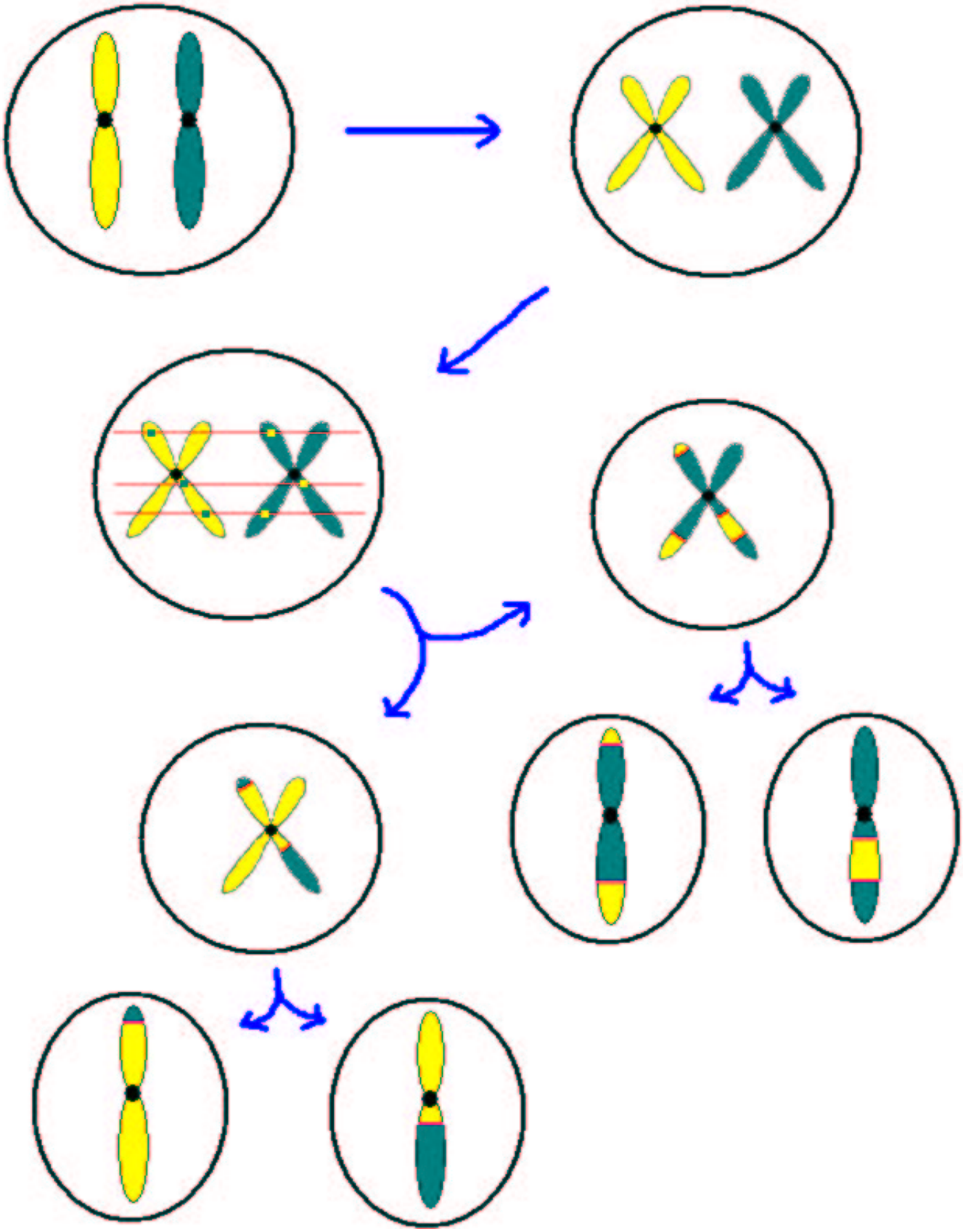
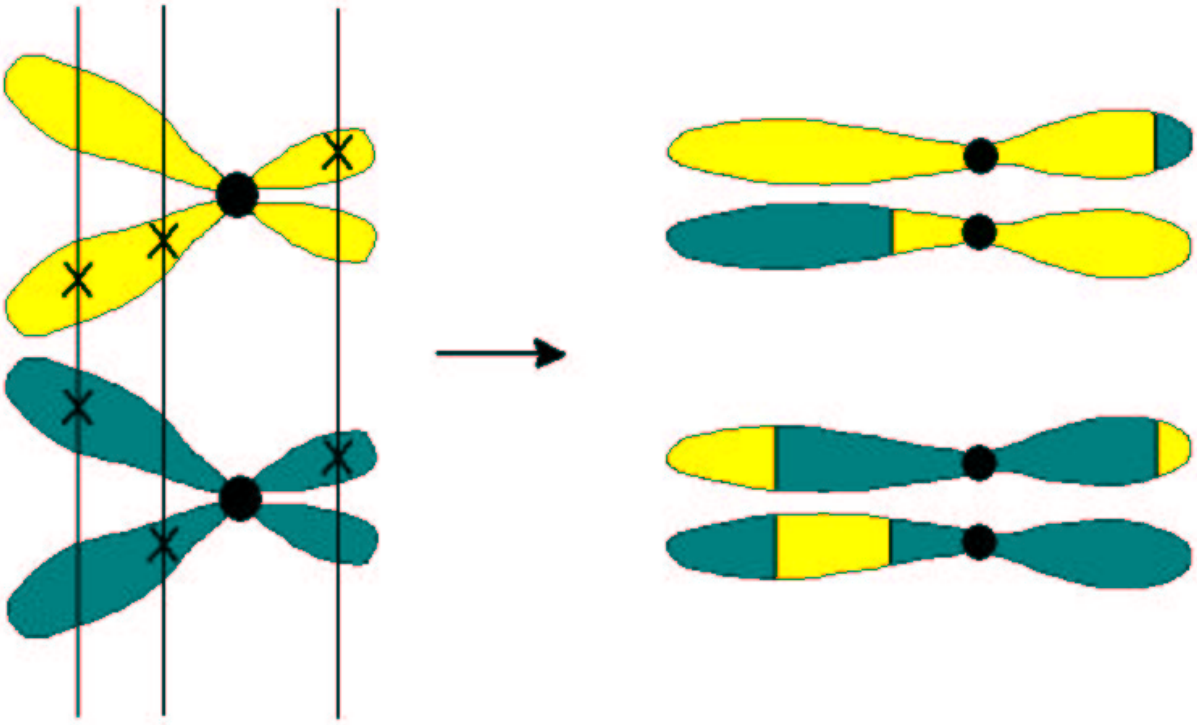
Crossover interference in the mouse

Karl W Broman
Dept of Biostatistics
Johns Hopkins University

Meiosis



Interference



- Strand choice

  → Chromatid interference

- Spacing

  → Chiasma (crossover) interference

Why study interference?

- Obtain a model of meiosis for simulation and analysis
- It's interesting
- Determine shortest possible distance between crossovers

Goals

- Compare stochastic models for meiosis
- Characterize the level of interference in the mouse
- Compare level of interference between chromosomes

Recombination

Crossovers on random meiotic product

Typical data: recombination information

We generally do not observe the locations of crossovers; rather, we observe the grandparental origin of DNA at a set of genetic markers.

Recombination across an interval indicates an odd number of crossovers.

Genetic distance

distance (cM) = average # crossovers in 100 meiotic products

per Morgan { 2 chiasmata on 4-strand bundle
1 crossover on meiotic product

Map function

recombination fraction as a function of genetic distance

Haldane	$r(d) = \frac{1}{2} [1 - \exp(-2d)]$
Kosambi	$r(d) = \frac{1}{2} \tanh(2d)$
Carter-Falconer	$d(r) = [\tanh^{-1} 2r + \tan^{-1} 2r] /4$

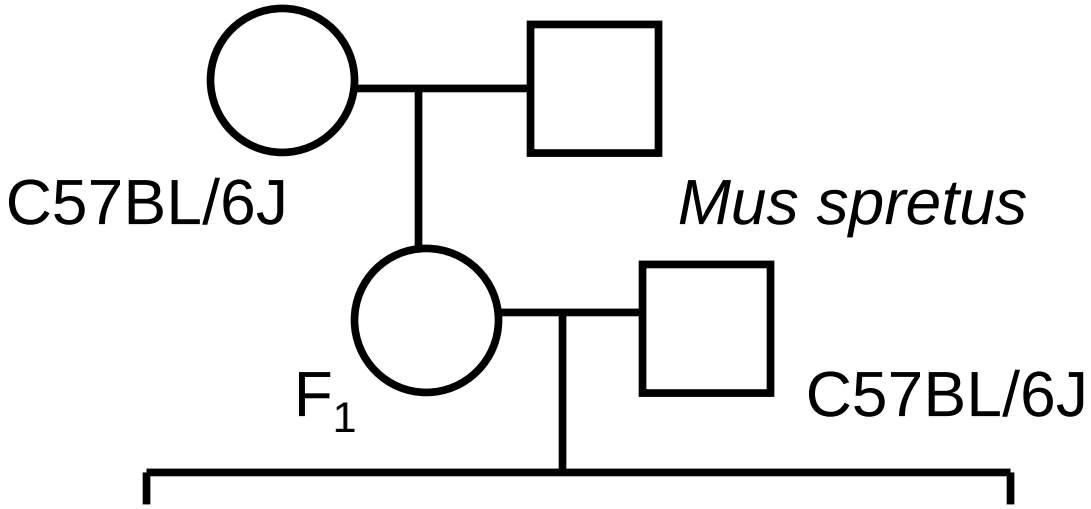
The usual data

- Lots of meioses
- A few linked markers
- Look at frequency of rare multiple recombination events

***Drosophila* data (Morgan et al 1935)**

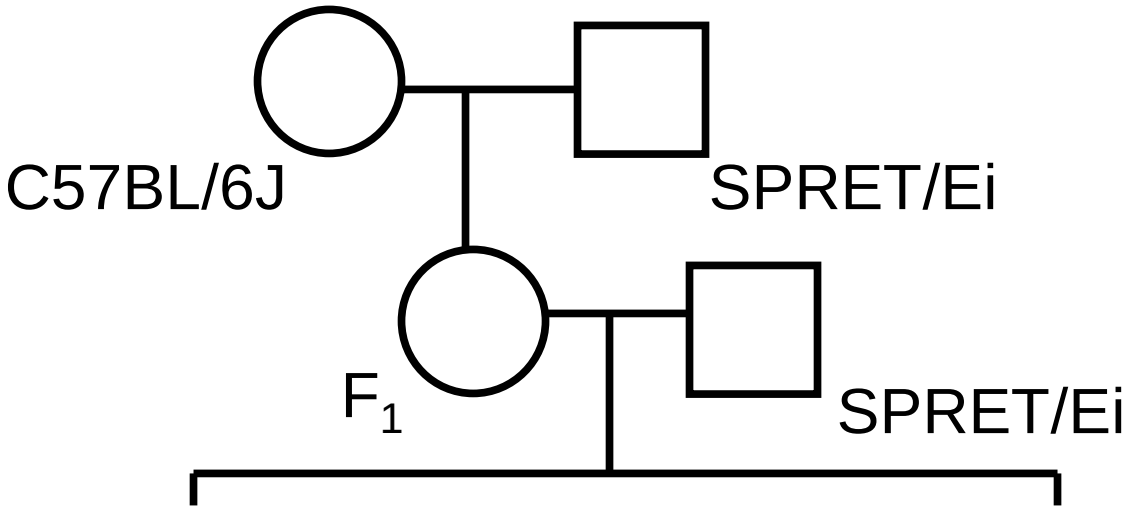
Event	Count	Event	Count
0000	10,431	1001	46
1000	771	0101	53
0100	1,579	0011	25
0010	1,221	1110	1
0001	1,994	1101	1
1100	4	1011	1
1010	7	0111	1
0110	4	1111	1

Our data



94 BSB progeny

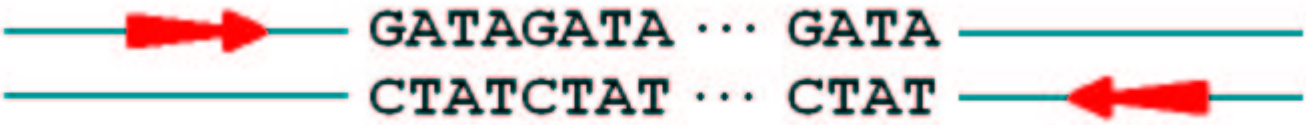
Typed at 1372 markers

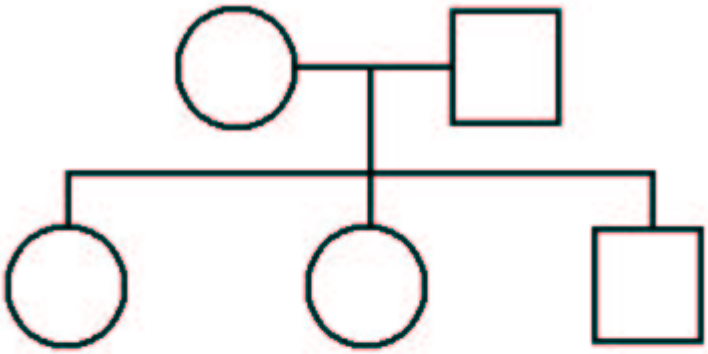


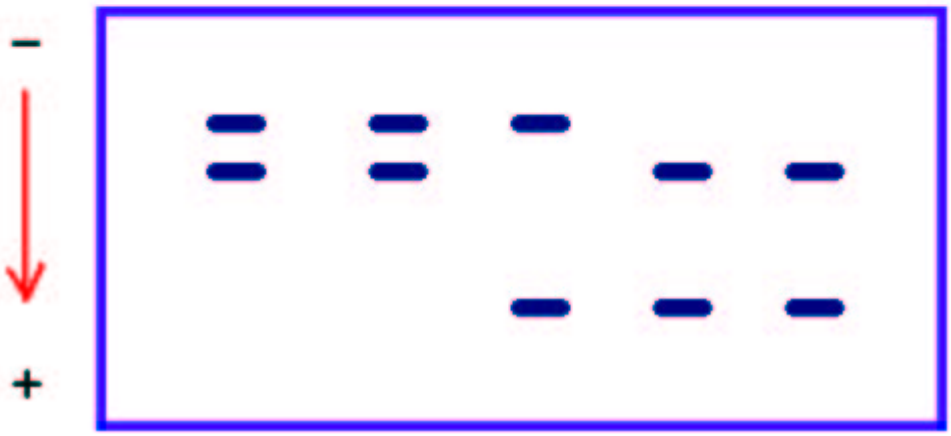
94 BSS progeny

Typed at 4913 markers

Genetic markers: STRPs or microsatellites

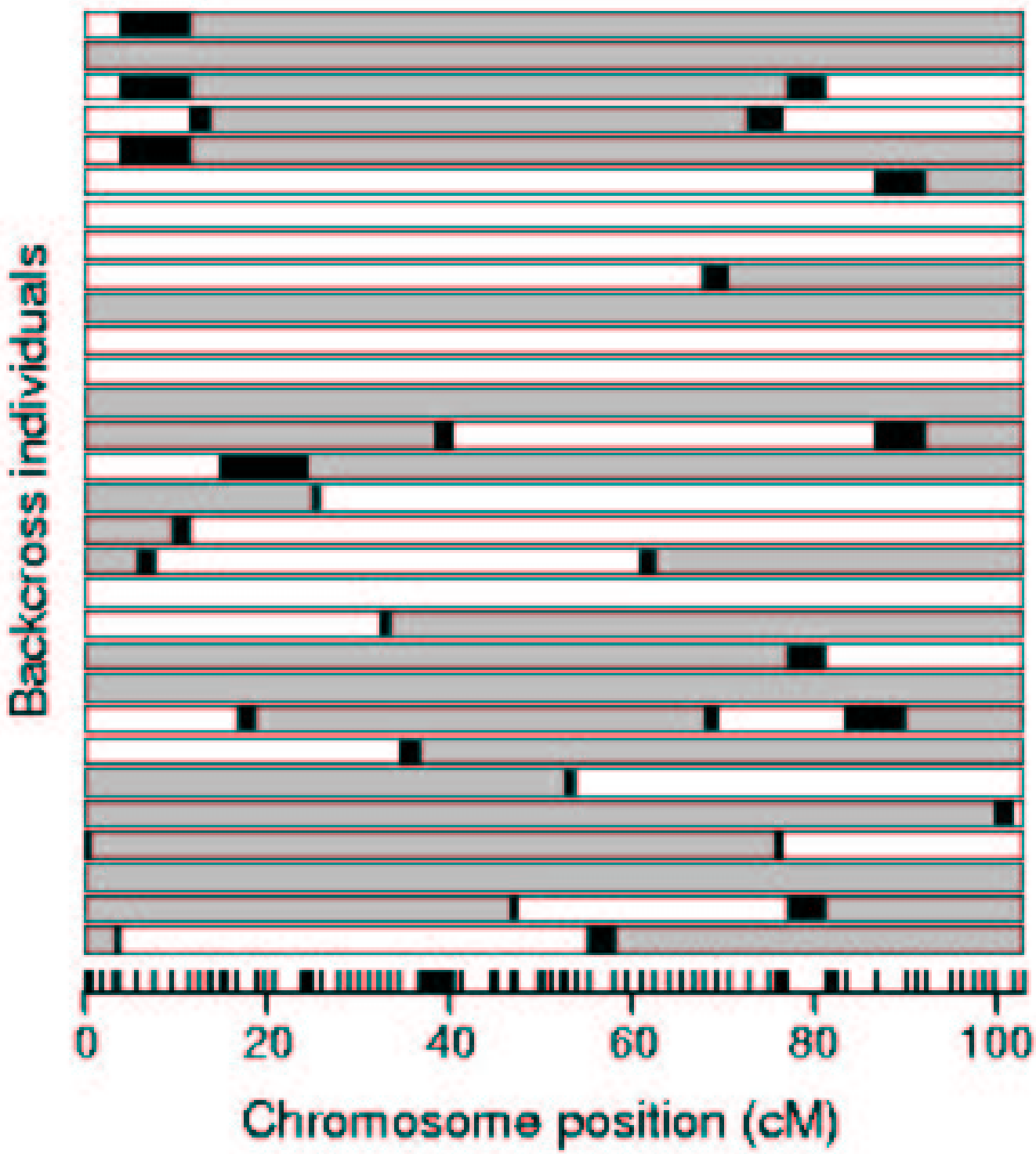




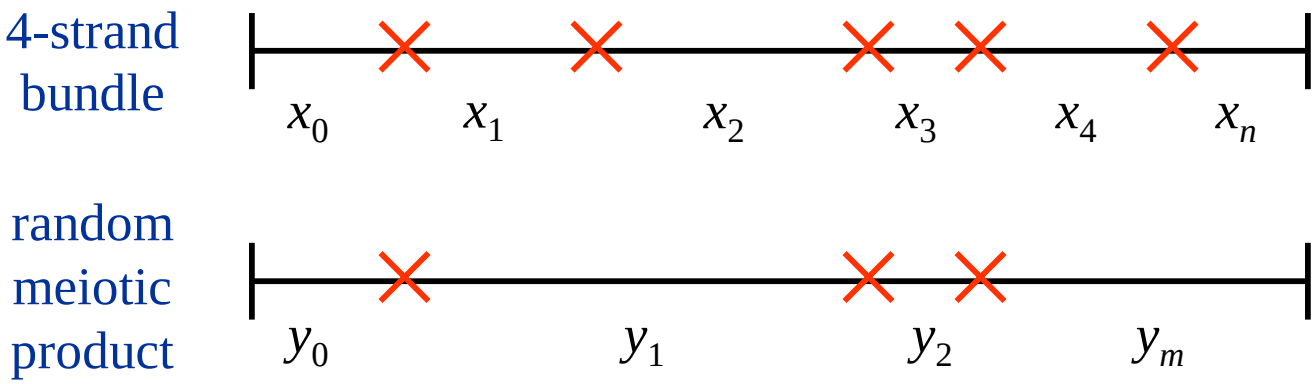


Basic methods

- Form integrated genetic map for the two crosses
- Identify intervals showing a recombination event
- Assume that recombination events indicate single crossovers, and that no double crossovers occurred
- Assume crossovers occurred at the center of the relevant interval (i.e., ignore interval censoring)
- Assume genetic distances known exactly (i.e., ignore sampling error)


Backcross individuals
0
20
40
60
80
100
Chromosome position (cM)

Models



- Count-location model

  $n \sim (p_0, p_1, p_2, \ldots)$

  locations | n ~ iid uniform

- Gamma model

  x_i's ~ stationary gamma renewal process (shape = u, rate = $2u$)

  y_i's ~ mixtures of gammas

Model fitting

- Count-location model

  m_i = # crossovers
  n_i = underlying # chiasmata

  $n_i \sim (p_0, p_1, p_2, \ldots)$
  $m_i \mid n_i \sim \text{binomial}(n_i, 1/2)$

  MLEs via a version of the EM algorithm

Model fitting

- Gamma model

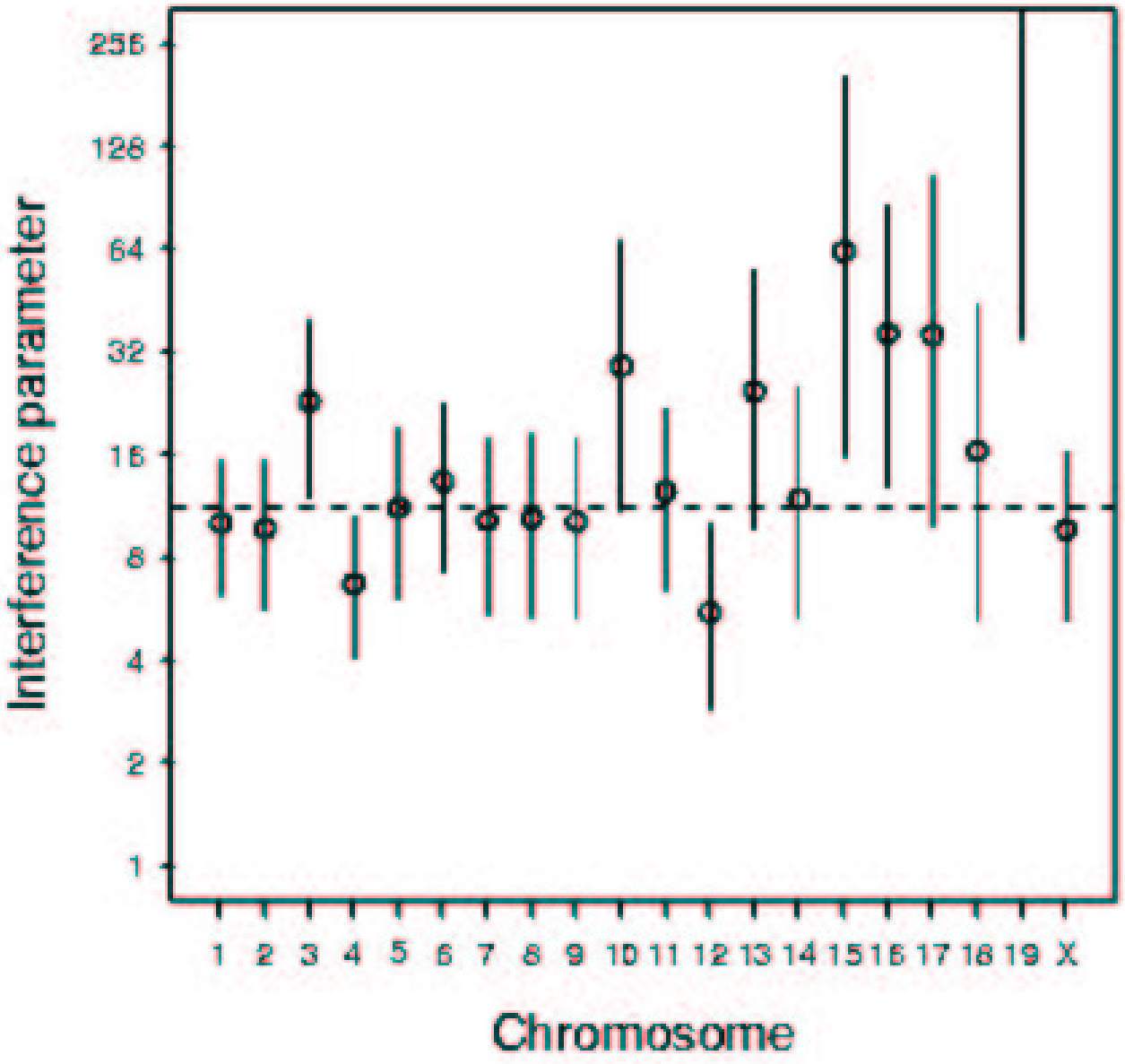
$x_1, x_2, \ldots \sim f(u, 2u)$
$x_0 \sim g = 2[1 - F(u, 2u)]$
x_i's independent

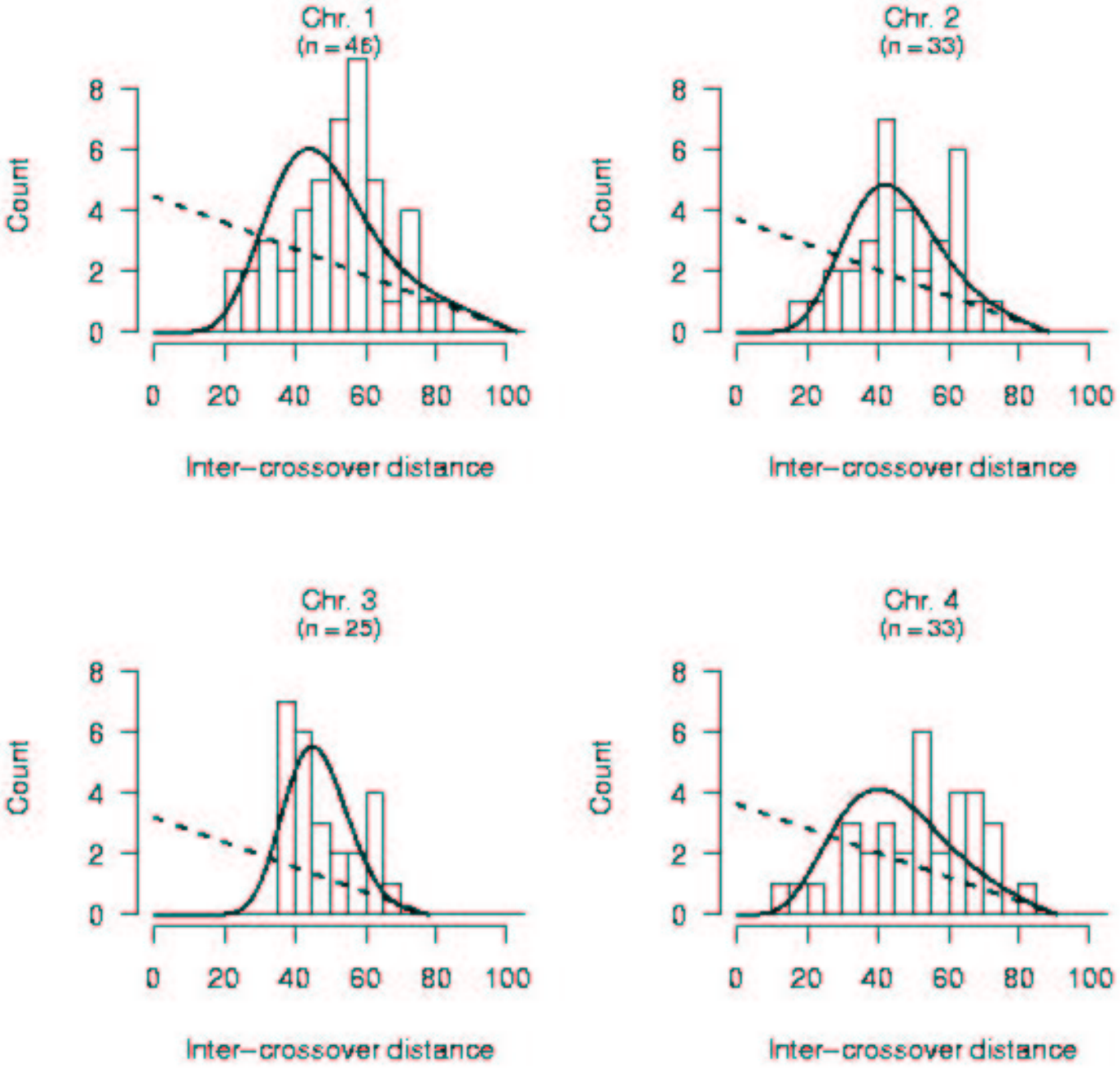
$y_1, y_2, \ldots \sim \Sigma\, (½)^k\, f(ku, 2u)$
$y_0 \sim ½\, g + \Sigma\, (½)^{(k+1)}\, g * f(ku, 2u)$
y_i's independent

- MLE of u using y_i's
- g calculated numerically
- Convolutions calculated numerically
- Maximization performed using a quasi-Newton method

Distributions of # XOs / chr

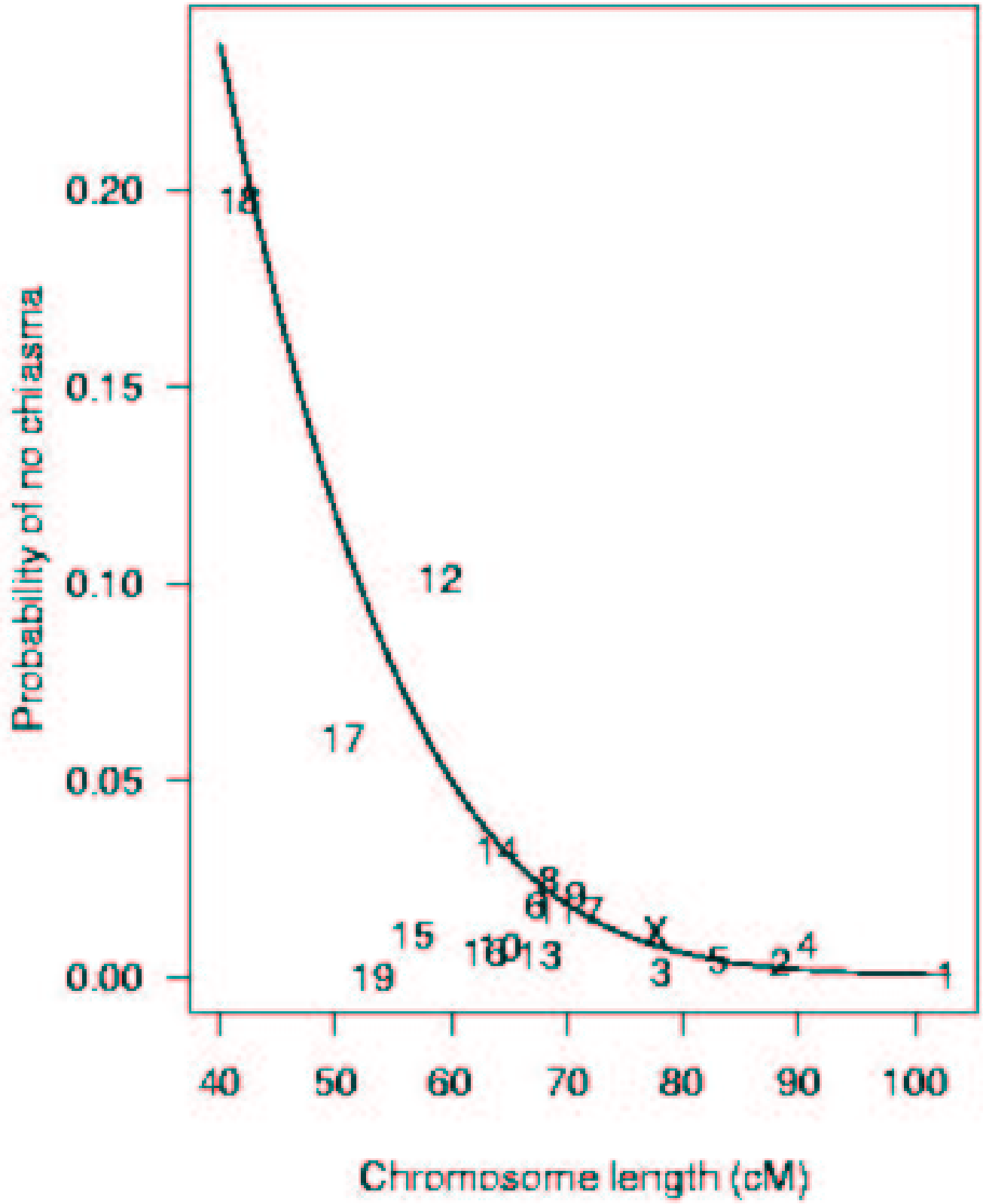
Source of Data	No. of Recombinations 0	1	2	3	χ^2
Chr. 1:					
Observed	53	87	46	2	
Poisson	70.3	69.2	34.0	14.5	23.9
Truncated Poisson	60.9	86.0	31.7	9.5	13.4
CL	52.7	88.6	44.6	2.1	0.1
Obligate-chiasma CL	50.6	91.9	43.4	2.2	0.5
Gamma	49.3	89.1	44.7	5.0	2.1
Chr. 14:					
Observed	76	108	4	0	
Poisson	101.4	62.6	19.3	4.7	56.1
Truncated Poisson	89.7	93.9	4.3	0.1	4.4
CL	89.5	93.8	4.6	0.1	4.4
Obligate-chiasma CL	89.5	93.9	4.5	0.1	4.3
Gamma	82.5	90.9	14.5	0.1	11.4

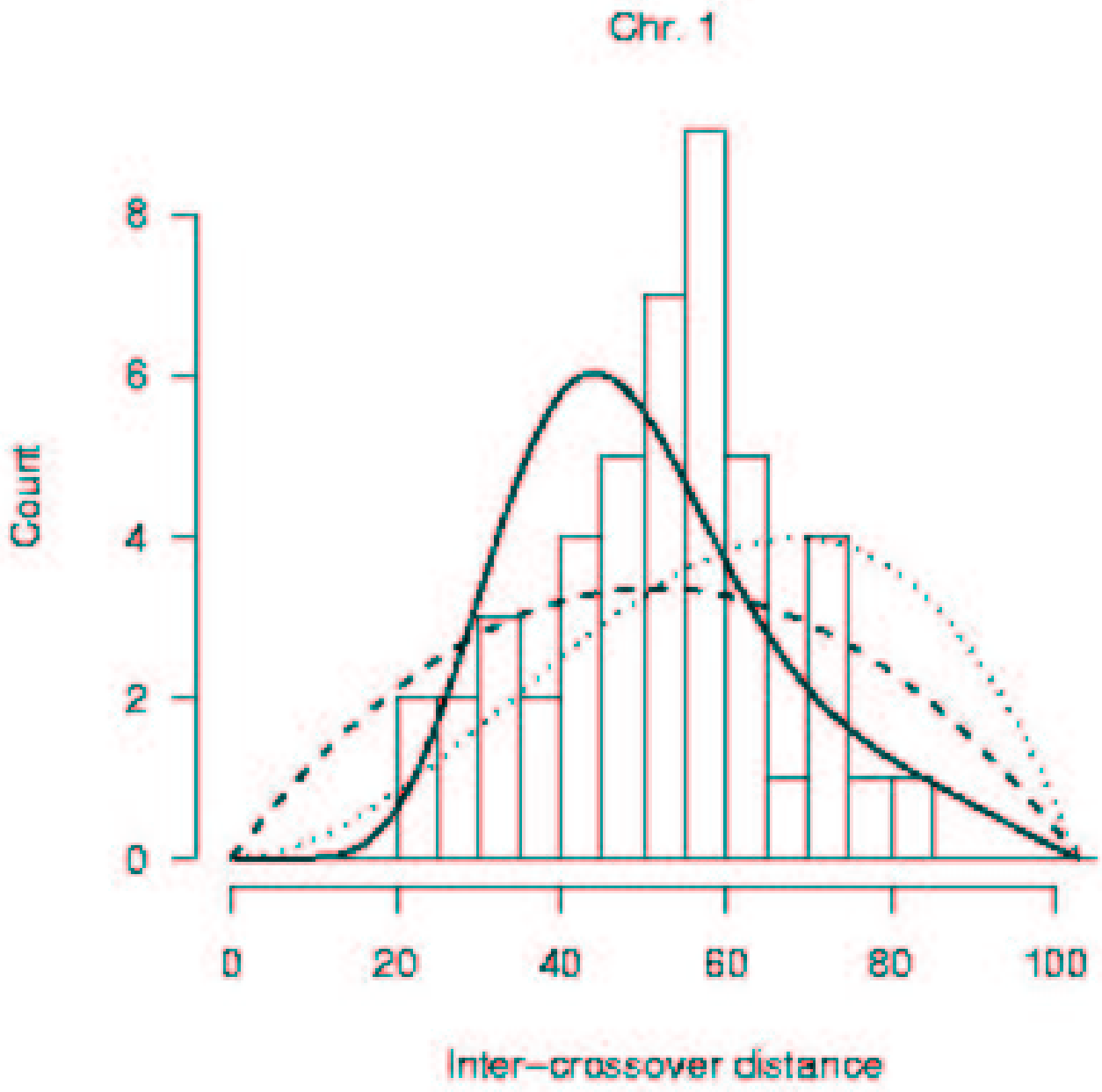

Interference parameter
256
128
64
32
16
8
4
2
1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 X
Chromosome


Chr. 1
(n = 46)
Chr. 2
(n = 33)
Chr. 3
(n = 25)
Chr. 4
(n = 33)
Count
0
2
4
6
8
0
20
40
60
80
100
Inter−crossover distance

Discussion

- Approximations
  - Correct marker order
  - Correct genetic distances
  - All crossovers observed
  - Interval censoring unimportant
  - Interference constant across chromosome
- Conclusions
  - Gamma model fits well
  - Count-location model fits poorly
  - Gamma parameter, $\nu \approx 11$ (stronger than Carter-Falconer, $\nu \approx 7.2$)
  - Apparent variation between chromosomes, with stonger interference in smaller chromosomes


Probability of no chiasma
0.20
0.15
0.10
0.05
0.00
40
50
60
70
80
90
100
Chromosome length (cM)


Chr. 1
Count
0
2
4
6
8
0
20
40
60
80
100
Inter−crossover distance