Computer simulations
The genomes of recombinant inbred lines

Karl Broman
Biostatistics & Medical Informatics, UW–Madison

kbroman.org
github.com/kbroman
@kwbroman
Course web: kbroman.org/AdvData
Intercross

\[ P_1 \times P_2 \rightarrow F_1 \times F_1 \rightarrow F_2 \]
Congenic line
Advanced intercross lines

P
A
B
F₂
F₃
F₄
F₇
F₁₀
Recombinant inbred lines

\[ P_1 \times P_2 \rightarrow F_1 \]

\[ F_1 \rightarrow F_2 \times F_1 \]

\[ F_2 \rightarrow F_3 \times F_2 \]

\[ F_3 \rightarrow F_4 \times F_3 \]

\[ \vdots \]

\[ F_\infty \]
Recombinant inbred lines

\[ P_1 \quad \vdash \quad \times \quad \vdash \quad P_2 \]

\[ F_1 \]

\[ F_2 \]

\[ F_3 \]

\[ F_4 \]

\[ \vdots \]

\[ F_\infty \]
Collaborative Cross

G₀

A × B

G₁

A B ×

G₂

A B C D

G₃

G₄

⋮

Gₙ
MAGIC

\[ G_0 \rightarrow A \xrightarrow{\times} B \]

\[ G_1 \rightarrow A \xrightarrow{\times} B \]

\[ G_2 \rightarrow ABCD \]

\[ G_3 \rightarrow \]

\[ G_4 \rightarrow \]

\[ \ldots \]

\[ G_\infty \rightarrow \]
Heterogeneous stock
Collaborative Cross

$G_0$

A ↓ B

$G_1$

A ↓ B

C × D

$G_2$

ABCD

$G_3$


$G_4$


...$

G_{\infty}$
CC genome
CC genome
Recombination fraction

\[
\begin{align*}
\text{A} & \quad \text{B} \\
\text{A} & \quad \text{B} \\
\text{A} & \quad \text{B} \\
(1-r)/2 & \quad (1-r)/2 & r/2 & \quad r/2
\end{align*}
\]

\( r \) is the "recombination fraction"
Simulation results

![Graph showing the relationship between recombination fraction and Pr (recombination in RIL). The graph plotting points indicates a nearly linear increase from 0.00 to 0.50 on the x-axis (recombination fraction) to 0.0 to 0.8 on the y-axis (Pr). The x-axis is labeled 'recombination fraction' and the y-axis is labeled 'Pr (recombination in RIL).']
INBREEDING AND LINKAGE*

J. B. S. HALDANE AND C. H. WADDINGTON

John Innes Horticultural Institution, London, England

Received August 9, 1930

TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-fertilization</td>
<td>358</td>
</tr>
<tr>
<td>Brother-sister mating. Sex-linked genes</td>
<td>360</td>
</tr>
<tr>
<td>Brother-sister mating. Autosomal genes</td>
<td>364</td>
</tr>
<tr>
<td>Parent and offspring mating. Sex-linked genes</td>
<td>367</td>
</tr>
<tr>
<td>Parent and offspring mating. Autosomal genes</td>
<td>368</td>
</tr>
<tr>
<td>Inbreeding with any initial population</td>
<td>370</td>
</tr>
<tr>
<td>Double crossing over</td>
<td>372</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>373</td>
</tr>
<tr>
<td>SUMMARY</td>
<td>374</td>
</tr>
<tr>
<td>LITERATURE CITED</td>
<td>374</td>
</tr>
</tbody>
</table>

When a heterozygous population is self-fertilized or inbred the ultimate result (apart from effects of mutation) is complete homozygosis. The final proportions of the various genotypes are usually independent of the system of inbreeding adopted, although, as JENNINGS (1916) and others have shown, the speed at which equilibrium is approached is greater in the case of self-fertilization than of brother-sister mating, and so on.

If however the population be heterozygous for linked genes, the final proportions depend on the system of mating, for crossing over can only occur in double heterozygotes, and the proportion of double heterozygotes falls off at a different rate in different mating systems. JENNINGS (1917) stated that he "would find it a relief if someone else would deal thoroughly with the laborious problem of the effects of inbreeding on two pairs of linked factors." This is the object of the present paper.

ROBBINS (1918) solved the problem in the case of self-fertilization.

In what follows we employ a direct method to obtain the final proportions of the population. The rate of approach can be calculated, but this is a very laborious process, and always involves the irrational roots of quadratic, sometimes those of quartic or higher equations. In each case we shall suppose that the number of dominant and recessive genes of each type in the population is equal throughout the progress of the inbreeding.

This enormously simplifies the mathematics. Thus a system of 55 equations would be required to deal with two pairs of linked factors in a population of 100 individuals, but if the number of different genes in the population is equal, a system of 27 equations is sufficient, and it can be seen from the tables that with this simplification the labor of calculation is not so great as it might appear from the number of equations.

MENDEL MEMORIAL FUND.

GENETICS 16: 357 J1 1931

* Part of the cost of the mathematical composition in this article is paid by the GALTON AND
Then \( c_n + \lambda d_n = c_n + \frac{1}{4}(1 - 2x)d_n + \frac{1}{2}\lambda(1 - 2x)d_n \)

\[
\therefore \quad \lambda = \frac{1 - 2x}{2 + 4x}.
\]

Then since \( d_\infty = 0 \), and \( c_1 = 0, d_1 = 2, \)

\[
c_\infty = c_\infty + \lambda d_\infty = c_1 + \lambda d_1 = \frac{1 - 2x}{1 + 2x}.
\]

Put \( y = D_\infty \) (the final proportion of crossover zygotes)

\[
\therefore \quad C_\infty + D_\infty = 1, \quad C_\infty - D_\infty = c_\infty \quad \therefore \quad y = \frac{1}{2}(1 - c_\infty).
\]

\[
\therefore \quad y = \frac{2x}{1 + 2x}.
\] (1.3)
Omitting some rather tedious algebra, the solution of these equations is:

\[
\begin{align*}
\zeta &= \frac{q}{2 - 3q}, \\
\theta &= \frac{2q}{2 - 3q}, \\
\kappa &= \frac{1}{2 - 3q}, \\
\lambda &= \frac{1 - 2q}{2 - 3q}, \\
\mu &= \frac{1 - 2q}{2 - 3q}, \\
\nu &= \frac{2q}{2 - 3q}
\end{align*}
\]

as may easily be verified.

\[
\therefore \quad c_\infty = c_n + 2e_n + \frac{1}{1 + 6x} [(1 - 2x)(d_n + 2f_n + 2j_n + \frac{1}{2}k_n) + 2g_n + 4x(h_n + i_n)]
\]  

(3.4)

and \(y = \frac{1}{2}(1 - c_\infty)\).

In the case considered, \(d_0 = 1\), \(\therefore \quad c_\infty = \zeta d_0 = 1 - 2x/1 + 6x\). Hence the proportion of crossover zygotes, \(y = 4x/1 + 6x\) (3.5).
Result for sib-mating

Omitting some rather tedious algebra, the solution of these equations is:

\[ \xi = \frac{q}{2 - 3q}, \quad \theta = \frac{2q}{2 - 3q}, \quad \kappa = \frac{1}{2 - 3q}, \]
\[ \lambda = \frac{1 - 2q}{2 - 3q}, \quad \mu = \frac{1 - 2q}{2 - 3q}, \quad \nu = \frac{2q}{2 - 3q} \]

as may easily be verified.

\[ \therefore \ c_\infty = c_n + 2e_n + \frac{1}{1 + 6x} \left[ (1 - 2x)(d_n + 2f_n + 2j_n + \frac{1}{2}k_n) \right. \]
\[ \left. + 2g_n + 4x(h_n + i_n) \right] \]  \hspace{1cm} (3.4)

and \( y = \frac{1}{2} (1 - c_\infty) \).

In the case considered, \( d_0 = 1, \therefore c_\infty = \xi d_0 = 1 - 2x / 1 + 6x \). Hence the proportion of crossover zygotes, \( y = 4x / 1 + 6x \) \hspace{1cm} (3.5).
Simulation results

\[ R = \frac{a r}{1 + b r} \]
Non-linear regression

\[
\text{out <- nls}( R \sim a*r/(1 + b*r), \\
data = \text{data.frame}(r=r, R=R), \\
\text{start = list}(a=4, b=6))
\]

\text{summary(out)}
Non-linear regression

out <- nls(R ~ a*r/(1 + b*r),
           data = data.frame(r=r, R=R),
           start = list(a=4, b=6))

summary(out)

<table>
<thead>
<tr>
<th>Estimate</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>7.016</td>
</tr>
<tr>
<td></td>
<td>0.011</td>
</tr>
<tr>
<td>b</td>
<td>6.023</td>
</tr>
<tr>
<td></td>
<td>0.016</td>
</tr>
</tbody>
</table>
Non-linear regression

```r
out <- nls(R ~ a*r/(1 + b*r),
            data = data.frame(r=r, R=R),
            start = list(a=4, b=6))
summary(out)
```

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>7.016</td>
<td>0.011</td>
<td>a</td>
<td>7.003</td>
<td>0.008</td>
</tr>
<tr>
<td>b</td>
<td>6.023</td>
<td>0.016</td>
<td>b</td>
<td>6.005</td>
<td>0.012</td>
</tr>
</tbody>
</table>

More data
Simulation results

\[ R = \frac{7r}{1 + 6r} \]

Pr (recombination in RIL) vs. recombination fraction

- Points on the graph represent data points for various recombination fractions.
- The equation \( R = \frac{7r}{1 + 6r} \) describes the relationship between the recombination fraction and the probability of recombination in RIL (recombinant inbred lines).

The graph shows a positive correlation between recombination fraction and the probability of recombination, with the probability increasing as the recombination fraction increases.
Markov chain

- Sequence of random variables \( \{X_0, X_1, X_2, \ldots \} \) satisfying
  \[
  \Pr(X_{n+1} \mid X_0, X_1, \ldots, X_n) = \Pr(X_{n+1} \mid X_n)
  \]
- Transition probabilities \( P_{ij} = \Pr(X_{n+1} = j \mid X_n = i) \)
- Here, \( X_n = \) “parental type” at generation \( n \).
- We are interested in absorption probabilities
  \[
  \pi_j = \Pr(X_n \rightarrow j \mid X_0)
  \]
Absorption probabilities

Consider the case of absorption into the state $AA | AA$

Let $h_i$ = probability, starting at $i$, of being absorbed into $AA | AA$.

Then $h_{AA|AA} = 1$ and $h_{AB|AB} = 0$.

Condition on the first step: $h_i = \sum_k P_{ik} h_k$

For selfing, this gives a system of 3 linear equations.
Equations for selfing

\[ C_n AABB \text{ and } aabb. \]
\[ D_n AAbb \text{ and } aaBB. \]
\[ E_n AABb, AaBB, Aabb, \text{ and } aaBb. \]
\[ F_n AB.ab. \]
\[ G_n Ab.aB. \]

We assume \( 2C_n + 2D_n + 4E_n + F_n + G_n = 2 \), so that \( C_1 = D_1 = E_1 = G_1 = 0 \), and \( F_1 = 2 \). Clearly \( E_\infty = F_\infty = G_\infty = 0 \), and \( D_\infty \) is the final proportion of crossover zygotes. Then considering the results of selfing each generation, we have:

\[
\begin{align*}
C_{n+1} &= C_n + \frac{1}{4}E_n + \frac{1}{4}(1 - \beta - \delta + \beta \delta)F_n + \frac{1}{4}\beta \delta G_n \\
D_{n+1} &= D_n + \frac{1}{2}E_n + \frac{1}{2}\beta \delta F_n + \frac{1}{4}(1 - \beta - \delta + \beta \delta)G_n \\
E_{n+1} &= E_n + \frac{1}{4}(\beta + \delta - 2\beta \delta)(F_n + G_n) \\
F_{n+1} &= \frac{1}{2}(1 - \beta - \delta + \beta \delta)F_n + \frac{1}{2}\beta \delta G_n \\
G_{n+1} &= \frac{1}{2}\beta \delta F_n + \frac{1}{2}(1 - \beta - \delta + \beta \delta)G_n
\end{align*}
\] (1.1)

Put \( y = D_\infty \) (the final proportion of crossover zygotes)

\[
\therefore C_\infty + D_\infty = 1, \quad C_\infty - D_\infty = c_\infty \quad \therefore y = \frac{1}{2}(1 - c_\infty).
\] (1.2)

\[
\therefore y = \frac{2x}{1 + 2x}.
\] (1.3)
Equations for sib-mating

<table>
<thead>
<tr>
<th>Typical mating</th>
<th>Number of types</th>
</tr>
</thead>
<tbody>
<tr>
<td>AABB × AABB</td>
<td>2</td>
</tr>
<tr>
<td>AAbb × AAbb</td>
<td>2</td>
</tr>
<tr>
<td>AABB × aabb</td>
<td>2</td>
</tr>
<tr>
<td>AAbb × aABb</td>
<td>2</td>
</tr>
<tr>
<td>AABB × AAbb</td>
<td>8</td>
</tr>
<tr>
<td>AAbb × AAbb</td>
<td>8</td>
</tr>
<tr>
<td>AABB × AB, ab</td>
<td>4</td>
</tr>
<tr>
<td>AABB × Ab, ab</td>
<td>4</td>
</tr>
<tr>
<td>AABB × Ab, ab</td>
<td>8</td>
</tr>
</tbody>
</table>

Typical mating

- \( C_{AB} = C_n + H + \frac{1}{2}(a^2 + \gamma + \theta)N + (1 + R + 1) + (a^2 + \gamma + \theta) \)
- \( D_{AB} = D_n + \frac{1}{2}(a^2 + \gamma)M + (1 + R + 1) + (a^2 + \gamma) \)
- \( E_{AB} = E_n + \frac{1}{2}(a^2 + \gamma + \theta)N + (1 + R + 1) + (a^2 + \gamma + \theta) \)
- \( F_{AB} = F_n + \frac{1}{2}(a^2 + \gamma + \theta)N + (1 + R + 1) + (a^2 + \gamma + \theta) \)
- \( G_{AB} = G_n + \frac{1}{2}(a^2 + \gamma + \theta)(U + V) + \frac{1}{2}(a^2 + \gamma + \theta)(W + 2X + Y) \)
- \( H_{AB} = H_n + \frac{1}{2}(a^2 + \gamma)N + (1 + R + 1) + (a^2 + \gamma) \)

Number of types

- \( N_{AB} = \frac{1}{2}(a^2 + \gamma + \theta)(U + V) + \frac{1}{2}(a^2 + \gamma + \theta)(W + 2X + Y) \)
- \( P_{AB} = \frac{1}{2}(a^2 + \gamma + \theta)(U + V) + \frac{1}{2}(a^2 + \gamma + \theta)(W + 2X + Y) \)
- \( Q_{AB} = 2G + \frac{1}{2}(U + V + 1) + (1 + R + 1) + (a^2 + \gamma + \theta) + (1 + \beta + \gamma) \)
- \( R_{AB} = \frac{1}{2}(a^2 + \gamma + \theta)(U + V) + \frac{1}{2}(a^2 + \gamma + \theta)(W + 2X + Y) \)
- \( S_{AB} = 2U + \frac{1}{2}(a^2 + \gamma + \theta)(U + V) + \frac{1}{2}(a^2 + \gamma + \theta)(W + 2X + Y) \)
- \( T_{AB} = \frac{1}{2}(a^2 + \gamma + \theta)(U + V) + \frac{1}{2}(a^2 + \gamma + \theta)(W + 2X + Y) \)

where:
- \( a, \gamma, \theta \) are coefficients
Result for sib-mating

Omitting some rather tedious algebra, the solution of these equations is:

\[
\begin{align*}
\zeta &= \frac{q}{2 - 3q}, \quad \theta = \frac{2q}{2 - 3q}, \quad \kappa = \frac{1}{2 - 3q}, \\
\lambda &= \frac{1 - 2q}{2 - 3q}, \quad \mu = \frac{1 - 2q}{2 - 3q}, \quad \nu = \frac{2q}{2 - 3q}
\end{align*}
\]

as may easily be verified.

\[
\therefore \quad c_\infty = c_n + 2e_n + \frac{1}{1 + 6x} \left[ (1 - 2x)(d_n + 2f_n + 2j_n + \frac{1}{2}k_n) \\
+ 2g_n + 4x(h_n + i_n) \right]
\]

(3.4)

and \( y = \frac{1}{2}(1 - c_\infty) \).

In the case considered, \( d_0 = 1 \), \( c_\infty = \zeta d_0 = 1 - 2x/1 + 6x \). Hence the proportion of crossover zygotes, \( y = 4x/1 + 6x \) (3.5).
3-point coincidence

- $r_{ij}$ = recombination fraction for interval (i, j)
  Assume $r_{12} = r_{23} = r$.

- **Coincidence** = $c = \frac{\Pr(\text{double recombinant})}{r^2}$
  $= \frac{\Pr(\text{rec’n in 23} \mid \text{rec’n in 12})}{\Pr(\text{rec’n in 23})}$

- No interference = 1
  Positive interference $< 1$
  Negative interference $> 1$

- Generally $c$ is a function of $r$
Coincidence

![Graph showing Coincidence vs. r]

- **Meiosis**
- **RILs by selfing**
- **RILs by sib-mating**
Coincidence

- Meiosis
- RILs by selfing
- RILs by sib-mating
- No interference
- Mouse interference
Coincidence in 8-way RILs

- The trick that allowed us to get the coincidence for 2-way RILs doesn’t work for 8-way RILs.

- It’s sufficient to consider 4-way RILs.

- Calculations for 3 points in 4-way RILs is still **astoundingly complex**.
  - 2 points in 2-way RILs by sib mating:
    - 55 parental types $\rightarrow$ **22 states** by symmetry
  - 3 points in 4-way RILs by sib mating:
    - 2,164,240 parental types $\rightarrow$ **137,488 states** by symmetry

- Even **counting** the states was difficult.
Coincidence

- Meiosis
- 2−way RILs by selfing
- 2−way RILs by sib−mating
- 8−way RILs by sib−mating
- No interference
- Mouse interference
The formula

\[
C = \frac{(1 + 6r)[280 + 1208r - 848r^2 + 5c(7 - 28r - 368r^2 + 344r^3) - 2c^2(49 - 324r + 452r^2)r^2 - 16c^3(1 - 2r)r^4]}{49(1 + 12r - 12cr^2)[5 + 10r - 4(2 + c)r^2 + 8cr^3]}
\]
3-point symmetry

\[ \Pr(M_2 = x \mid M_1 = A, M_2 \neq A, M_3 = A) \]

![Graph showing recombination fraction with different symptoms: No interference, Positive interference, with lines for x = B, x = C, and x = E.](image)
Markov property

$$\log_2 \left\{ \frac{\Pr(M_3 = A \mid M_2 = A, M_1 = x)}{\Pr(M_3 = A \mid M_2 = A)} \right\}$$
Markov property

\[
\log_2 \left\{ \frac{\Pr(M_3 = A \mid M_2 = B, M_1 = x)}{\Pr(M_3 = A \mid M_2 = B)} \right\}
\]
Markov property

\[
\log_2 \left\{ \frac{\Pr(M_3 = A \mid M_2 = C, M_1 = x)}{\Pr(M_3 = A \mid M_2 = C)} \right\}
\]
Markov property

$$\log_2 \left\{ \frac{\Pr(M_3 = A \mid M_2 = E, M_1 = x)}{\Pr(M_3 = A \mid M_2 = E)} \right\}$$
Whole genome simulations

- 2-way selfing, 2-way sib-mating, 8-way sib-mating
- Mouse-like genome, 1665 cM
- Strong positive crossover interference
- Inbreed to complex fixation
- 10,000 simulation replicates
No. generations to fixation

- 2-way selfing: mean = 10.5
- 2-way sib-mating: mean = 35.6
- 8-way sib-mating: mean = 38.9
No. generations to 99% fixation

- 2-way selfing: mean = 8.0
- 2-way sib-mating: mean = 23.5
- 8-way sib-mating: mean = 26.7
Percent genome not fixed

- 2-way selfing
- 2-way sib-mating
- 8-way sib-mating
- Average
- 95th percentile

No. generations

Percent not yet fixed
Segment lengths

- 2-way selfing: median = 23.7 cM
- 2-way sib-mating: median = 12.9 cM
- 8-way sib-mating: median = 8.5 cM
Segment lengths

- 2-way selfing: median = 23.7 cM
- 2-way sib-mating: median = 12.9 cM
- 8-way sib-mating: median = 8.5 cM

Two chromosomes
- X chromosome

Segment lengths (cM)
Probability a segment is inherited intact

- 2-way selfing
- 2-way sib-mating
- 8-way sib-mating

![Graph showing the probability of segment inheritance intact over different lengths and mating systems.](image-url)
Length of smallest segment

95th %ile = 0.26 cM

95th %ile = 0.58 cM

95th %ile = 2.2 cM
No. segments < 1 cM

- 2-way selfing: mean = 1.4
- 2-way sib-mating: mean = 5.2
- 8-way sib-mating: mean = 11.2
Collaborative Cross

G\(_0\) A \(\times\) B

G\(_1\) A \(\downarrow\) B

G\(_2\) ABCD

G\(_3\)

G\(_4\)

\(\vdots\)

G\(_\infty\)
The PreCC

What happens at $G_2F_k$?

$\Pr(g_1 = i)$ as a function of $k$

$\Pr(g_1 = i, g_2 = j)$ as a function of $k$ and the recombination fraction
<table>
<thead>
<tr>
<th>Chr.</th>
<th>Individual</th>
<th>Prototype</th>
<th>No. states</th>
<th>Probability of each</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Random</td>
<td>AA</td>
<td>4</td>
<td>(\frac{1}{4(1+6r)}\left(\frac{6r^2-7r-3s}{4(1+6r)s}\right) + \frac{1}{4(1+6r)s}\left(1-2r+s\right))</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AB</td>
<td>4</td>
<td>(\frac{1}{2(1+6r)}\left(\frac{10r^2-r-3s}{4(1+6r)s}\right) + \frac{1}{4(1+6r)s}\left(1-2r+s\right))</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AC</td>
<td>8</td>
<td>(\frac{1}{r}\left(\frac{2r^2+3r+s}{4(1+6r)s}\right) + \frac{1}{4(1+6r)s}\left(1-2r+s\right))</td>
</tr>
<tr>
<td>X</td>
<td>Female</td>
<td>AA</td>
<td>2</td>
<td>(\frac{1}{3(1+4r)}\left(\frac{6r^2+3r+3s}{4(1+4r)s}\right) + \frac{1}{4(1+4r)s}\left(1-r+t\right))</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AB</td>
<td>2</td>
<td>(\frac{2r}{3(1+4r)}\left(\frac{1}{3(1+r)}\right) + \frac{1}{2(1+4r)s}\left(\frac{2r^3+6r^2+(2r^2+r)t}{4(1+4r)s}\right))</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AC</td>
<td>4</td>
<td>(\frac{2r}{3(1+4r)}\left(\frac{1}{2(1+r)}\right) + \frac{1}{4(1+4r)s}\left(\frac{9r^2+5r+rt}{2(1+4r)s}\right))</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>AA</td>
<td>2</td>
<td>(\frac{1}{3(1+4r)}\left(\frac{1}{3(1+r)}\right) + \frac{1}{4(1+4r)s}\left(\frac{r^2-(8r^2+r^2-3r)t-10r^2+5r}{2(1+4r)s}\right))</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AB</td>
<td>2</td>
<td>(\frac{2r}{3(1+4r)}\left(\frac{1}{3(1+r)}\right) + \frac{1}{4(1+4r)s}\left(\frac{r^2+(5r^2-r)t-10r^2+5r^2}{4(1+4r)s}\right))</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AC</td>
<td>4</td>
<td>(\frac{2r}{3(1+4r)}\left(\frac{1}{3(1+r)}\right) + \frac{1}{4(1+4r)s}\left(\frac{r^2+(5r^2-r)t-10r^2+5r^2}{4(1+4r)s}\right))</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CC</td>
<td>1</td>
<td>(\frac{1}{3(1+4r)}\left(\frac{1}{3(1+r)}\right) + \frac{1}{4(1+4r)s}\left(\frac{r^2+(5r^2-r)t-10r^2+5r^2}{4(1+4r)s}\right))</td>
</tr>
</tbody>
</table>

\(s = \sqrt{4r^2-12r+5}\) and \(r = \sqrt{r^2-10r+5}\); the autosomal haplotype probabilities are valid for \(|r| < \frac{1}{2}\).
Computer simulations are hugely valuable.
Uses of simulations

- Study probabilities
- Estimate power/sample size
- Evaluate performance of a method
- Evaluate sensitivity/robustness of a method
Relative advantages?

- Simulations
- Numerical calculations
- Analytic calculations
References

► Haldane & Waddington (1931) Inbreeding and Linkage. Genetics 16:357–374


