True, False, False},
FrameLabel → {"Migration rate"},
LabelStyle → labelstyle,
ImagePadding → {{50, 15}, {40, 10}},
Epilog → {
  Rotate[Text[
    Style["Probability of rescue", labelstyle], Scaled@{-0.125, 0.5}],  $\pi/2$ ]
],
PlotRangeClipping → False,
PlotRangePadding → None
]
)

```

```

plotPrescueMIG[{0.2, 0.13}, 0.05, 104, 0.5,
  Table[10i/10, {i, -19, 0, 4}] // N, Table[10i/10, {i, -17, 0, 4}] // N]

```



Note that, if  $ne/n$  is constant, so is the ratio of migration to coalescence

$$\begin{aligned}
 \text{In[630]:= } & \frac{\text{pmig}[k, \tau]}{\text{pcoal}[k, \tau] /. n[\tau] \rightarrow ne[\tau]} /. ne[\tau] \rightarrow NeN n[\tau] \\
 \text{Out[630]= } & \frac{2 m NeN}{-1 + k}
 \end{aligned}$$

We can therefore use Ewen's sampling formula to get the distribution of migrant haplotypes among a sample, replacing  $\theta$  with

```
In[631]:= Solve[ $\left(2 \text{ne}[\tau] \text{x}[\tau] \text{pmut}[k, \tau] /. u \rightarrow \frac{\theta}{4 \text{ne}[\tau]} /. \text{x}[\tau] \rightarrow 0\right) ==$   

 $(2 \text{ne}[\tau] \text{x}[\tau] \text{pmig}[k, \tau] /. n[\tau] \rightarrow \text{ne}[\tau] / \text{NeN}), \theta]$  // Flatten
```

```
Out[631]:=  $\{\theta \rightarrow 2 \text{m NeN}\}$ 
```

For example, the expected number of migrant haplotypes found in a sample of size  $n$  is (Ewens 2004 p 336)

```
In[632]:= nMIG = Sum[ $\frac{\theta}{j - 1 + \theta}$ , {j, k}]
```

```
Out[632]:=  $-\theta \text{PolyGamma}[0, \theta] + \theta \text{PolyGamma}[0, k + \theta]$ 
```

Try that out with simulation data

```
In[633]:= predictNest[nval_] :=  

(  

  params = {k → nval,  $\theta \rightarrow 2 \text{m NeN}$ };  

  nest = nMIG /. params /. NeN → 4 / 7;  

  nest  

)  

getNest[dval_, sval_, ms_] :=  

(  

  folder = StringForm["nestablish_MIG_d`_s`",  

    NumberForm[dval, {3, 2}], NumberForm[sval, {3, 2}]];  

  Table[data = Import[datadir <> ToString[folder] <> "/data/nestablish_m" <>  

    ToString[NumberForm[m, {4, 3}]] <> ".txt", "Table"] // Flatten;  

    data = Select[data, # > 0 &];  

    {m, Mean[data], StandardDeviation[data] /  $\sqrt{\text{Length}[data]}$ },  

    {m, ms}  

  ]  

)  

plotNestablishMIGrescue[svals_, dval_, Nval_, mss_] :=  

(  

  (*rescue theory*)  

  theory = predictNest[2 Nval];  

  (*constant simulations*)  

  Psweep = getNest[0, svals[[1]], mss[[1]]];  

  (*rescue simulations*)  

  Prescue = getNest[dval, svals[[1]], mss[[2]]];  

  (*constant simulations*)  

  Psweep2 = getNest[0, svals[[2]], mss[[3]]];  

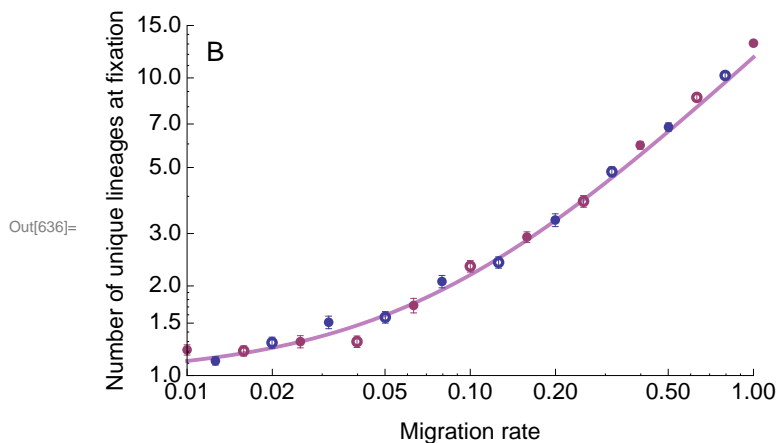
  (*rescue simulations*)  

  Prescue2 = getNest[dval, svals[[2]], mss[[4]]];  


```

```
Show[
  LogLogPlot[theory, {m, 0.01, 1}, PlotStyle →
    Directive[Purple, Thick, Opacity[0.5]], PlotRange → {{10-2, 1}, {1, 15}}
  ],
  ErrorListPlot[Prescue, PlotStyle → AbsolutePointSize[5]] /.
    {x_Real, y_Real} → {Log@x, Log@y},
  ErrorListPlot[Psweep, PlotStyle → Directive[AbsolutePointSize[5],
    defaultcolors[[2]]] /. {x_Real, y_Real} → {Log@x, Log@y},
  ErrorListPlot[Prescue2, PlotMarkers → Graphics[{Thickness[0.4], Circle[]},
    ImageSize → 6]] /. {x_Real, y_Real} → {Log@x, Log@y},
  ErrorListPlot[Psweep2, PlotStyle → defaultcolors[[2]], PlotMarkers →
    Graphics[{defaultcolors[[2]], Thickness[0.4], Circle[]}, ImageSize → 6]] /.
    {x_Real, y_Real} → {Log@x, Log@y},
  (*PlotRange→Log@{{10-7, 10-2}, {1, 100}}, *)
  Frame → {True, True, False, False},
  FrameLabel → {"Migration rate"},
  LabelStyle → labelstyle,
  ImagePadding → {{50, 15}, {40, 10}},
  Epilog → {
    Text[Style["B", letterstyle], Scaled@letterposition],
    Rotate[Text[Style["Number of unique lineages at fixation", labelstyle],
      Scaled@{-0.125, 0.5}],  $\pi/2$ ]
  },
  PlotRangeClipping → False,
  PlotRangePadding → None
]
```

```
plotNestablishMIGrescue[{0.2, 0.13}, 0.05, 104,
  {Table[10i/10, {i, -20, 0, 4}], Table[10i/10, {i, -19, 0, 4}],
  Table[10i/10, {i, -18, 0, 4}], Table[10i/10, {i, -17, 0, 4}]} // N]
Export[imagedir <> "NumberEstMIG.pdf", %];
```



And the probability of a soft sweep is

```
In[638]:= psoftMIG = 1 - Product[ $\frac{j}{j + \theta}$ , {j, k - 1}]
```

```
Out[638]:= 1 -  $\frac{(-1 + k)!}{\text{Pochhammer}[1 + \theta, -1 + k]}$ 
```

and try that with simulation data

```
In[639]:= predictPsoft[nval_] :=
(
  params = {k → nval, θ → 2 m NeN};
  p = psoftMIG /. params /. NeN → 4 / 7;
  p // N
)

getPsoft[dval_, sval_, ms_] :=
(
  folder = StringForm["nestablish_MIG_d`_s`"],
  NumberForm[dval, {3, 2}], NumberForm[sval, {3, 2}];
  Table[data = Import[datadir <> ToString[folder] <> "/data/nestablish_m" <>
    ToString[NumberForm[m, {4, 3}]] <> ".txt", "Table"] // Flatten;
    data = Length[Select[data, # > 1 &]];
    {m, data / 100},
    {m, ms}
  ]
)

plotPsoftMIG[Nval_, dval_, svals_, mss_] :=
(
  (*theory*)
  theory = predictPsoft[2 Nval];

  (*constant simulations*)
  Psweep = getPsoft[0, svals[[1]], mss[[1]]];
  (*rescue simulations*)
  Prescue = getPsoft[dval, svals[[1]], mss[[2]]];
  (*constant simulations*)
  Psweep2 = getPsoft[0, svals[[2]], mss[[3]]];
  (*rescue simulations*)
  Prescue2 = getPsoft[dval, svals[[2]], mss[[4]]];

  Show[
    LogLinearPlot[{, , theory}, {m, 0.01, 1},
      PlotStyle → Directive[Purple, Thick, Opacity[0.5]], PlotRange → {0, 1}],
    ListLogLinearPlot[Prescue, PlotStyle → AbsolutePointSize[5],
      PlotRange → {0, 1}],
    ListLogLinearPlot[Psweep, PlotStyle →
      Directive[AbsolutePointSize[5], defaultcolors[[2]]],
    ListLogLinearPlot[Prescue2, PlotMarkers →
      Graphics[{Thickness[0.4], Circle[]}, ImageSize → 6]],
    ListLogLinearPlot[Psweep2, PlotMarkers →
      Graphics[{defaultcolors[[2]], Thickness[0.4], Circle[]}, ImageSize → 6]],
```



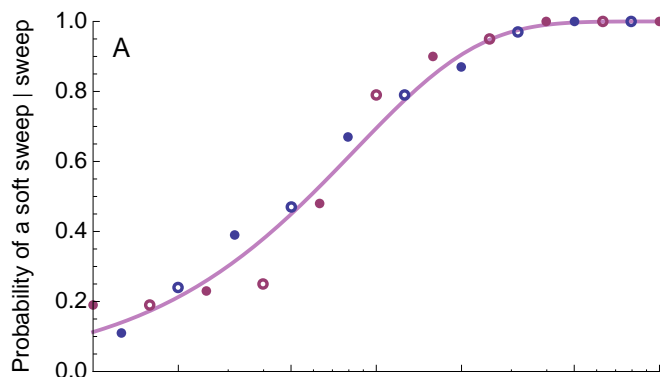
```

(*PlotRange→{Log@{10-2,1},{0,1}},*)
Frame→{True,True,False,False},
FrameTicksStyle→{FontColor→White, Automatic, Automatic, Automatic},
LabelStyle→labelstyle,
ImagePadding→{{50,15},{40,10}},
PlotRange→All,
Epilog→{
  Text[Style["A", letterstyle], Scaled@letterposition],
  Rotate[Text[Style["Probability of a soft sweep | sweep", labelstyle],
    Scaled@{-0.125, 0.5}],  $\pi/2$ ]
},
PlotRangeClipping→False,
PlotRangePadding→None
]
)

plotPsoftMIG[104, 0.05, {0.2, 0.13},
  {Table[10i/10, {i, -20, 0, 4}], Table[10i/10, {i, -19, 0, 4}],
    Table[10i/10, {i, -18, 0, 4}], Table[10i/10, {i, -17, 0, 4}]} // N]
Export[imagedir <> "PsoftMIG.pdf", %];

```

Out[642]=



## Effective initial allele frequencies and the backward-time dynamics

The PDF of waiting times until the first successful migrant allele arrives, given one arrives, is a truncated exponential

```

In[644]:= 1 - PDF[PoissonDistribution[m ρ t], X] /. X → 0;
Ft = D[%, t] / PMIG
Simplify[% ==  $\frac{\text{PDF}[\text{ExponentialDistribution}[m \rho], t]}{\text{PMIG}}, t > 0]$ 
Integrate[%, {t, 0,  $\frac{\text{Log}[N0]}{d}}$ ] == 1

Out[645]=  $\frac{e^{-m t \rho} m \rho}{1 - N0^{-\frac{m \rho}{d}}}$ 

Out[646]= True

Out[647]= True

```

So the expected number of copies of the beneficial allele at time t, while rare, is

```

In[648]:= Solve[1 == N0 Exp[-d t], t] /. C[1] → 0 // Flatten;
Integrate[ $\frac{\text{Exp}[\epsilon (t - \tau)]}{\rho} (\text{Ft} /. t \rightarrow \tau), \{\tau, 0, t /. \%\}$ ]

Out[649]=  $\frac{e^{t \epsilon} m \left(1 - N0^{-\frac{\epsilon + m \rho}{d}}\right)}{\left(1 - N0^{-\frac{m \rho}{d}}\right) (\epsilon + m \rho)}$ 

```

so that it is as if the initial frequency was

```

In[650]:= Solve[1 == N0 Exp[-d t], t] /. C[1] → 0 // Flatten;
Integrate[ $\frac{\text{Exp}[\epsilon (t - \tau)]}{\rho} (\text{Ft} /. t \rightarrow \tau), \{\tau, 0, t /. \%\}$ ];
Simplify[ $\frac{1}{2 N0} \%$ ] /. t → 0;

% ==  $\frac{1}{2 N0} \frac{1}{\rho} \frac{m \rho \left(1 - N0^{-\frac{\epsilon + m \rho}{d}}\right)}{(\epsilon + m \rho) \left(1 - N0^{-\frac{m \rho}{d}}\right)}$  // Simplify

q0rescueMIG =  $\frac{1}{2 N0} \frac{1}{\rho} \frac{m \rho \left(1 - N0^{-\frac{\epsilon + m \rho}{d}}\right)}{(\epsilon + m \rho) \left(1 - N0^{-\frac{m \rho}{d}}\right)}$ ;

```

Out[653]= True

When  $\frac{m \rho}{d}$  is small this is nearly independent of m (as in the case of rescue by DNM)

```

In[655]:= Series[q0rescueMIG /. m → mpoverd d / ρ, {mpoverd, 0, 0}] // Normal // Simplify

Out[655]=  $\frac{d \left(1 - N0^{-\frac{\epsilon}{d}}\right)}{2 N0 \epsilon \rho \text{Log}[N0]}$ 

```

In a population of constant size the waiting time is a simple exponential so that waiting time factor is

```
In[656]:= Simplify[PDF[ExponentialDistribution[m ρ], τ], τ ≥ 0];
Simplify[Integrate[Exp[ε (t - τ)] %, {τ, 0, ∞}], {m > 0, ρ > 0, ε > 0}];
% /. t → 0
```

$$\text{Out[658]} = \frac{m \rho}{\epsilon + m \rho}$$

so it is as if the initial frequency was

```
In[659]:= Simplify[PDF[ExponentialDistribution[m ρ], τ], τ ≥ 0];
Simplify[Integrate[Exp[ε (t - τ)] %, {τ, 0, ∞}], {m > 0, ρ > 0, ε > 0}];
% /. t → 0;
```

$$\frac{1}{2 N_0} \frac{1}{\rho} \%$$

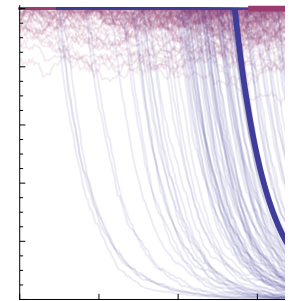
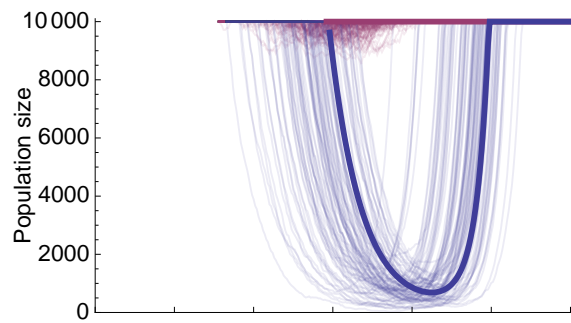
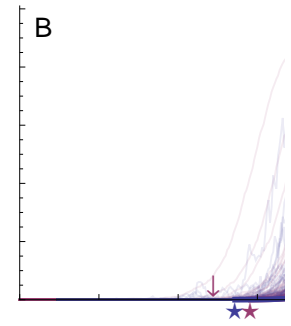
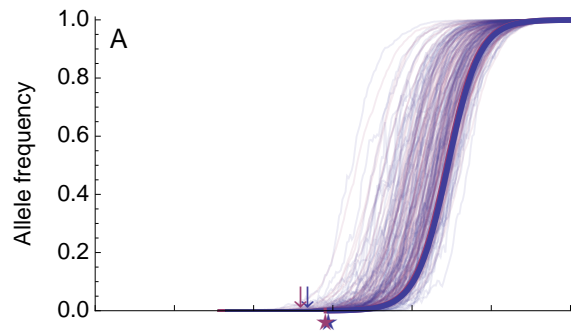
$$\text{Out[662]} = \frac{m}{2 N_0 (\epsilon + m \rho)}$$

But as in the case of de novo mutation, we do not really care about the timing of the sweep in a constant population size, just the effective initial allele frequency at the time this sweep begins

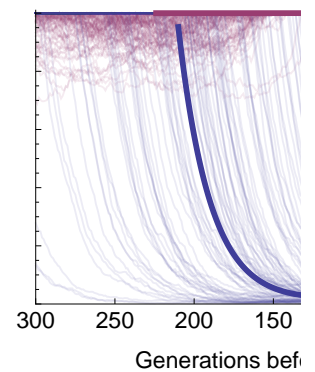
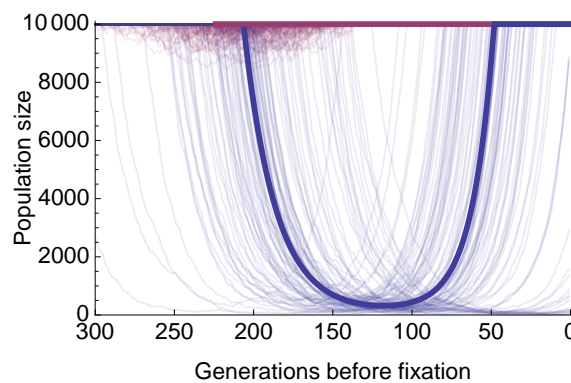
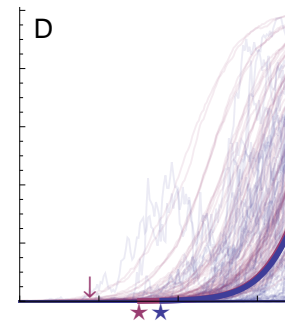
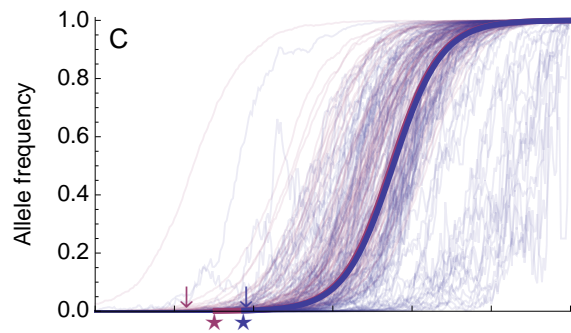
$$\text{In[663]} := q_{0\text{sweepMIG}} = \frac{1}{2 N_0 \rho};$$

Plot some dynamics

```
In[664]:= letters = {"A", "B"}, {"C", "D"};
ms = {1, 10.^-1};
ss = {0.2, 0.13};
GraphicsGrid[
  Table[
    Table[
      plotDynamicsBackwards[10^4, 0.05, ss[[i]],
        0.5, 0, 0, ms[[j]], 300, letters[[i, j]], 1, 100, Median],
      {j, Length[ms]}],
    {i, Length[ss]}],
  Spacings → {0, 0}
]
```



Out[667]=



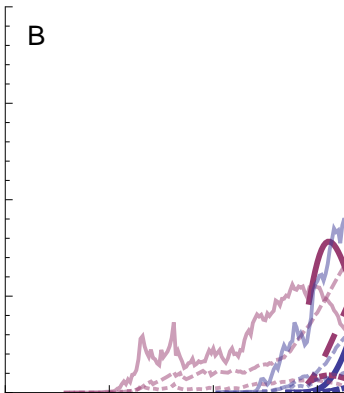
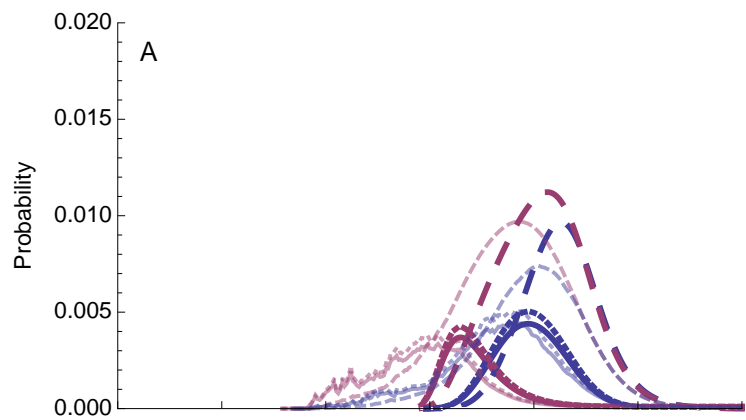
## The coalescent

Plot the backwards time dynamics and the probability of the possible events

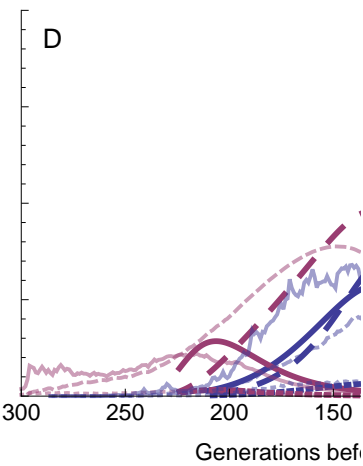
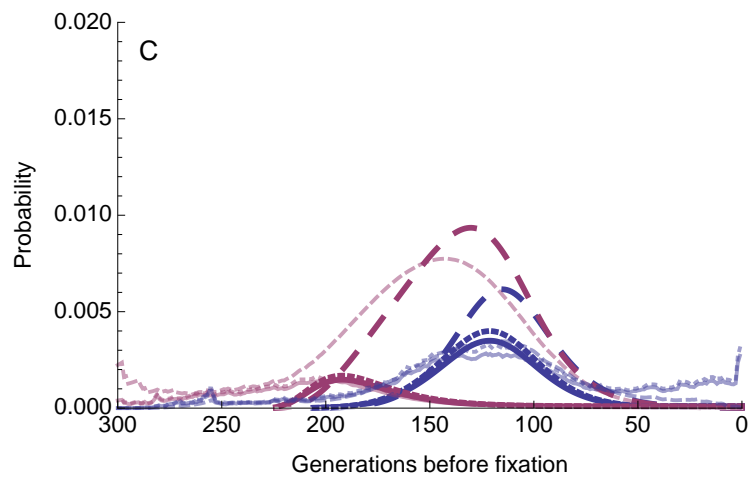
```

In[668]:= letters = {"A", "B"}, {"C", "D"};
ms = {1, 10.-1};
ss = {0.2, 0.13};
GraphicsGrid[
  Table[
    Table[
      plotCoalescentSims[104, 0.05, ss[[i]], 0.5,
        0, 0, ms[[j]], 2, 0.01, 300, 0.02, letters[[i, j]], 1, 100],
      {j, Length[ms]}],
    {i, Length[ss]}],
  Spacings → {0, 0}
]

```



Out[671]=



# Discussion

## Minimum HIV population size

Equation 22 in Orr & Unckless 2014 PLoS Genetics gives the minimum population size under rescue by DNM as

$$\text{In}[672]:= \text{nmin}[s\_ , d\_ , n0\_ ] := \frac{n0 s}{s - d} (2 n0 s)^{-d/s};$$

where  $d$  is the decline rate of the wildtype,  $n0$  is the initial number of wildtypes, and  $s$  is the selective advantage of the mutant.

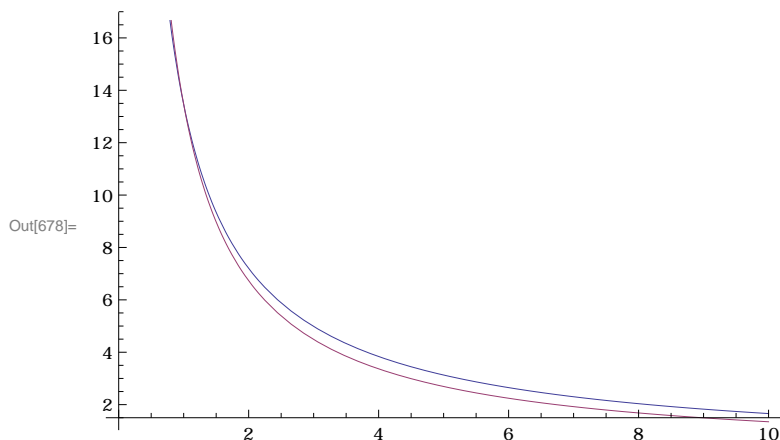
Thus for a given  $s$  and  $n0$  the minimum population size possible is  $\frac{e \text{Log}[2 n0 s]}{2 s}$

$$\begin{aligned} \text{In}[673]:= & \frac{n0 s}{s - d} (2 n0 s)^{-d/s}; \\ & \text{Solve}[D[\%, d] == 0, d] // \text{FullSimplify}; \\ & \%\ /. \% // \text{FullSimplify}; \\ & \%[[1]] == \frac{e \text{Log}[2 n0 s]}{2 s} // \text{Simplify} \end{aligned}$$

Out[676]= True

which, because  $\text{Log}[s]$  changes much slower than  $1/s$ , is nearly proportional to  $1/s$

$$\begin{aligned} \text{In}[677]:= & \left\{ \frac{e \text{Log}[2 n0 s]}{2 s}, \frac{e \text{Log}[2 n0 1]}{2 s} \right\} /. n0 \rightarrow 10^4; \\ & \text{Plot}[\%, \{s, 0.01, 10\}] \end{aligned}$$



The minimum population size for  $s=0.05$  and  $n0 = 10^4$  (Harris et al 2018 fig 3A) that is compatible with this model of rescue is

$$\text{In}[679]:= \frac{(2 n0 s)^{\frac{1}{\text{Log}[2 n0 s]}} \text{Log}[2 n0 s]}{2 s} /. n0 \rightarrow 10^4 /. s \rightarrow \{0.05\}$$

Out[679]= {187.772}