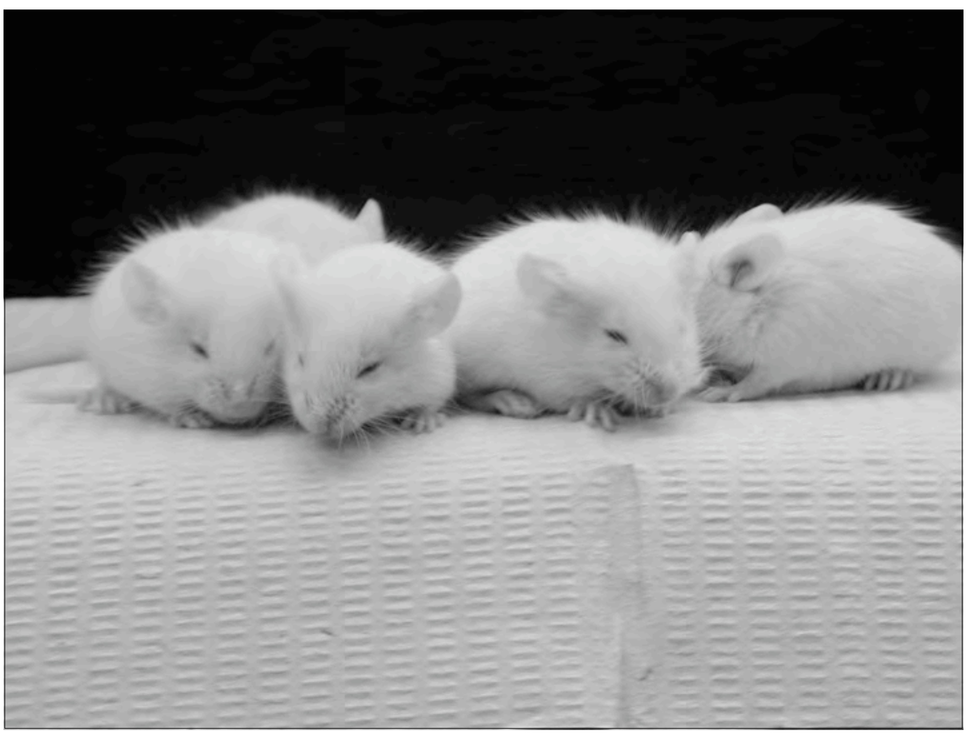
Super computing for “classical” genetics

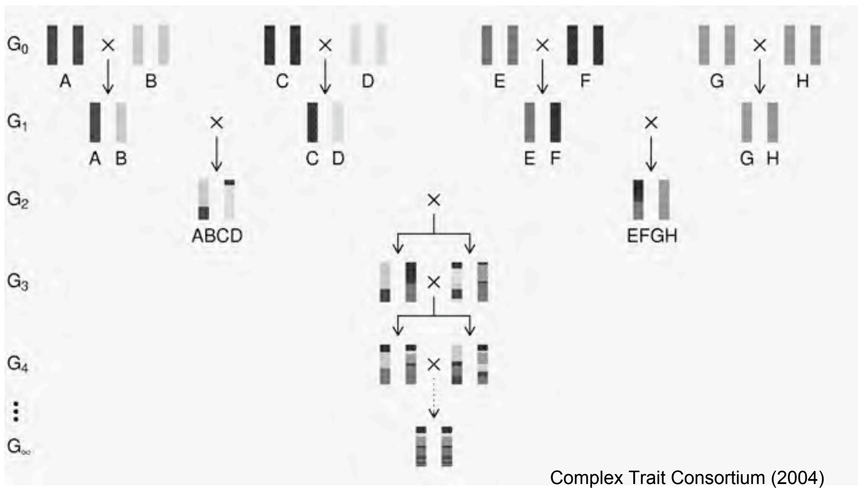
Karl W Broman

Department of Biostatistics
Johns Hopkins Bloomberg School of Public Health

http://www.biostat.jhsph.edu/~kbroman

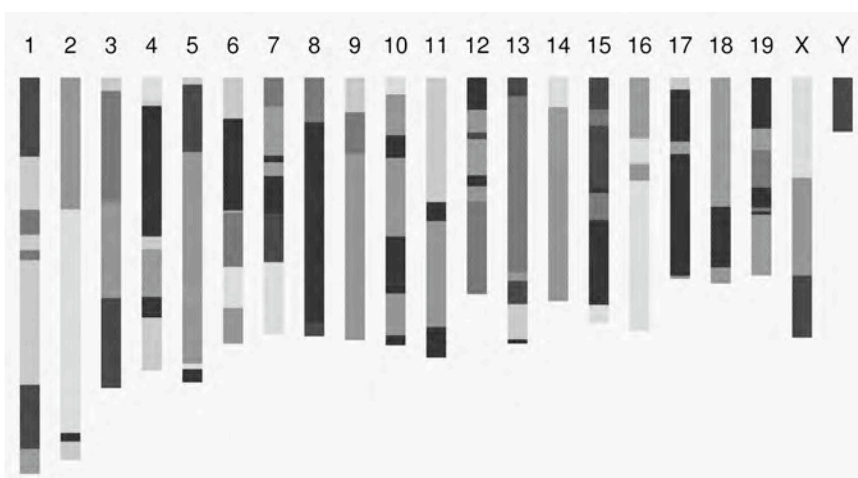


The Collaborative Cross



Complex Trait Consortium (2004)
Nat Genet 36:1133-1137

Genome of an 8-way RI



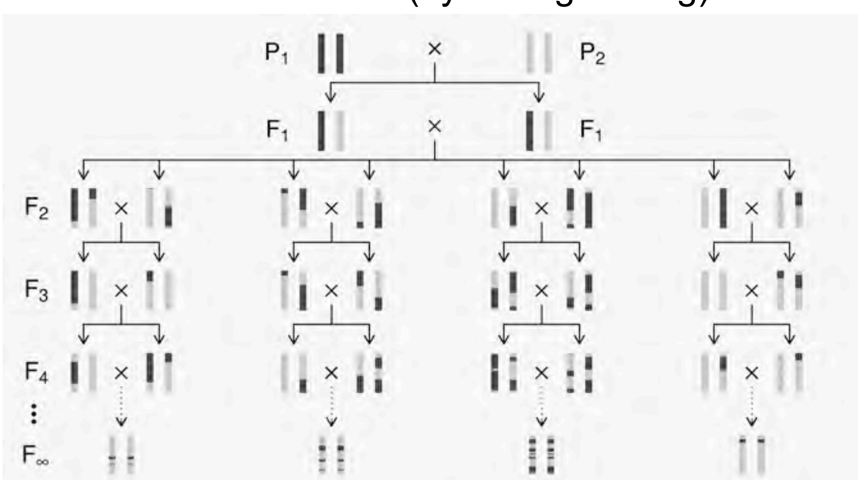
2 points in an RIL

- r = recombination fraction = probability of a recombination in the interval in a random meiotic product.
- R = analogous thing for the RIL = probability of different alleles at the two loci on a random RIL chromosome.

5

Recombinant inbred lines

(by sibling mating)



6

Haldane & Waddington 1931

INBREEDING AND LINKAGE*

J. B. S. HALDANE AND C. H. WADDINGTON
John Innes Horticultural Institution, London, England

Received August 9, 1930

Genetics 16:357-374

TABLE OF CONTENTS

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.

7

Equations for sib-mating

Typical mating	*Number of types*	
AABB×*AABB*	2	$C_{n+1}=C_n+H+\frac{1}{4}(\alpha^2+\gamma^2)L+\frac{1}{4}(\beta^2+\delta^2)N+\frac{1}{8}Q+\frac{1}{8}R+\frac{1}{8}(\alpha^2+\gamma^2)U+\frac{1}{8}(\beta^2+\delta^2)V+\frac{1}{16}\alpha^2\gamma^2W+\frac{1}{16}(\alpha^2\delta^2+\beta^2\gamma^2)X+\frac{1}{16}\beta^2\delta^2Y.$
AAbb×*AAbb*	2	$D_{n+1}=D+I+\frac{1}{4}(\alpha^2+\gamma^2)M+\frac{1}{4}(\beta^2+\delta^2)P+\frac{1}{8}Q+\frac{1}{8}S+\frac{1}{8}(\beta^2+\delta^2)U+\frac{1}{8}(\alpha^2+\gamma^2)V+\frac{1}{16}\beta^2\delta^2W+\frac{1}{16}(\alpha^2\delta^2+\beta^2\gamma^2)X+\frac{1}{16}\alpha^2\gamma^1Y.$
AABB×*aabb*	2	$E_{n+1}=\frac{1}{16}\alpha^2\gamma^2W+\frac{1}{16}(\alpha^2\delta^2+\beta^2\gamma^2)X+\frac{1}{16}\beta^2\delta^2Y.$
AAbb×*aaBB*	2	$F_{n+1}=\frac{1}{16}\beta^2\delta^2W+\frac{1}{16}(\alpha^2\delta^2+\beta^2\gamma^2)X+\frac{1}{16}\alpha^2\gamma^2Y.$
AABB×*AAbb*	8	$G_{n+1}=\frac{1}{16}(\alpha\beta+\gamma\delta)(U+V)+\frac{1}{16}\alpha\beta\gamma\delta(W+2X+Y).$
AABB×*AABb*	8	$H_{n+1}=\frac{1}{2}H$ … $U+\frac{1}{16}($ … $(\alpha\delta+\beta$ …
AAbb×*AABb*	8	$I_{n+1}=\frac{1}{2}I+$ … $U+\frac{1}{16}($ … $(\alpha\delta+\beta$ …
AABB×*Aabb*	8	$J_{n+1}=\frac{1}{16}($ … $\beta\delta)(\alpha\delta$ …
AAbb×*AaBB*	8	$K_{n+1}=\frac{1}{16}($ … $\beta\delta)(\alpha\delta$ …
AABB×*AB.ab*	4	$L_{n+1}=\frac{1}{4}(\alpha$ … $\alpha^2\gamma^2W$ …
AAbb×*Ab.aB*	4	$M_{n+1}=\frac{1}{4}($ … $\beta^2\delta^2W$ …

Typical mating	*Number of types*	
AABB×*Ab.aB*	4	$N_{n+1}=\frac{1}{8}R+\frac{1}{8}(\alpha\beta+\gamma\delta)(U+V)+\frac{1}{8}\alpha\beta\gamma\delta(W+2X+Y).$
AAbb×*AB.ab*	4	$P_{n+1}=\frac{1}{8}S+\frac{1}{8}(\alpha\beta+\gamma\delta)(U+V)+\frac{1}{8}\alpha\beta\gamma\delta(W+2X+Y).$
AABb×*AABb*	4	$Q_{n+1}=2G+\frac{1}{2}(H+I+J+K)+\frac{1}{4}(\alpha^2+\gamma^2)(L+M)+\frac{1}{4}(\beta^2+\delta^2)(N+P)+\frac{1}{4}Q+\frac{1}{8}(R+S+T)+\frac{1}{8}(\alpha^2+\alpha\beta+\beta^2+\gamma^2+\gamma\delta+\delta^2)(U+V)+\frac{1}{16}(\alpha\delta+\beta\gamma)^2(W+Y)+\frac{1}{8}(\alpha\gamma+\beta\delta)^2X.$
AABb×*AaBB*	4	$R_{n+1}=\frac{1}{4}(\beta^2+\delta^2)L+\frac{1}{4}(\alpha^2+\gamma^2)N+\frac{1}{8}R+\frac{1}{8}(\beta+\delta)U+\frac{1}{8}(\alpha+\gamma)V+\frac{1}{16}(\alpha\delta+\beta\gamma)^2(W+Y)+\frac{1}{8}(\alpha\gamma+\beta\delta)^2X.$
AABb×*Aabb*	4	$S_{n+1}=\frac{1}{4}(\beta^2+\delta^2)M+\frac{1}{4}(\alpha^2+\gamma^2)P+\frac{1}{8}S+\frac{1}{8}(\alpha+\gamma)U+\frac{1}{8}(\beta+\delta)V+\frac{1}{16}(\alpha\delta+\beta\gamma)^2(W+Y)+\frac{1}{8}(\alpha\gamma+\beta\delta)^2X.$
AABb×*aaBb*	4	$T_{n+1}=\frac{1}{8}(\alpha\beta+\gamma\delta)(U+V)+\frac{1}{16}(\alpha\delta+\beta\gamma)^2(W+Y)+\frac{1}{8}(\alpha\gamma+\beta\delta)^2X.$
AABb×*AB.ab*	8	$U_{n+1}=\frac{1}{2}J+\frac{1}{4}(\alpha\beta+\gamma\delta)(L+N)+\frac{1}{8}(S+T)+\frac{1}{8}(\alpha+\gamma)U+\frac{1}{8}(\beta+\delta)V+\frac{1}{8}\alpha\gamma(\beta\gamma+\alpha\delta)W+\frac{1}{8}(\alpha\gamma+\beta\delta)(\alpha\delta+\beta\gamma)X+\frac{1}{8}\beta\delta(\beta\gamma+\alpha\delta)Y.$
AABb×*Ab.aB*	8	$V_{n+1}=\frac{1}{2}K+\frac{1}{4}(\alpha\beta+\gamma\delta)(M+P)+\frac{1}{8}(R+T)+\frac{1}{8}(\beta+\delta)U+\frac{1}{8}(\alpha+\gamma)V+\frac{1}{8}\beta\delta(\beta\gamma+\alpha\delta)W+\frac{1}{8}(\alpha\gamma+\beta\delta)(\alpha\delta+\beta\gamma)X+\frac{1}{8}\alpha\gamma(\beta\gamma+\alpha\delta)Y.$
AB.ab×*AB.ab*	1	$W_{n+1}=2(E+J)+\frac{1}{2}(\alpha^2+\gamma^2)L+\frac{1}{2}(\beta^2+\delta^2)N+\frac{1}{4}(S+T)+\frac{1}{4}(\alpha^2+\gamma^2)U+\frac{1}{4}(\beta^2+\delta^2)V+\frac{1}{4}\alpha^2\gamma^2W+\frac{1}{4}(\alpha^2\delta^2+\beta^2\gamma^2)X+\frac{1}{4}\beta^2\delta^2Y.$
AB.ab×*Ab.aB*	2	$X_{n+1}=\frac{1}{2}T+\frac{1}{2}(\alpha\beta+\gamma\delta)(U+V)+\frac{1}{2}\alpha\beta\gamma\delta(W+2X+Y).$
Ab.aB×*Ab.aB*	1	$Y_{n+1}=2(F+K)+\frac{1}{2}(\alpha^2+\gamma^2)M+\frac{1}{2}(\beta^2+\delta^2)P+\frac{1}{4}(R+T)+\frac{1}{4}(\beta^2+\delta^2)U+\frac{1}{4}(\alpha^2+\gamma^2)V+\frac{1}{4}\beta^2\delta^2W+\frac{1}{4}(\alpha^2\delta^2+\beta^2\gamma^2)X+\frac{1}{4}\alpha^2\gamma^2Y.$

8

Result for sib-mating

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

$$\zeta = \frac{q}{2-3q}, \quad \theta = \frac{2q}{2-3q}, \quad \kappa = \frac{1}{2-3q},$$

$$\lambda = \frac{1-2q}{2-3q}, \quad \mu = \frac{1-2q}{2-3q}, \quad \nu = \frac{2q}{2-3q}$$

as may easily be verified.

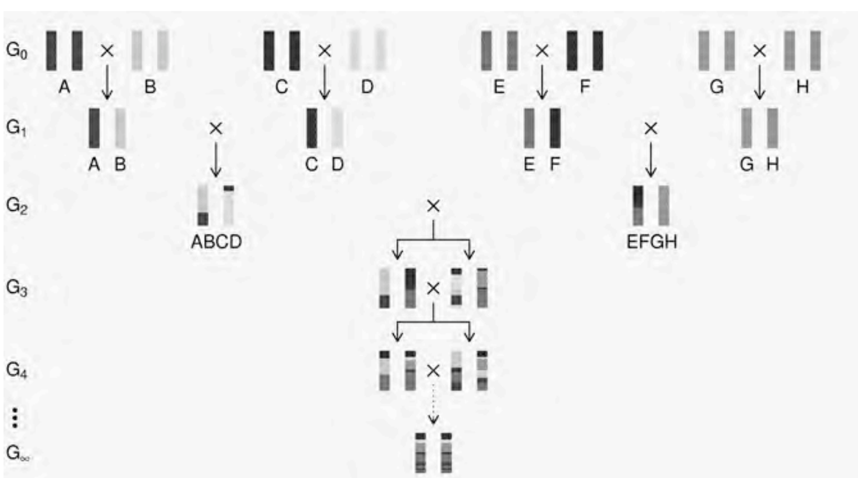
$$\therefore \; c_\infty = c_n + 2e_n + \frac{1}{1+6x}\left[(1-2x)(d_n + 2f_n + 2j_n + \tfrac{1}{2}k_n) + 2g_n + 4x(h_n + i_n)\right] \quad (3.4)$$

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

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

9

The "Collaborative Cross"



10

8-way RILs

Autosomes

$Pr(G_1 = i) = 1/8$

$Pr(G_2 = j \mid G_1 = i) = r / (1+6r)$ for $i \neq j$

$Pr(G_2 \neq G_1) = 7r / (1+6r)$

X chromosome

$Pr(G_1=A) = Pr(G_1=B) = Pr(G_1=E) = Pr(G_1=F) = 1/6$

$Pr(G_1=C) = 1/3$

$Pr(G_2=B \mid G_1=A) = r / (1+4r)$

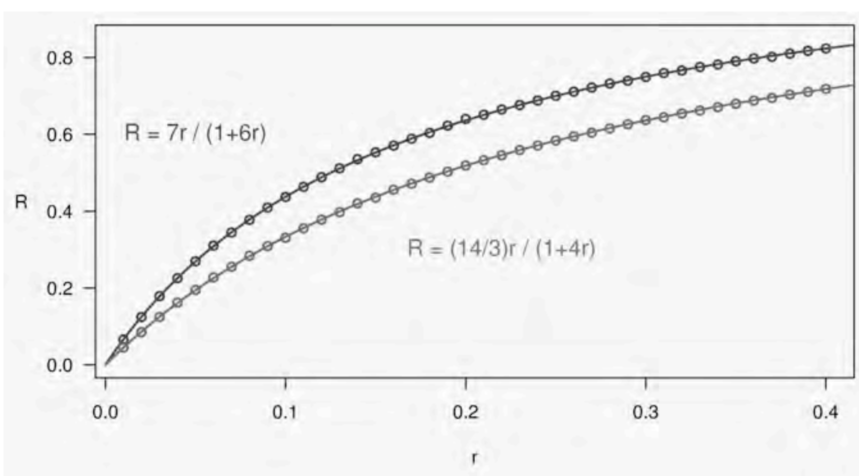
$Pr(G_2=C \mid G_1=A) = 2r / (1+4r)$

$Pr(G_2=A \mid G_1=C) = r / (1+4r)$

$Pr(G_2 \neq G_1) = (14/3)\ r / (1+4r)$

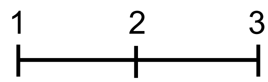
11

Computer simulations



12

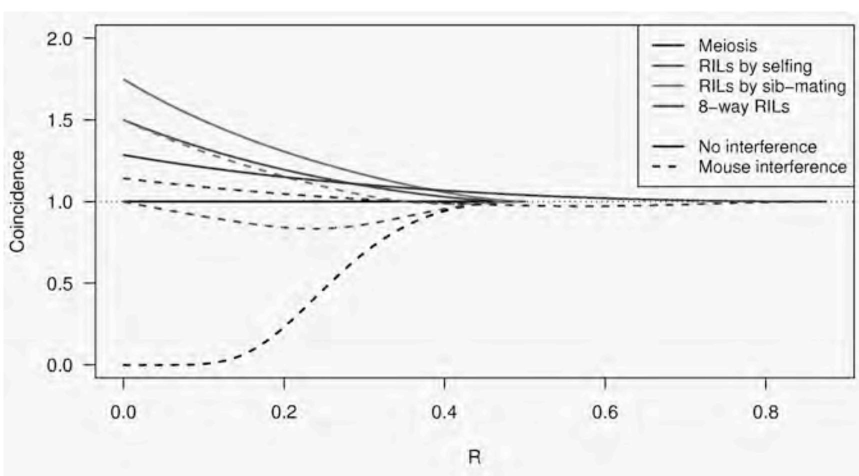
3-point coincidence



- r_{ij} = recombination fraction for interval i,j;
  assume $r_{12} = r_{23} = r$
- Coincidence = c = Pr(double recombinant) / r^2
  = Pr(rec'n in 23 | rec'n in 12) / Pr(rec'n in 23)
- No interference → = 1
  Positive interference → < 1
  Negative interference → > 1
- Generally c is a function of r.

13

Coincidence



14

Summary

- The Collaborative Cross could provide "one-stop shopping" for gene mapping in the mouse.
- Use of such 8-way RILs requires an understanding of the breakpoint process.
- We've extended Haldane & Waddington's results to the case of 8-way RILs: $R = 7r / (1 + 6r)$.
- We've shown clustering of breakpoints in RILs by sib-mating, even in the presence of strong crossover interference.
- Broman KW (2005) The genomes of recombinant inbred lines. *Genetics* 169:1133-1146

15