Genetic maps

past, present and future

Karl W Broman

Biostatistics & Medical Informatics
University of Wisconsin – Madison

`www.biostat.wisc.edu/~kbroman`

Genetic maps

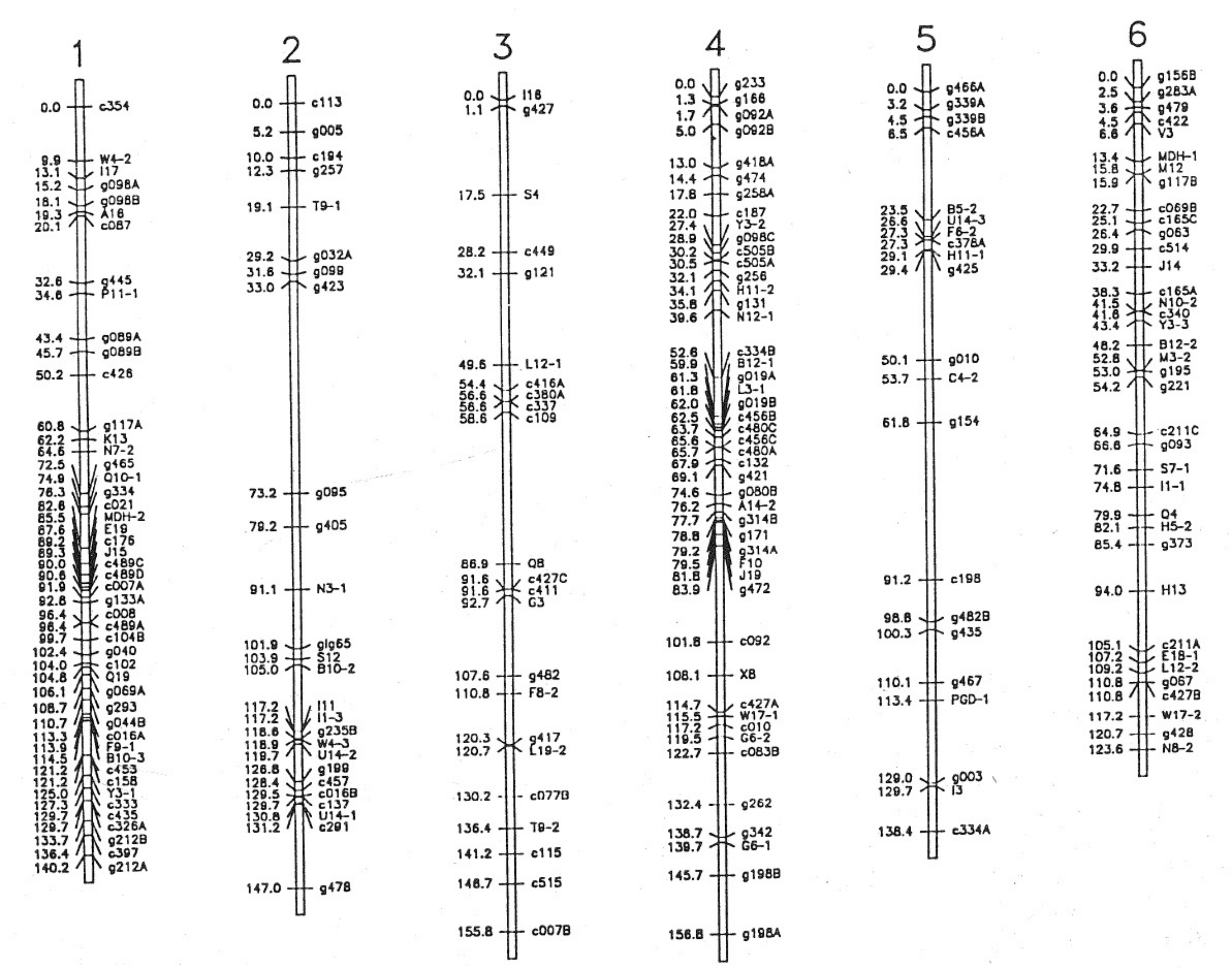
from my past, present and future

Karl W Broman

Biostatistics & Medical Informatics
University of Wisconsin – Madison

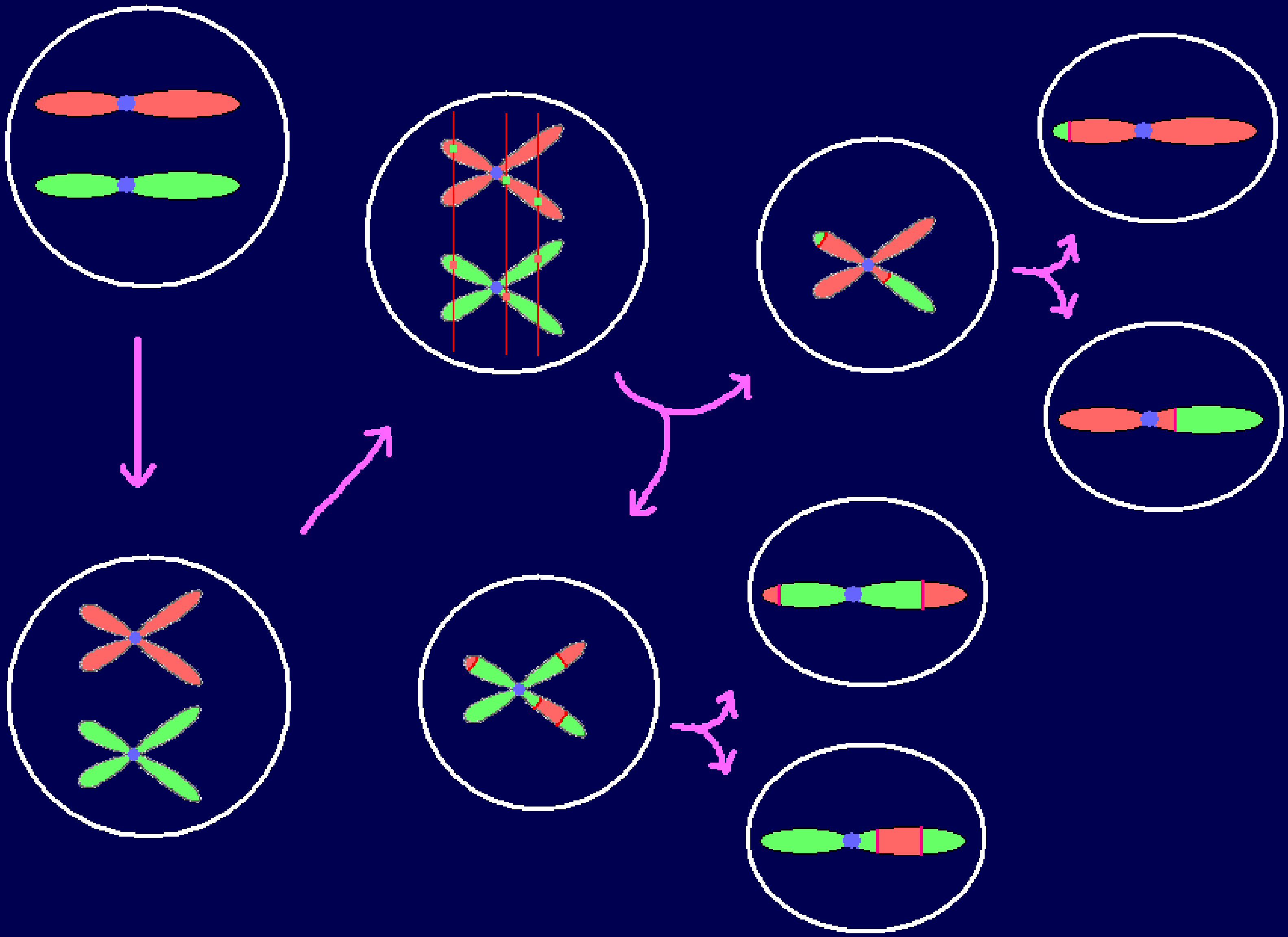
`www.biostat.wisc.edu/~kbroman`

Eucalypt genetic map



Byrne et al., Theor Appl Genet 91:869–875, 1995

Meiosis



Genetic distance

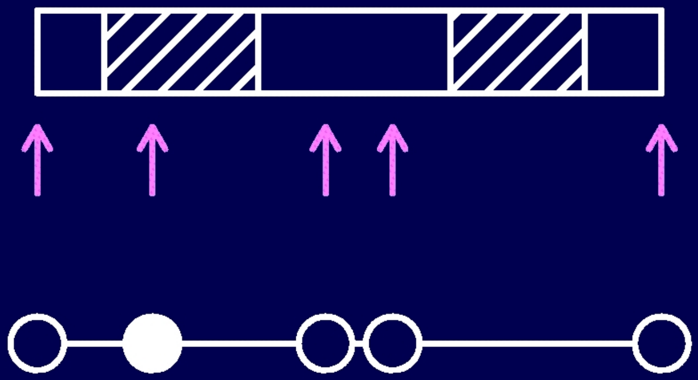
- Genetic distance between two markers (in cM) =

  Average number of crossovers in the interval
  in 100 meiotic products.

- “Intensity” of the crossover point process

- Recombination rate varies by
  - Organism
  - Sex
  - Chromosome
  - Position on chromosome

Recombination fraction



Crossover
process

Marker
data

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

Recombination across an interval indicates an odd number of crossovers.

Recombination fraction =

Pr(recombination in interval) = Pr(odd no. XOs in interval)

Map functions

- A map function relates the genetic length of an interval and the recombination fraction.

$$r = M(d)$$

- Map functions are related to crossover interference, but a map function is not sufficient to define the crossover process.

- Haldane map function: no crossover interference

- Kosambi: similar to the level of interference in humans

- Carter-Falconer: similar to the level of interference in mice

Ordering markers

$$\begin{array}{c|c} \mathsf{A} & \mathsf{a} \\ \mathsf{B} & \mathsf{b} \\ \mathsf{C} & \mathsf{c} \end{array} \quad \longrightarrow \quad \begin{array}{cc} \mathsf{ABC} & \mathsf{abc} \\ \mathsf{ABc} & \mathsf{abC} \\ \mathsf{Abc} & \mathsf{aBC} \\ \mathsf{AbC} & \mathsf{aBc} \end{array}$$

Marker orders: A–B–C A–C–B B–A–C

With M markers, there are $M!/2$ possible orderings.

For $M = 100$, $M!/2 \approx 10^{157}$

Ordering markers

$$\begin{array}{c|c} \mathsf{A} & \mathsf{a} \\ \mathsf{B} & \mathsf{b} \\ \mathsf{C} & \mathsf{c} \end{array} \quad \longrightarrow \quad \begin{array}{cc} \mathsf{ABC} & \mathsf{abc} \\ \mathsf{ABc} & \mathsf{abC} \\ \mathsf{Abc} & \mathsf{aBC} \\ \mathsf{AbC} & \mathsf{aBc} \end{array}$$

Marker orders: A–B–C A–C–B B–A–C

With M markers, there are $M!/2$ possible orderings.

For $M = 100$, $M!/2 \approx 10^{157}$

Ordering markers

$$\begin{array}{c|c} A & a \\ B & b \\ C & c \end{array} \longrightarrow \begin{array}{cc} ABC & abc \\ ABc & abC \\ Abc & aBC \\ AbC & aBc \end{array}$$

Marker orders: A–B–C A–C–B B–A–C

With M markers, there are $M!/2$ possible orderings.

For $M = 100$, $M!/2 \approx 10^{157}$

Ordering markers

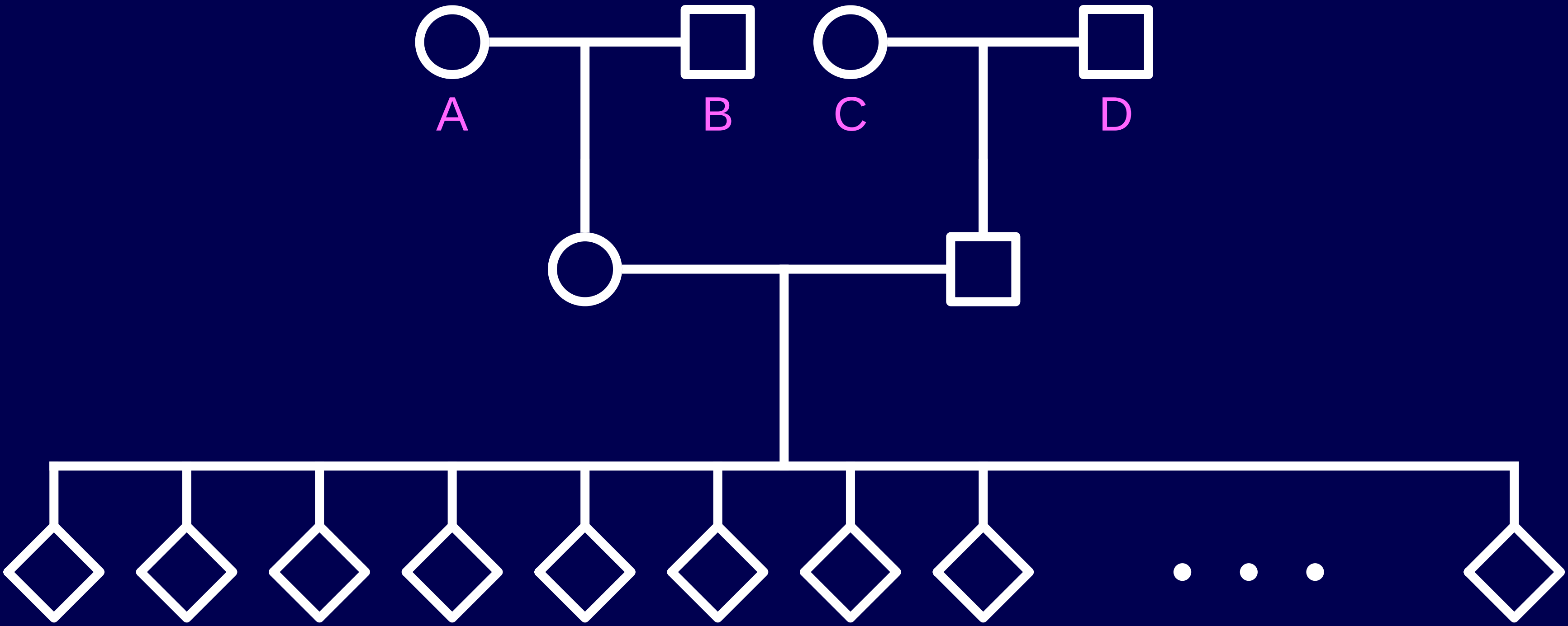
$$\begin{array}{c|c} \mathsf{A} & \mathsf{a} \\ \mathsf{B} & \mathsf{b} \\ \mathsf{C} & \mathsf{c} \end{array} \quad \longrightarrow \quad \begin{array}{cc} \mathsf{ABC} & \mathsf{abc} \\ \mathsf{ABc} & \mathsf{abC} \\ \mathsf{Abc} & \mathsf{aBC} \\ \mathsf{AbC} & \mathsf{aBc} \end{array}$$

Marker orders: A–B–C A–C–B B–A–C

With M markers, there are $M!/2$ possible orderings.

For $M = 100$, $M!/2 \approx 10^{157}$

Eucalypt pedigree

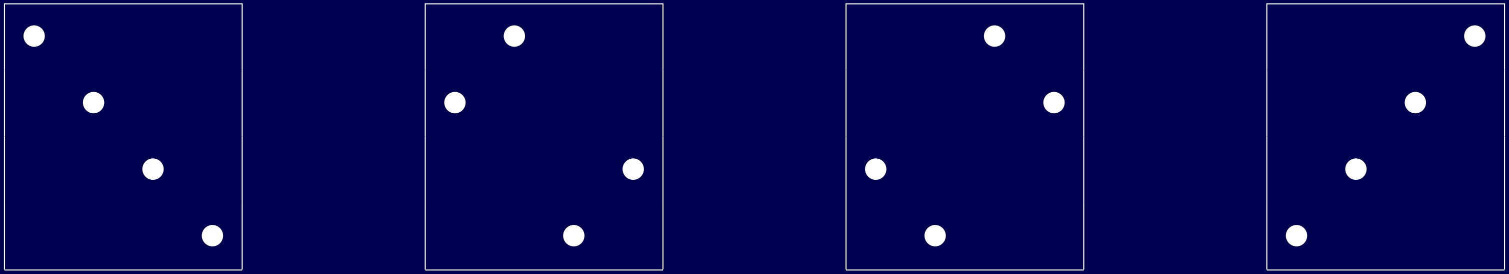


118 progeny

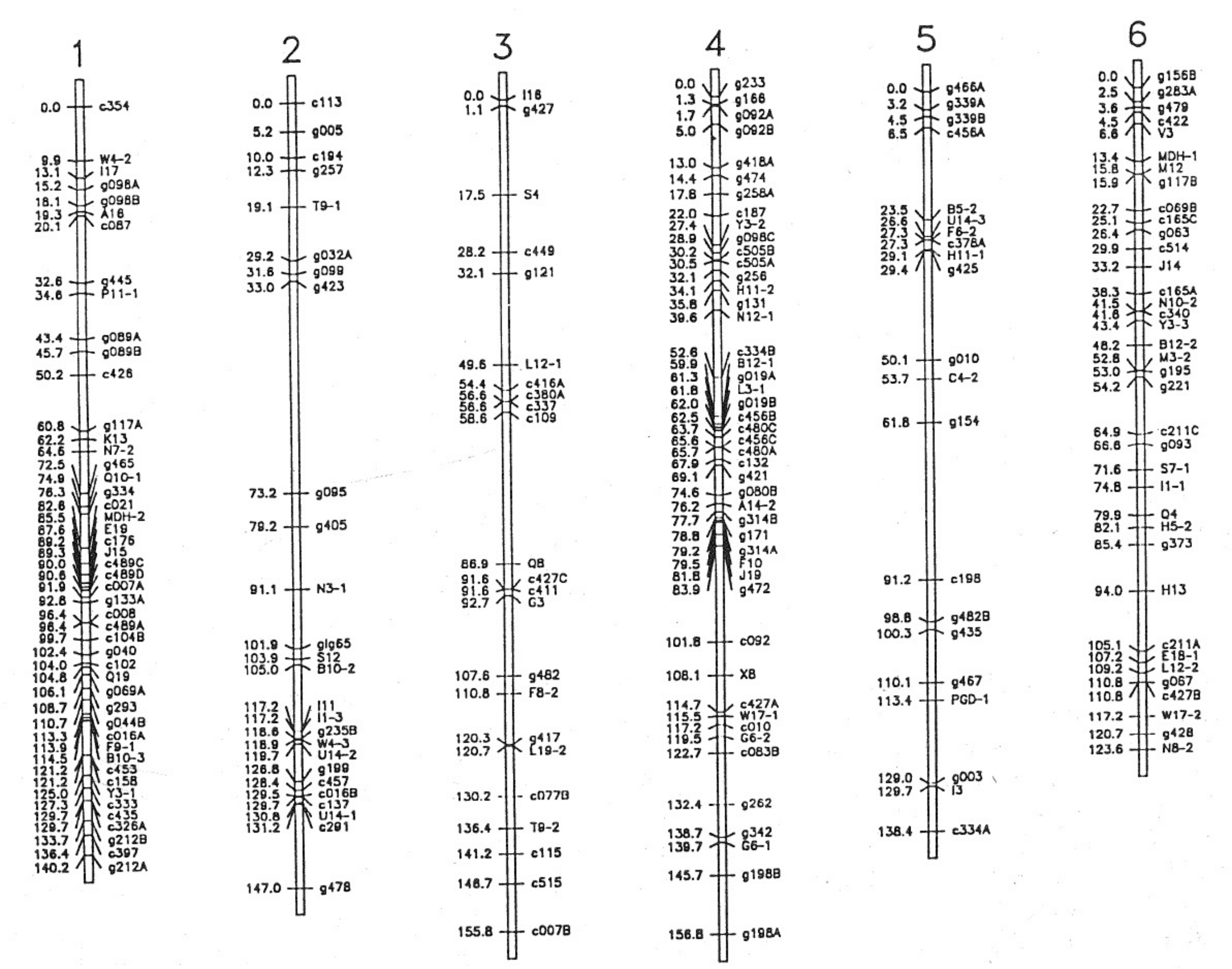
Inferring phase

$$\left(\begin{array}{c|c} A_1 & B_1 \\ A_2 & B_2 \end{array} \text{ or } \begin{array}{c|c} A_1 & B_1 \\ B_2 & A_2 \end{array}\right) \times \left(\begin{array}{c|c} C_1 & D_1 \\ C_2 & D_2 \end{array} \text{ or } \begin{array}{c|c} C_1 & D_1 \\ D_2 & C_2 \end{array}\right)$$

	locus 1			
locus 2	AC	AD	BC	BC
AC	●			
AD		●		
BC			●	
BD				●



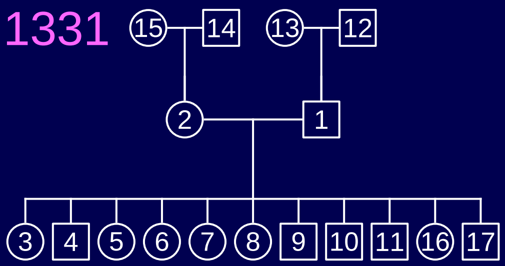
Eucalypt genetic map



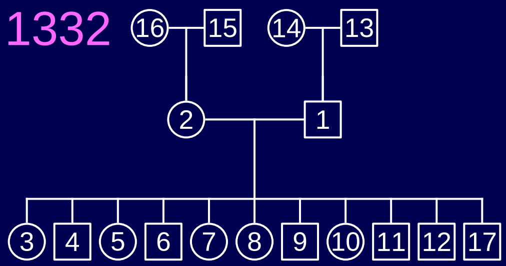
Byrne et al., Theor Appl Genet 91:869–875, 1995

CEPH pedigrees

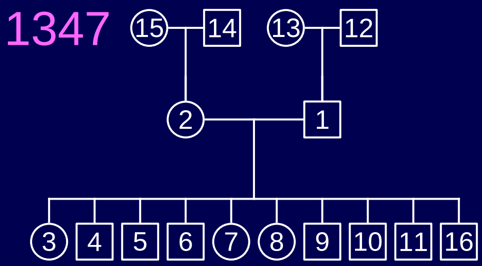
1331



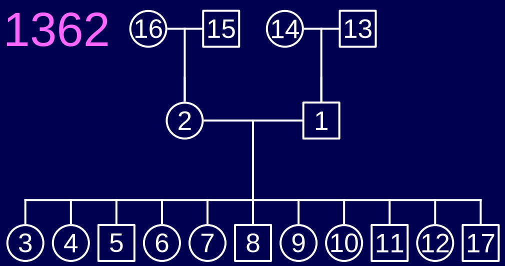
1332



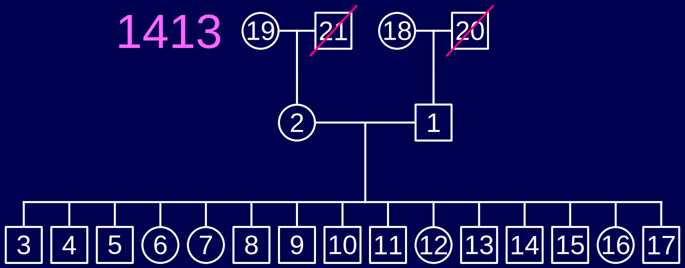
1347



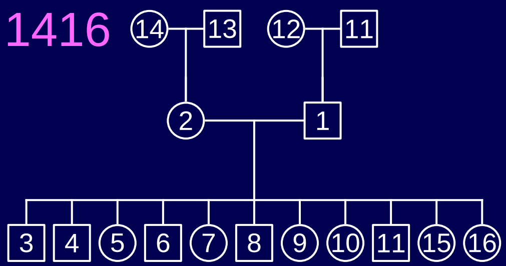
1362



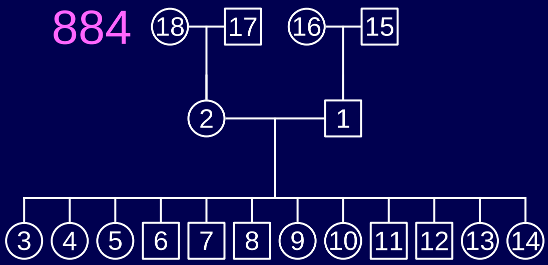
1413



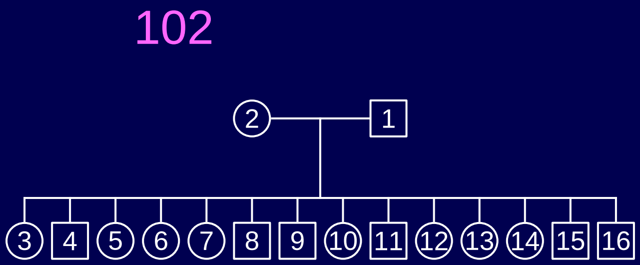
1416



884



102



Marshfield maps: Tasks

- Assemble data
- Understand marker names
  - AFM, UT, CHLC (GATA etc.), Mfd, D*S*
- Identify cryptic duplicates
- Order markers and identify genotyping errors
  - Removed 764 / 969,425 genotypes

CRIMAP chrompic

```
1332-03 ma -11-i--11--111-i111-11-1111i--1111i-1111-i--11---1--11-1111-1-1i1---1...
1332-03 pa 0000----0000000o00o00-000-000-0000o00-000-00000-00001---000-00-o000-0...

1332-04 ma -11-i--11--111-1111-11-i111i--i1111-1111-i--11---1--11-1111-1-11i--11...
1332-04 pa 1111----111111111i11-1i1-111-i111i11-111-11111-11111---111-11-1i1111...

1332-05 ma -11-i--11--111-i111-11-1111o--0000o-0000-o--00---0--00-0000-0-0o0--00...
1332-05 pa 0000----0000000o00o00-000-111-1111i11-111-1111--11111---111-11-i11111...

1332-06 ma -00-o--00--000-o000-00-0000o--0000o-0000-o--00---0--00-0000-1-11i--11...
1332-06 pa 1111----1111111i11i11-111-111-1111i11-111-11111-11111---111-11-1i1111...

1332-07 ma -00-o--00--000-o000-00-0000o--0000o-0000-o--00---0--00-0000-0-0o0--00...
1332-07 pa 1111----1111111i11i11-111-111-1111i11-111-1111--11111---111-11-i11111...

1332-08 ma -10-o--00--000-00-0-00-0000o--o0000-0000-o--00---0--11-1111-1-1i1--11...
1332-08 pa 0000----000000000-o00-010-000-o000o00-000-00000-00000---000-00-o00000...

1332-10 ma -11-i--1---111-i111-11-1111i--1111i-1111-i--11---1--11-1111-1-1i1--11...
1332-10 pa 1000-----000000o00o00-000-000-0000o00-000-00000-00000---000-00-o00000...

1332-11 ma -11-o--00--000-o000-00-0000o--0000o-0000-o--00---0--00-0000-0-0o0--00...
1332-11 pa 1111----1111111i11i11-111-111-1111i11-111-11111-11111---111-11-i11111...

1332-12 ma -00-i--11--111-i111-11---11i--1111i-1111-i--11---1--11-1111-1-1i1---1...
1332-12 pa 0000----0000000o00o00-0---000-0000o00-000-00000-00000---000-00-o000-0...

1332-17 ma -11-i--1---11--i111-1--1111i--1111i-1111-i--11---1--11-1100-0-00o--00...
1332-17 pa 0000-----0000--o00o00-000-000-0000o-0-000-0000--00000---000-00-0o0000...
```

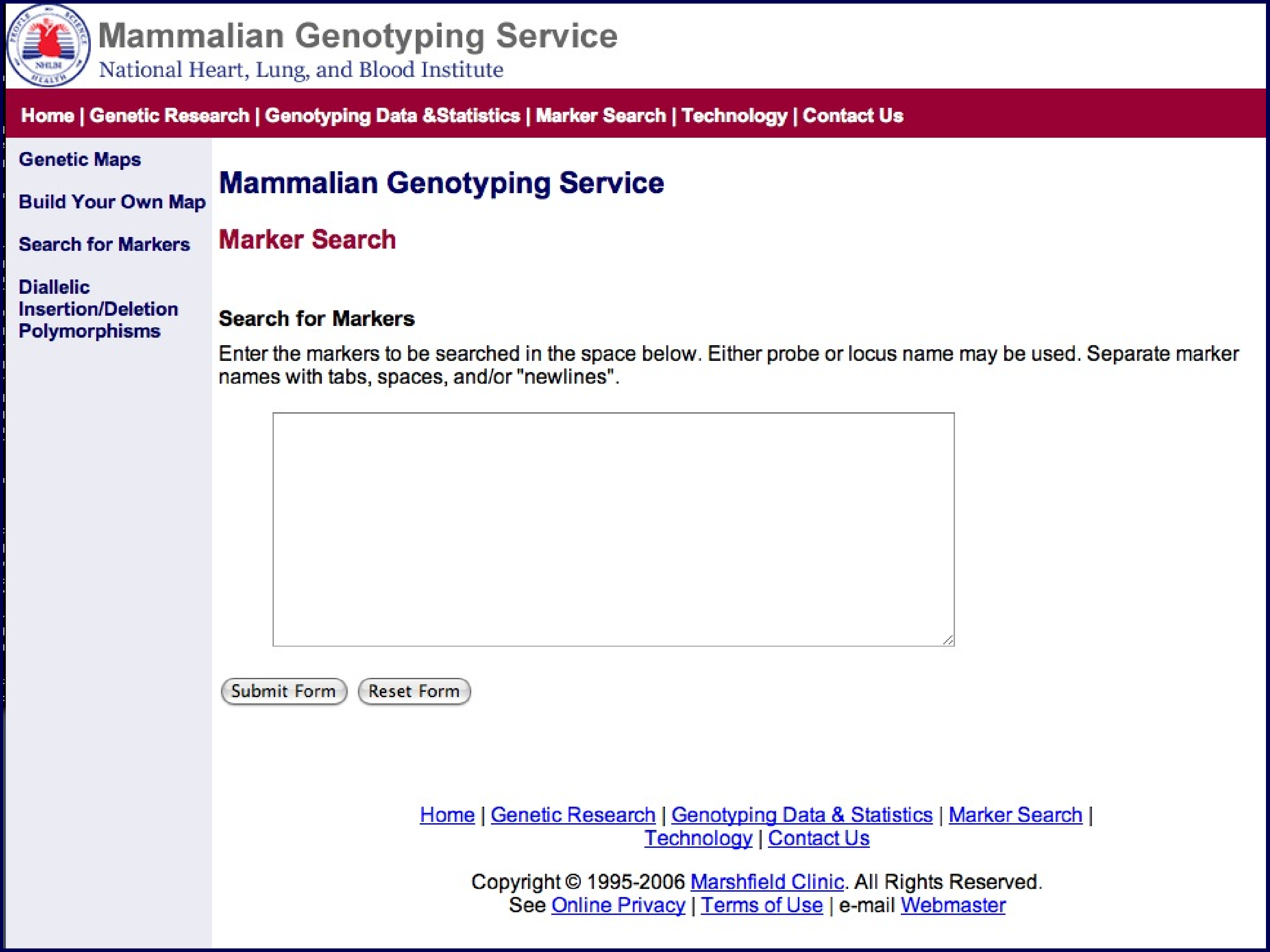
CRIMAP chrompic

```
1332-03 ma -11-i--11--111-i111-11-1111i--1111i-1111-i--11---1--11-1111-1-1i1---1...
1332-03 pa 0000----0000000o00o00-000-000-0000o00-000-00000-00001---000-00-o000-0...

1332-04 ma -11-i--11--111-1111-11-i111i--i1111-1111-i--11---1--11-1111-1-11i--11...
1332-04 pa 1111----111111111i11-1i1-111-i111i11-111-11111-11111---111-11-1i1111...

1332-05 ma -11-i--11--111-i111-11-1111o--0000o-0000-o--00---0--00-0000-0-0o0--00...
1332-05 pa 0000----0000000o00o00-000-111-1111i11-111-1111--11111---111-11-i11111...

1332-06 ma -00-o--00--000-o000-00-0000o--0000o-0000-o--00---0--00-0000-1-11i--11...
1332-06 pa 1111----1111111i11i11-111-111-1111i11-111-11111-11111---111-11-1i1111...

1332-07 ma -00-o--00--000-o000-00-0000o--0000o-0000-o--00---0--00-0000-0-0o0--00...
1332-07 pa 1111----1111111i11i11-111-111-1111i11-111-1111--11111---111-11-i11111...

1332-08 ma -10-o--00--000-00-0-00-0000o--o0000-0000-o--00---0--11-1111-1-1i1--11...
1332-08 pa 0000----000000000-o00-010-000-o000o00-000-00000-00000---000-00-o00000...

1332-10 ma -11-i--1---111-i111-11-1111i--1111i-1111-i--11---1--11-1111-1-1i1--11...
1332-10 pa 1000-----000000o00o00-000-000-0000o00-000-00000-00000---000-00-o00000...

1332-11 ma -11-o--00--000-o000-00-0000o--0000o-0000-o--00---0--00-0000-0-0o0--00...
1332-11 pa 1111----1111111i11i11-111-111-1111i11-111-11111-11111---111-11-i11111...

1332-12 ma -00-i--11--111-i111-11---11i--1111i-1111-i--11---1--11-1111-1-1i1---1...
1332-12 pa 0000----0000000o00o00-0---000-0000o00-000-00000-00000---000-00-o000-0...

1332-17 ma -11-i--1---11--i111-1--1111i--1111i-1111-i--11---1--11-1100-0-00o--00...
1332-17 pa 0000-----0000--o00o00-000-000-0000o-0-000-0000--00000---000-00-0o0000...
```

Top of chr 22

Marker	Dnumber	sex-ave(cM)	female(cM)	male(cM)
1 ATA2G02	Unknown	0.00	0.00	0.00
		1.79	0.00	2.60
2 GATA198B05	Unknown	1.79	0.00	2.60
		2.27	3.32	0.00
3 AFM217xf4	D22S420	4.06	3.32	2.60
		4.26	4.51	5.42
4 AFM288we5	D22S427	8.32	7.83	8.02
		5.25	7.52	3.00
5 265yf5	D22S425	13.57	15.35	11.02
		0.03	0.00	0.65
6 GGAA10F06	D22S686	13.60	15.35	11.67
		0.84	0.00	0.82
7 AFMa037zd1	D22S539	14.44	15.35	12.49
		0.00	0.00	0.00
8 AFM292va9	D22S446	14.44	15.35	12.49
		3.27	5.91	0.00
9 Mfd51	D22S257	17.71	21.26	12.49

Marker search



Mammalian Genotyping Service

National Heart, Lung, and Blood Institute

Mammalian Genotyping Service

Marker Search

Search for Markers

Enter the markers to be searched in the space below. Either probe or locus name may be used. Separate marker names with tabs, spaces, and/or "newlines".

Copyright © 1995-2006 Marshfield Clinic. All Rights Reserved.
See Online Privacy | Terms of Use | e-mail Webmaster

http://research.marshfieldclinic.org/genetics/MarkerSearch/searchMarkers.asp

10th worst graph

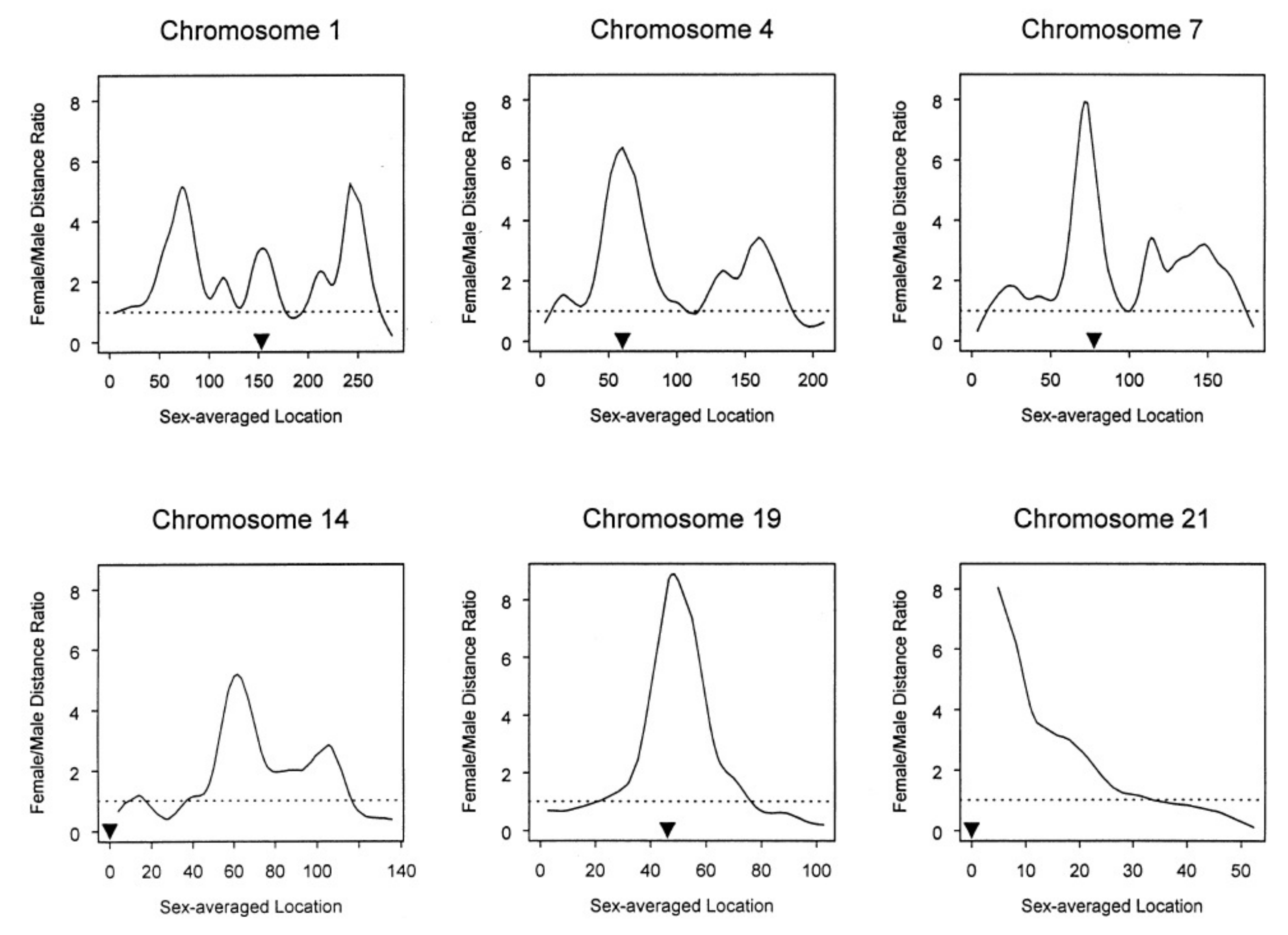
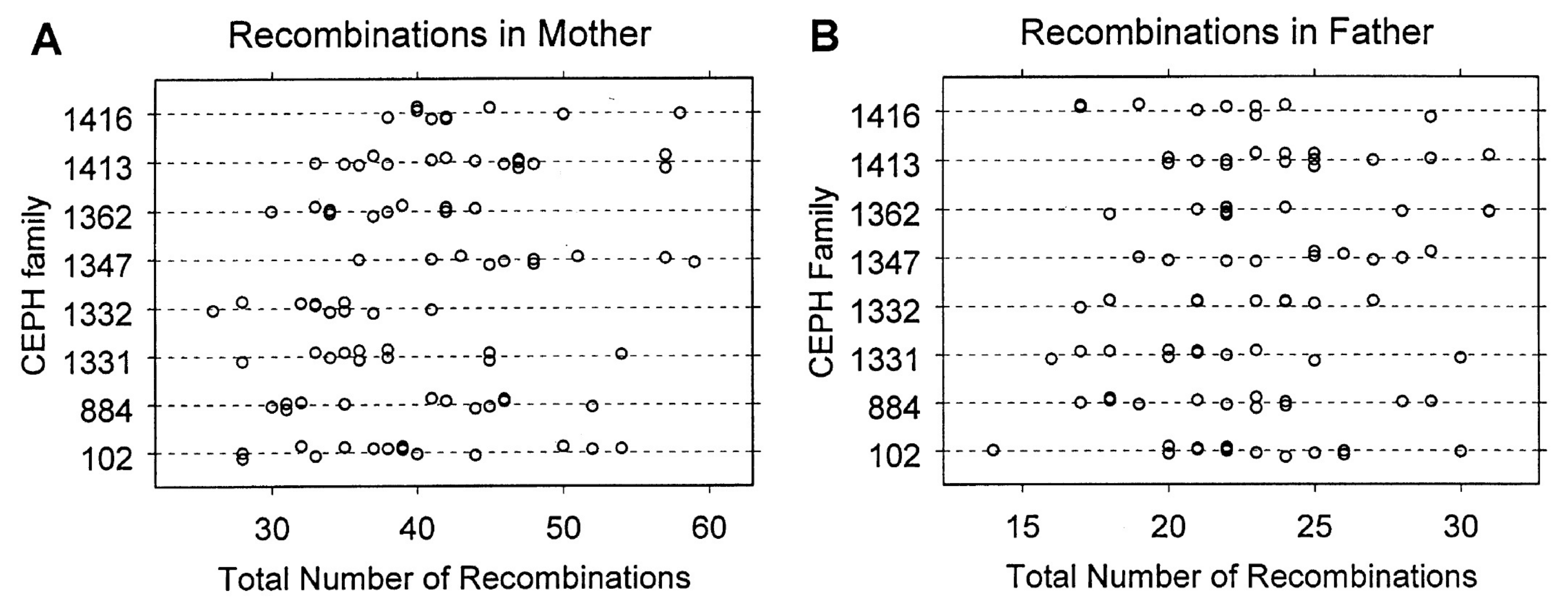


Figure 1 Plots of the female:male genetic-distance ratio against sex-averaged genetic location (in cM) along six selected chromosomes. Approximate locations of the centromeres are indicated by the triangles. The dashed lines correspond to equal female and male distances.

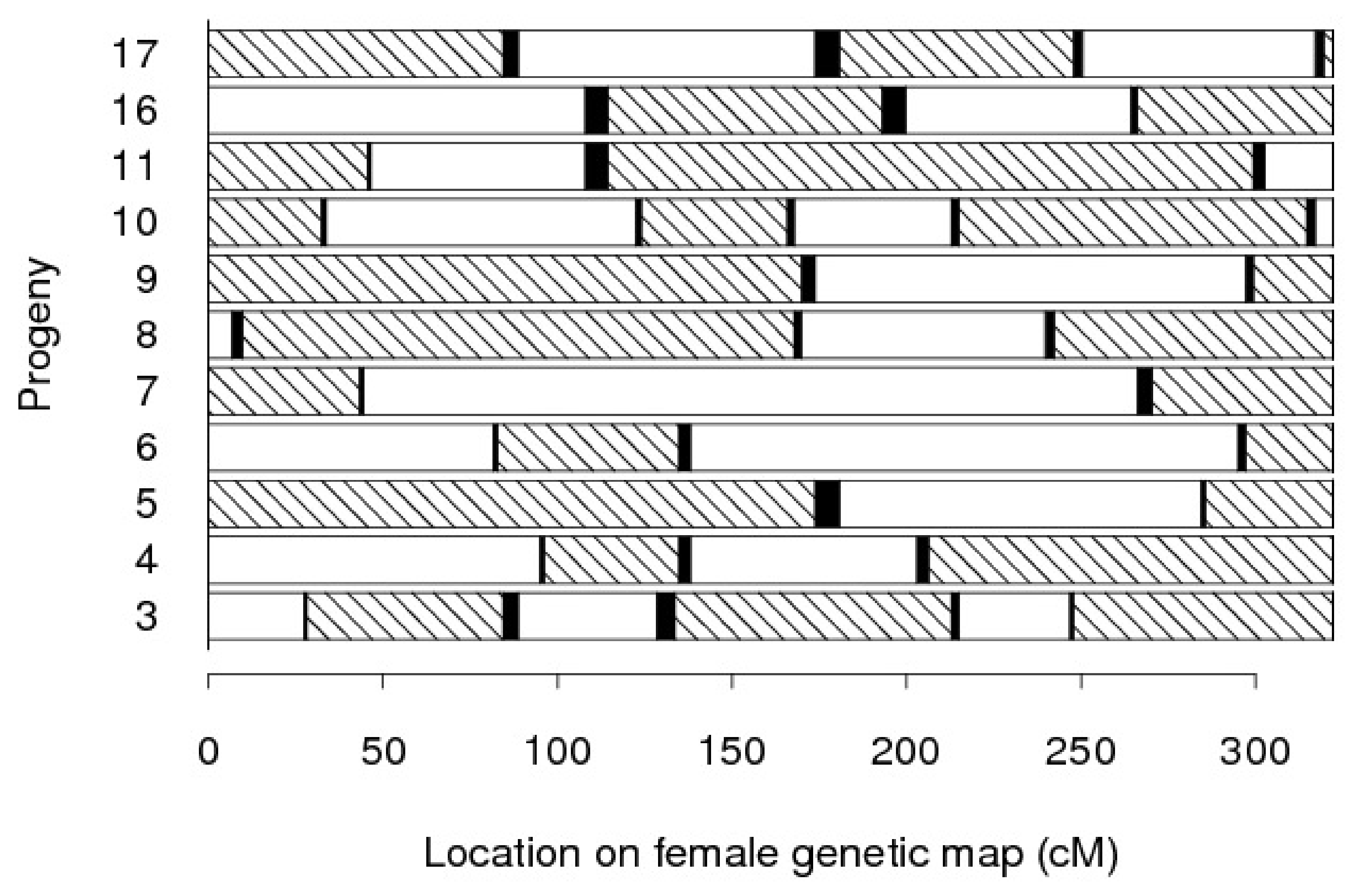
Broman et al., Am J Hum Genet 63: 861–869, 1998

Total no. crossovers



Broman et al., Am J Hum Genet 63: 861–869, 1998

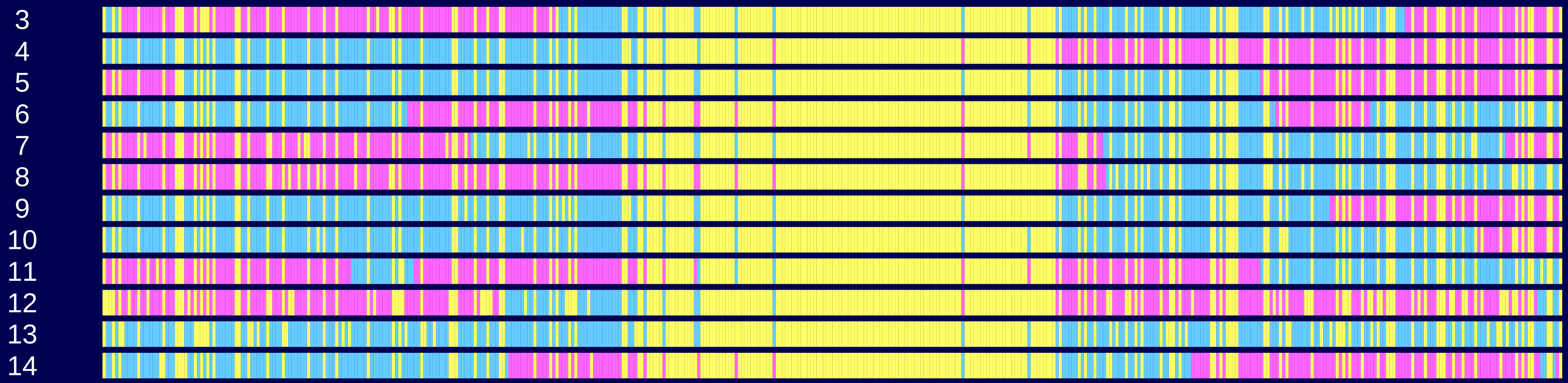
Crossover locations



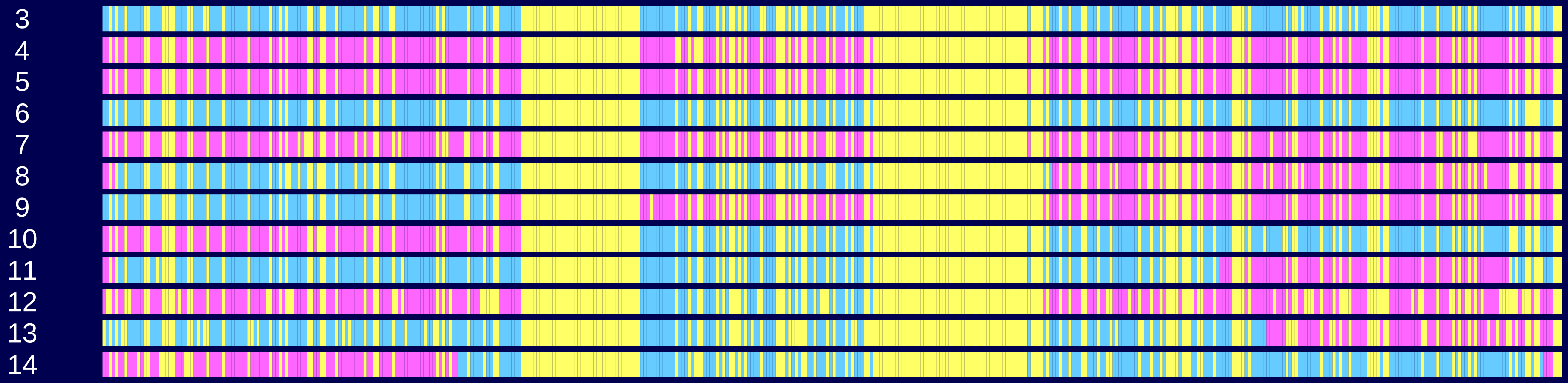
Broman and Weber, Am J Hum Genet 66:1911–1926, 2000

Family 884, chr 6

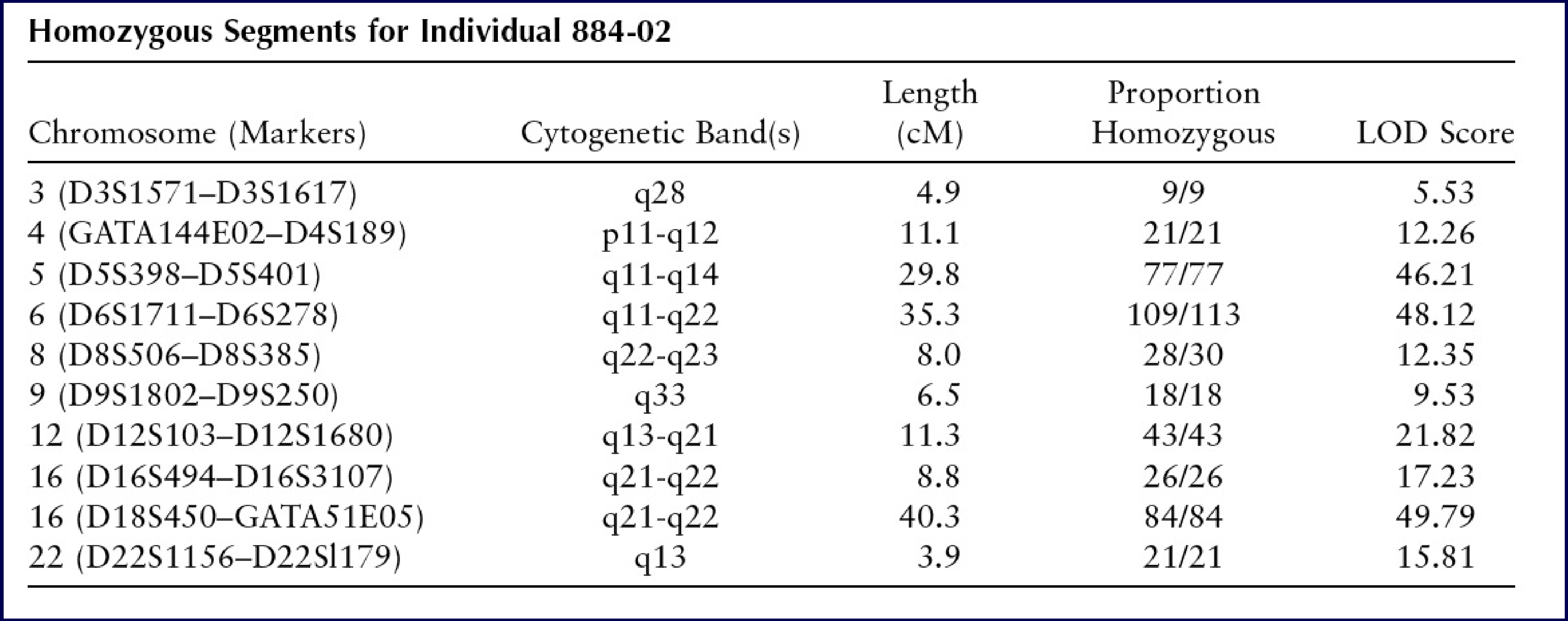
Maternal chromosomes



Paternal chromosomes



Autozygosity

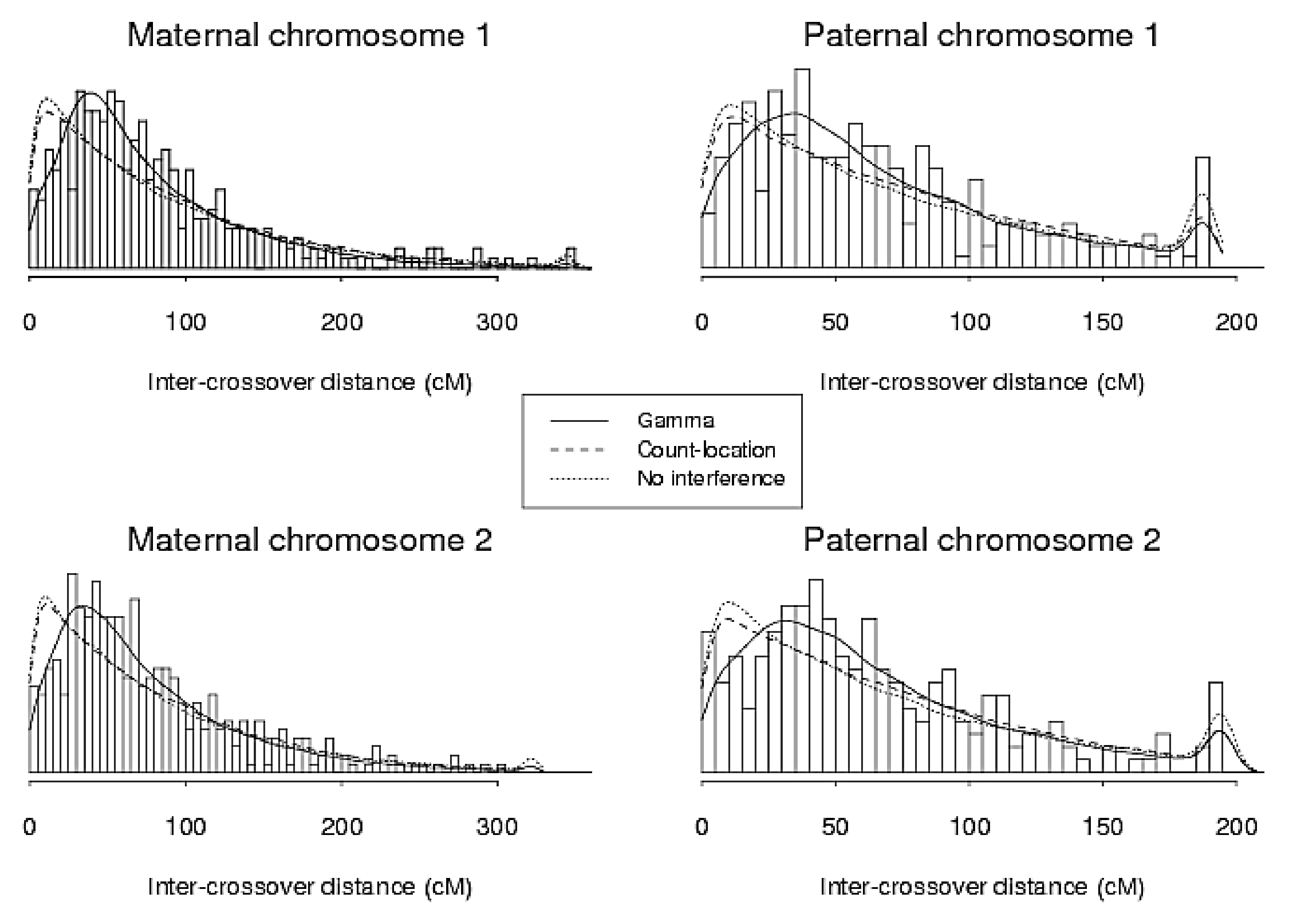


Homozygous Segments for Individual 884-02

Chromosome (Markers)	Cytogenetic Band(s)	Length (cM)	Proportion Homozygous	LOD Score
3 (D3S1571–D3S1617)	q28	4.9	9/9	5.53
4 (GATA144E02–D4S189)	p11-q12	11.1	21/21	12.26
5 (D5S398–D5S401)	q11-q14	29.8	77/77	46.21
6 (D6S1711–D6S278)	q11-q22	35.3	109/113	48.12
8 (D8S506–D8S385)	q22-q23	8.0	28/30	12.35
9 (D9S1802–D9S250)	q33	6.5	18/18	9.53
12 (D12S103–D12S1680)	q13-q21	11.3	43/43	21.82
16 (D16S494–D16S3107)	q21-q22	8.8	26/26	17.23
16 (D18S450–GATA51E05)	q21-q22	40.3	84/84	49.79
22 (D22S1156–D22Sl179)	q13	3.9	21/21	15.81

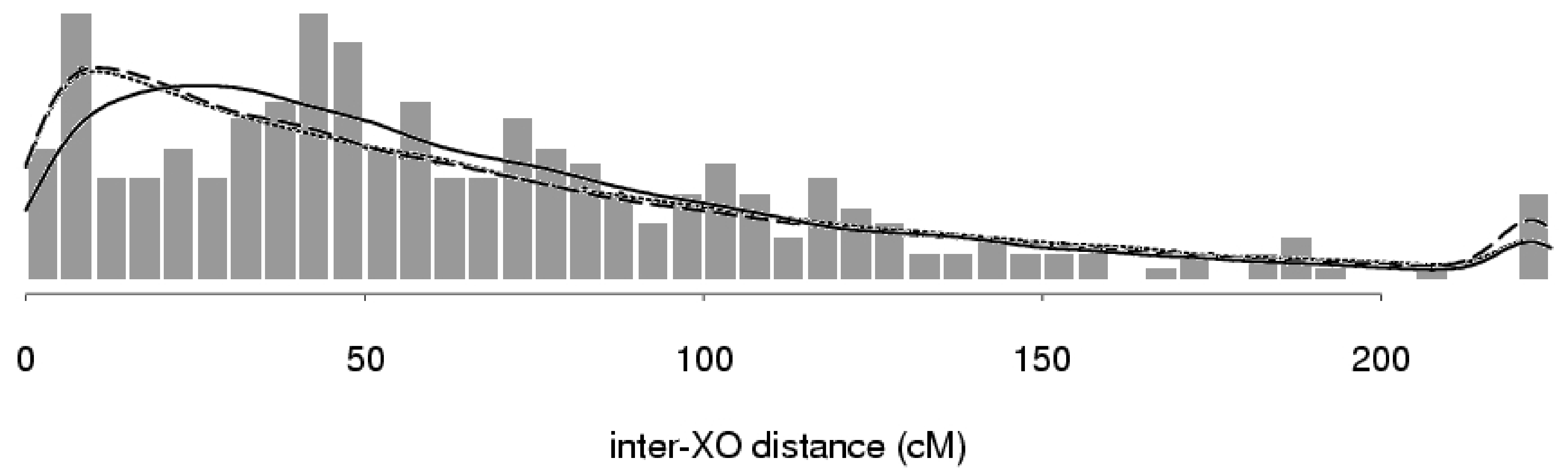
Broman and Weber, Am J Hum Genet 65:1493–1500, 1999

Crossover interference

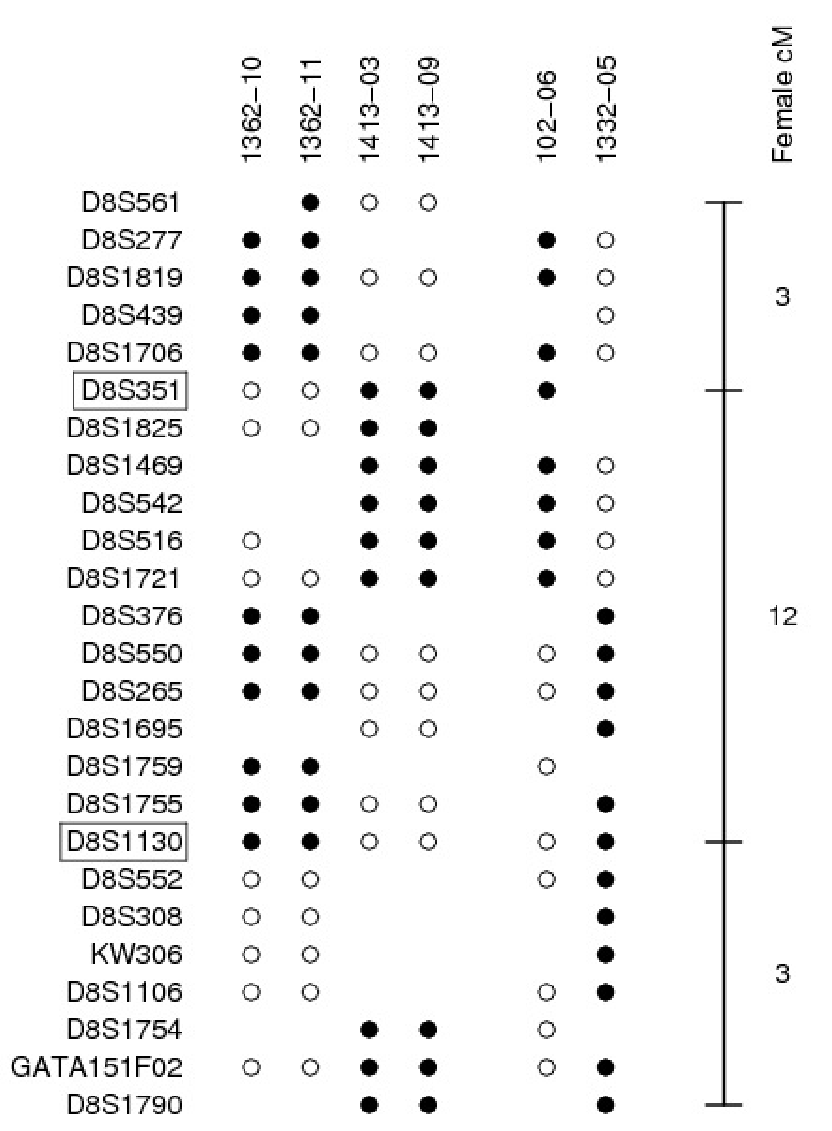


Broman and Weber, Am J Hum Genet 66:1911–1926, 2000

Maternal chr 8



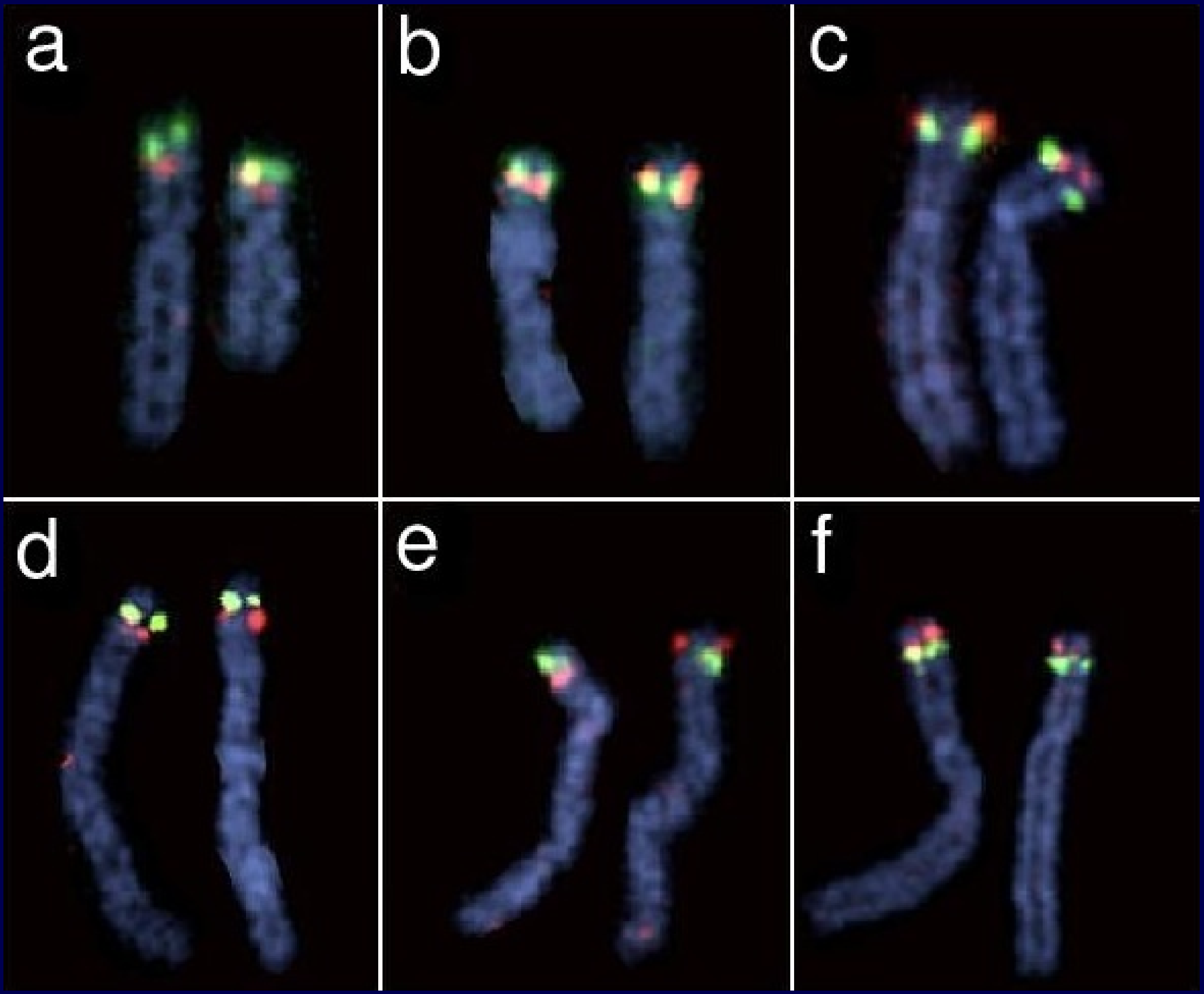
Apparent triple XOs



	1362–10	1362–11	1413–03	1413–09	102–06	1332–05	Female cM
D8S561		●	○	○			
D8S277	●	●			●	○	
D8S1819	●	●	○	○	●	○	
D8S439	●	●				○	3
D8S1706	●	●	○	○	●	○	
D8S351	○	○	●	●	●		
D8S1825	○	○	●	●			
D8S1469			●	●	●	○	
D8S542			●	●	●	○	
D8S516	○		●	●	●	○	
D8S1721	○	○	●	●	●	○	
D8S376	●	●				●	12
D8S550	●	●	○	○	○	●	
D8S265	●	●	○	○	○	●	
D8S1695			○	○		●	
D8S1759	●	●			○		
D8S1755	●	●	○	○		●	
D8S1130	●	●	○	○	○	●	
D8S552	○	○			○	●	
D8S308	○	○				●	
KW306	○	○				●	
D8S1106	○	○			○	●	3
D8S1754			●	●	○		
GATA151F02	○	○	●	●	○	●	
D8S1790			●	●		●	

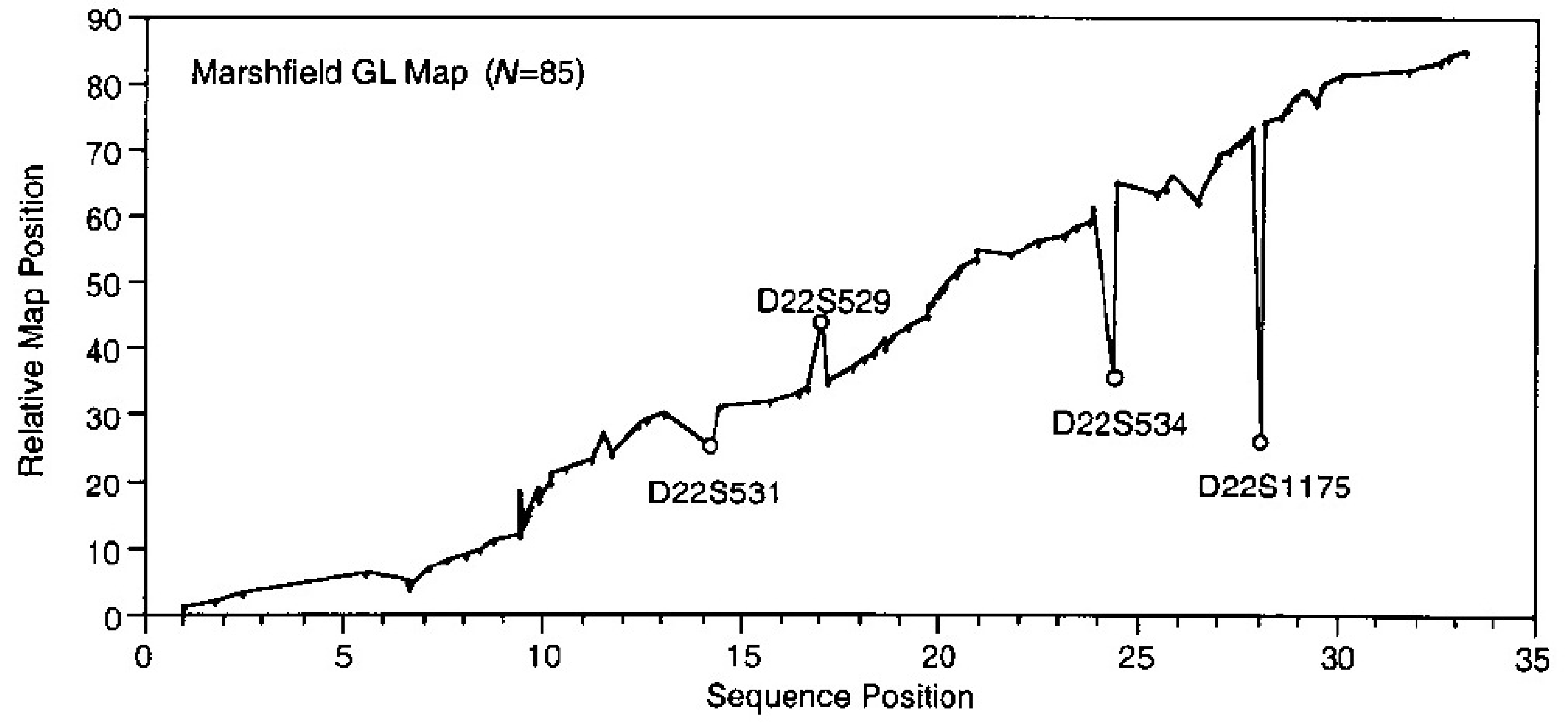
Broman et al., In: *Science and Statistics: A Festschrift for Terry Speed*, 2003

Chr 8p inversion



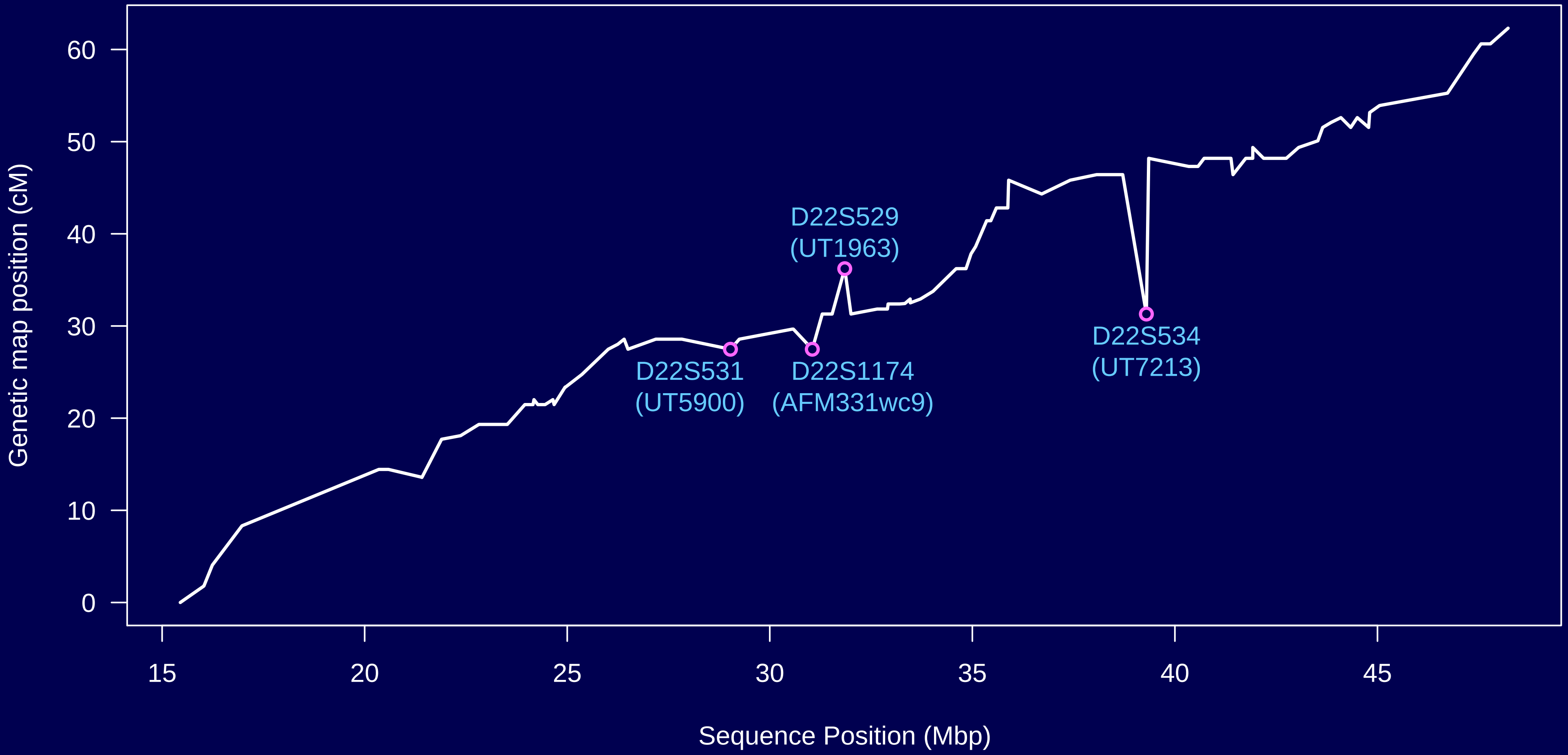
Broman et al., In: *Science and Statistics: A Festschrift for Terry Speed*, 2003

Comparison to sequence



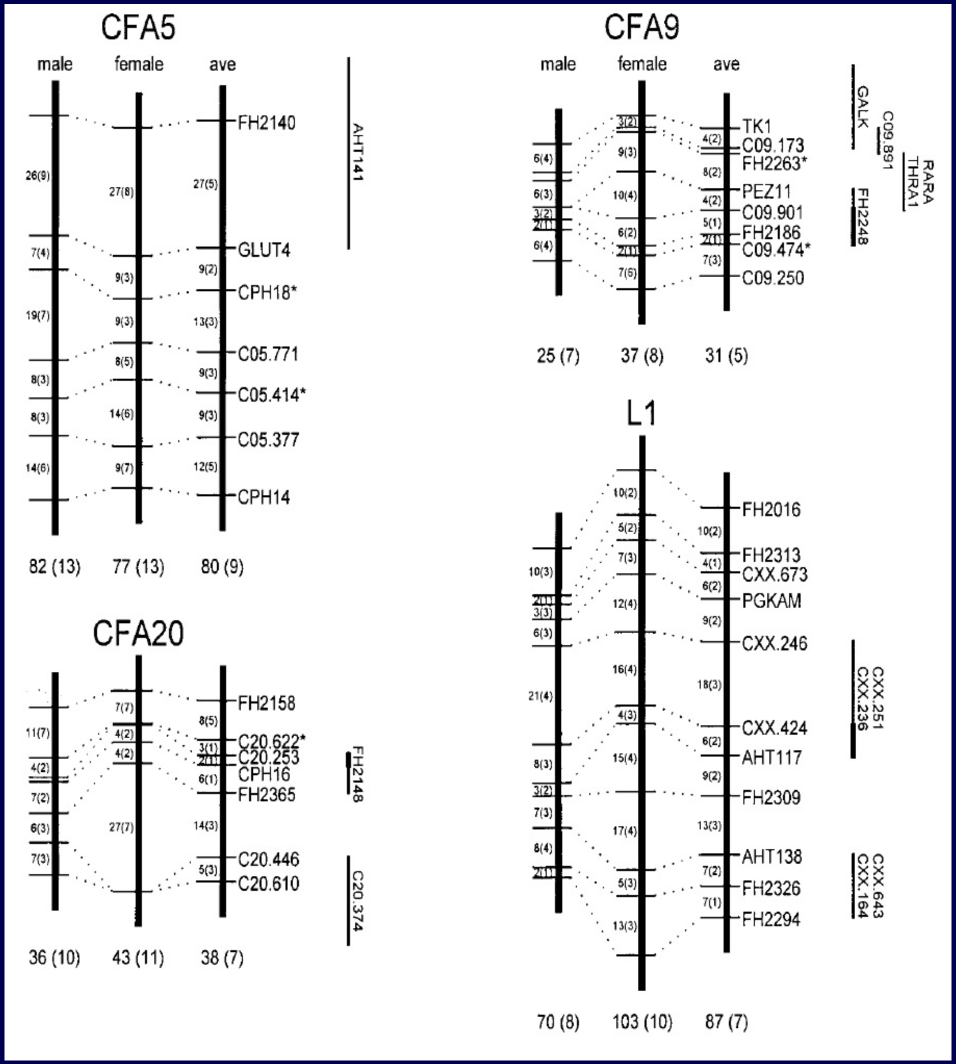
Matise et al., Am J Hum Genet 70:1398–1410, 2002

Comparison to sequence
(revisited)



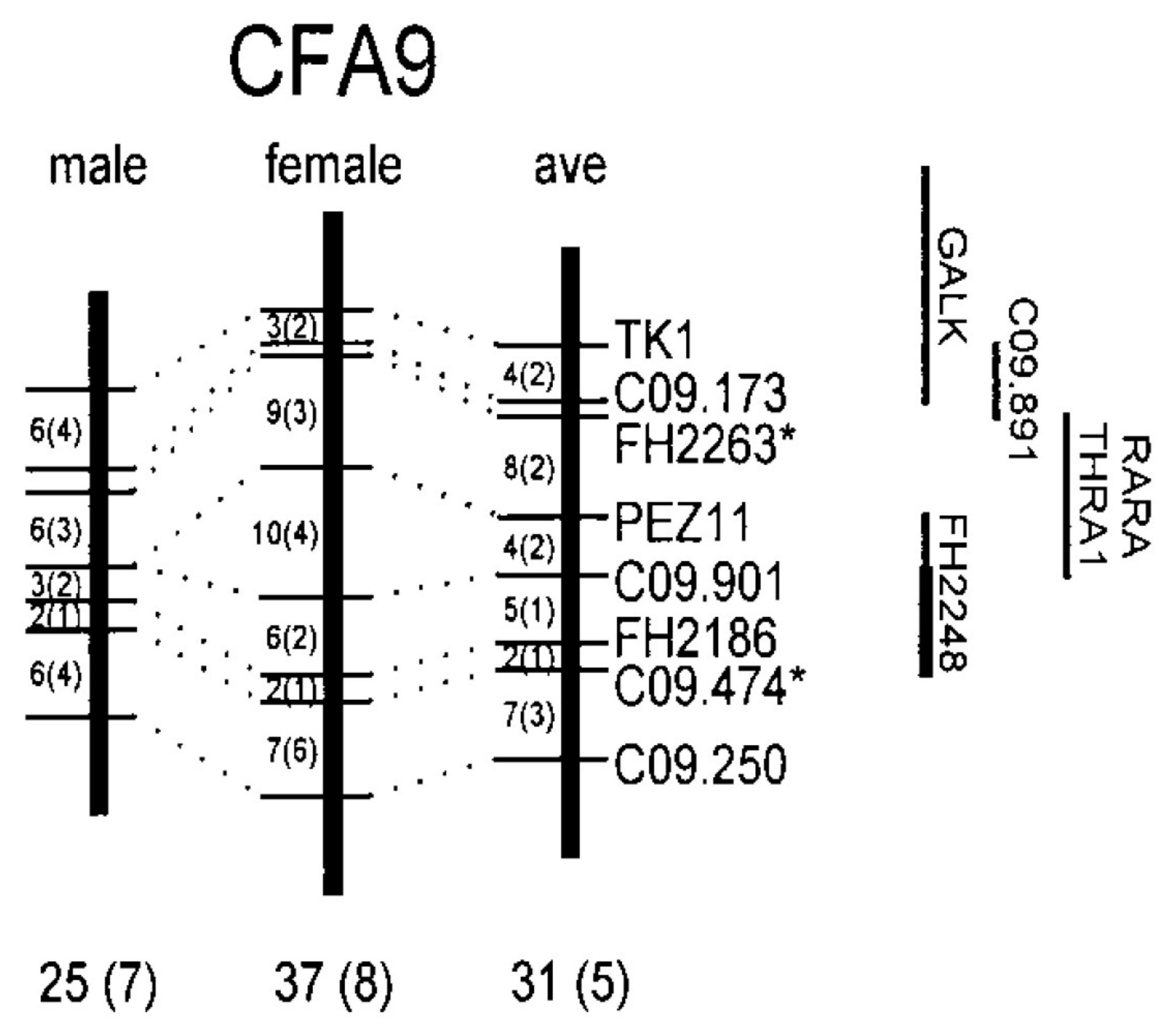
Thanks to UCSC and Ensembl!

Dog genetic map



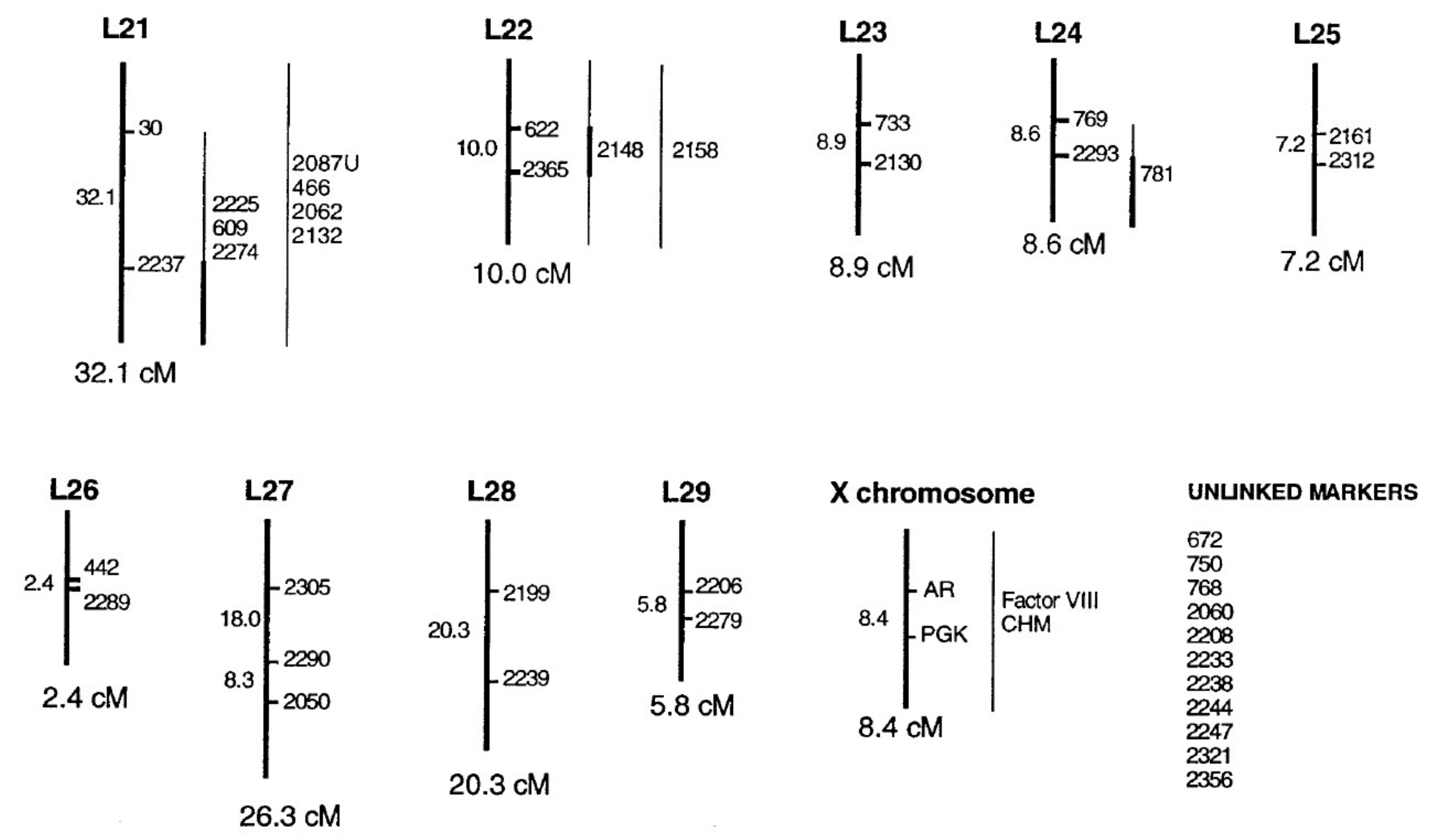
Neff et al., Genetics 151:803–820, 1999

Dog genetic map



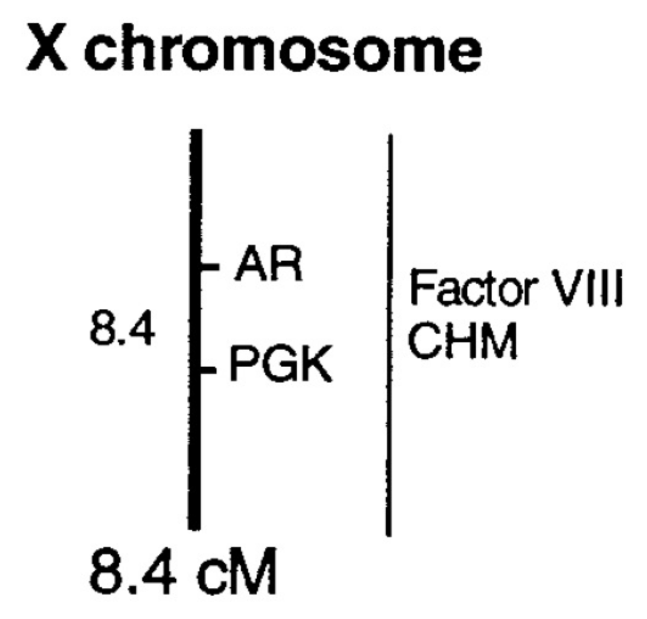
Neff et al., Genetics 151:803–820, 1999

The first dog map



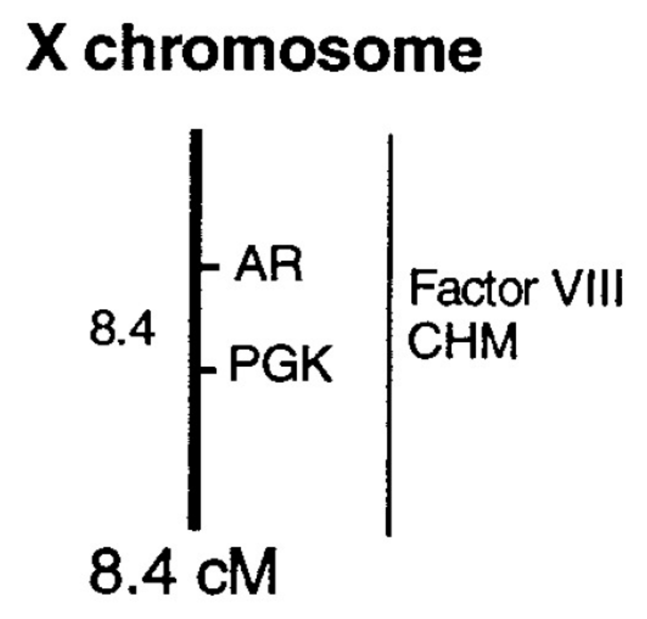
Mellersh et al., Genomics 46:326–336, 1997

The dog X

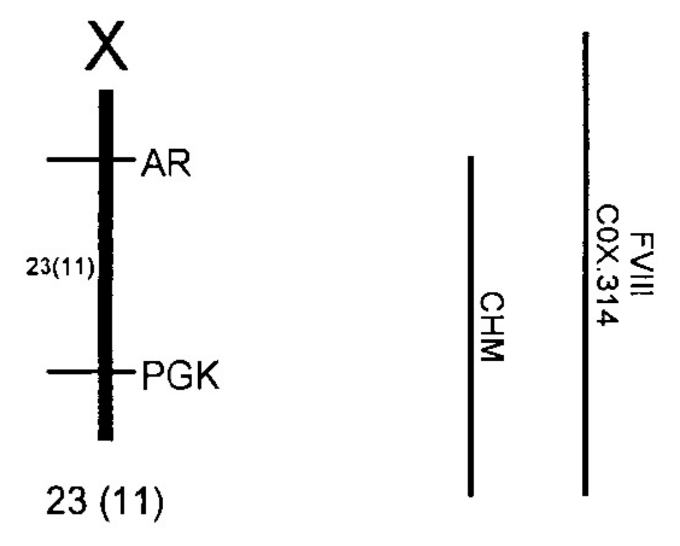


Mellersh et al., Genomics 46:326–336, 1997

The dog X



Mellersh et al., Genomics 46:326–336, 1997



Neff et al., Genetics 151:803–820, 1999

JAX backcrosses

(C57BL/6J × SPRET/Ei) × SPRET/Ei

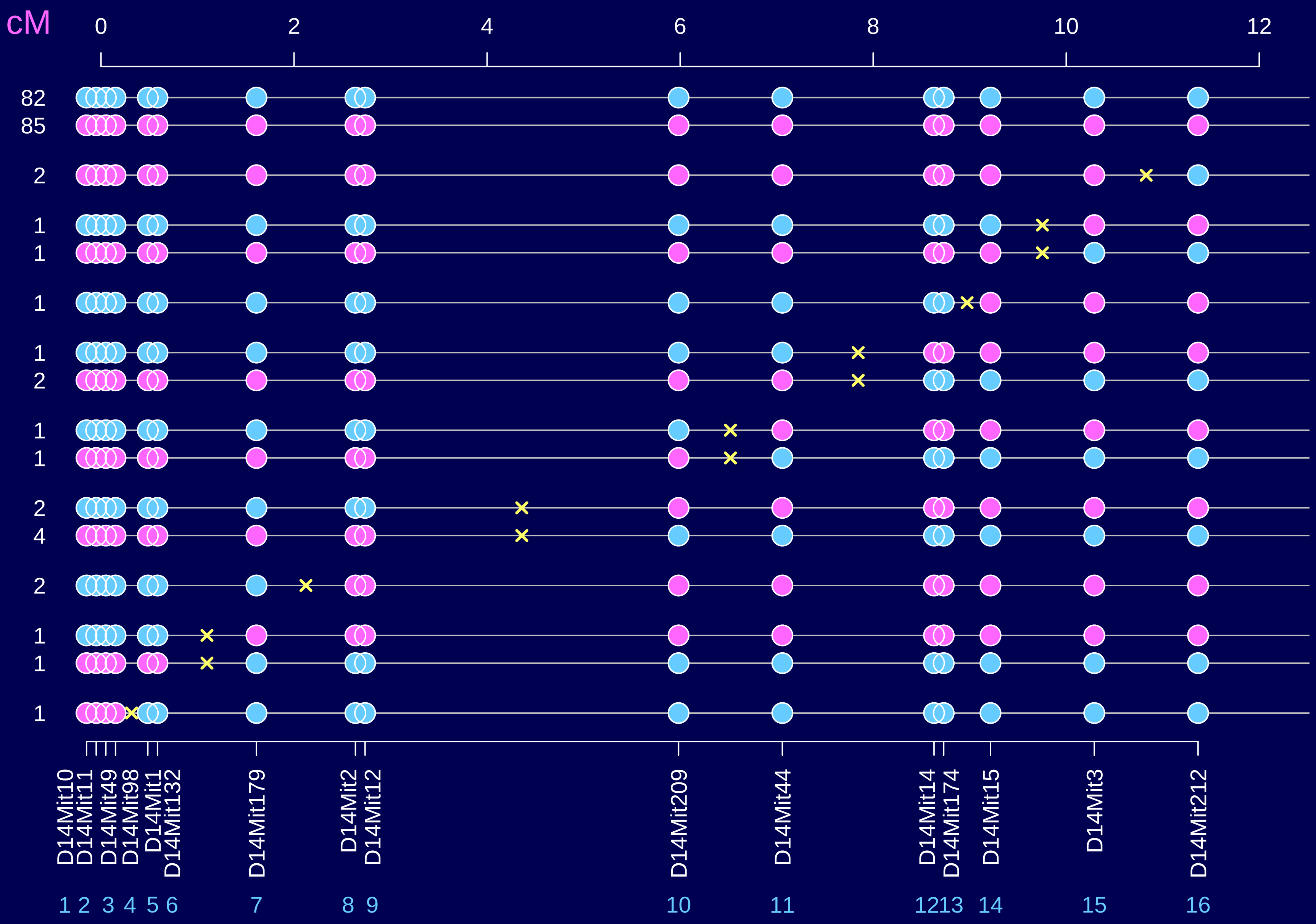
- 94 individuals
- 4913 markers

(C57BL/6J × SPRET/Ei) × C57BL/6J

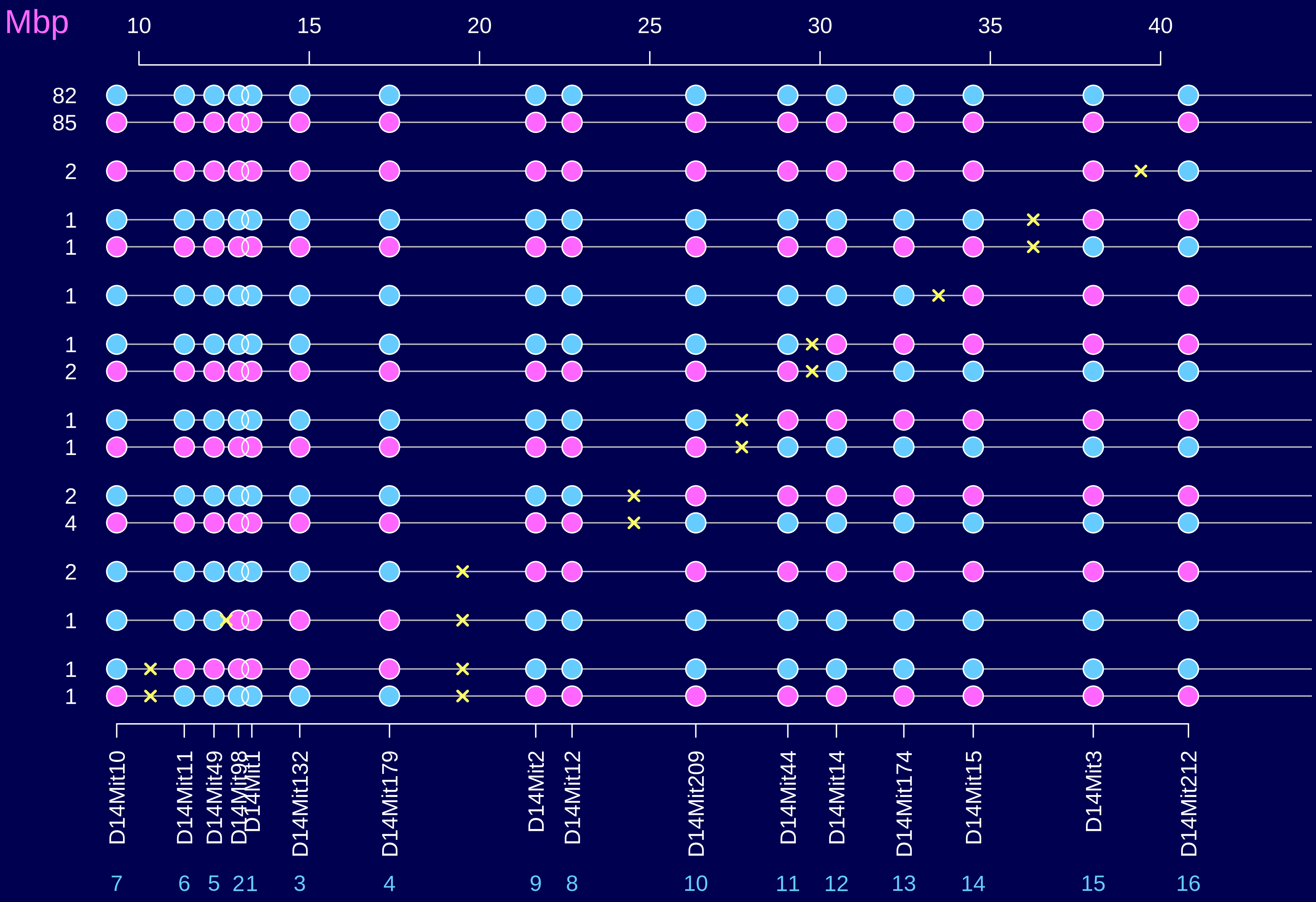
- 94 individuals
- 1659 markers

Rowe et al., Mamm Genome 5:253–274, 1994
Broman et al., Genetics 160:1123–1131, 2002

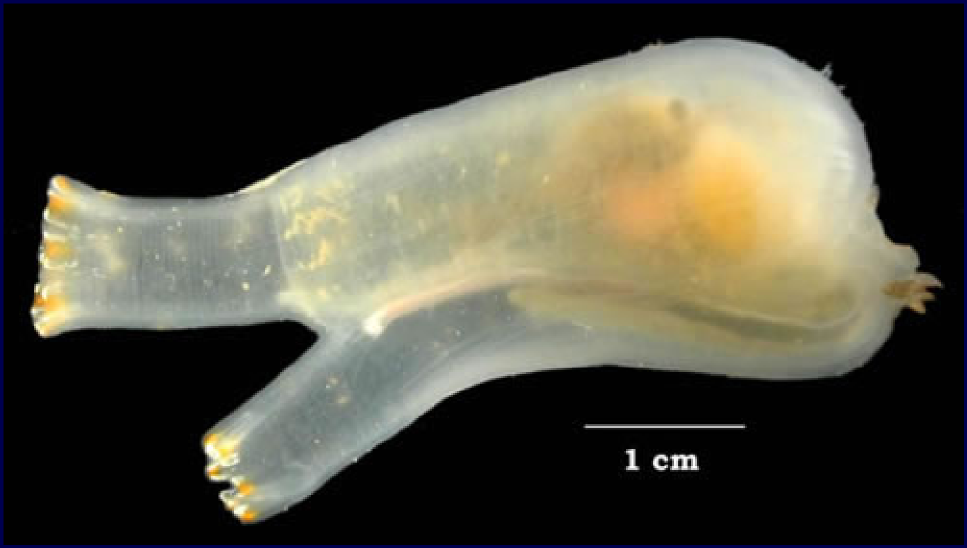
Mouse chr 14p



Mouse chr 14p

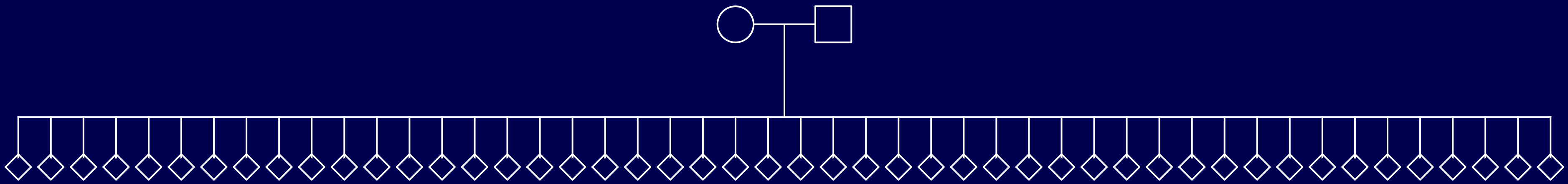


C. savignyi (sea squirt)



Aim: Build linkage map to improve long-range ordering of draft genome sequence (which is currently represented as a few hundred “reftigs”)

C. savignyi pedigree



48 progeny

Markers: PCR amplicon (primers in exons, spanning an intron), digested with a restriction enzyme

⟶ 2, 3 or 4 banding patterns

Tricky bits: Which marker “phenotype” corresponds to which genotype?

Using information on locations of markers within “reftigs”

Example of two markers

	99481-HaeIII			
114467-MboI	AC	AD	BC	BC
AC	8	3	1	0
AD	0	10	0	1
BC	2	0	11	1
BD	0	1	0	9

Example of two markers

	99481-HaeIII			
114467-MboI	AC	AD	BC	BC
AC	8	3	1	0
AD	0	10	0	1
BC	2	0	11	1
BD	0	1	0	9

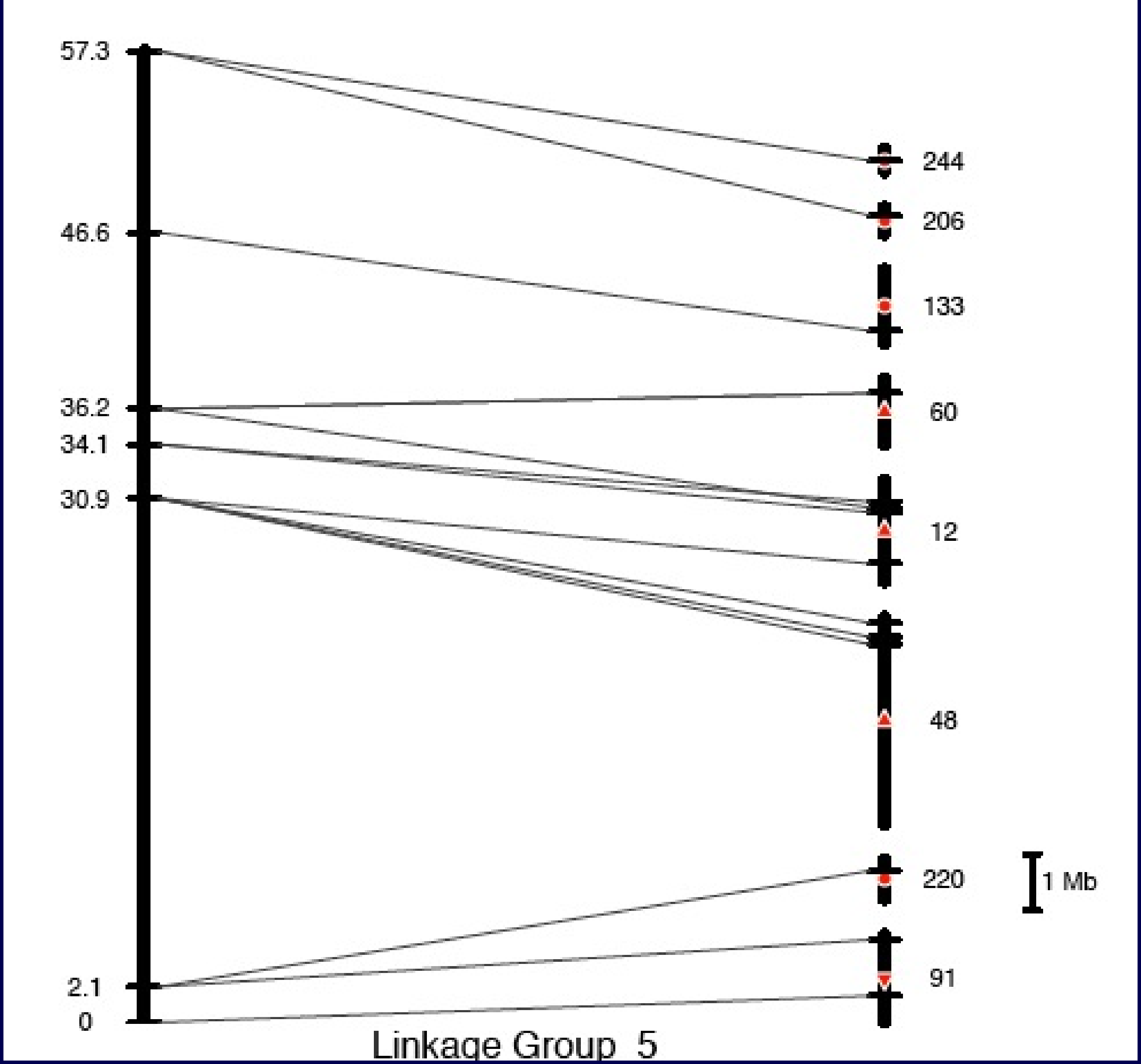
Example of two markers

	99481-HaeIII			
114467-MboI	AC	AD	BC	BC
AC	8	3	1	0
AD	0	10	0	1
BC	2	0	11	1
BD	0	1	0	9

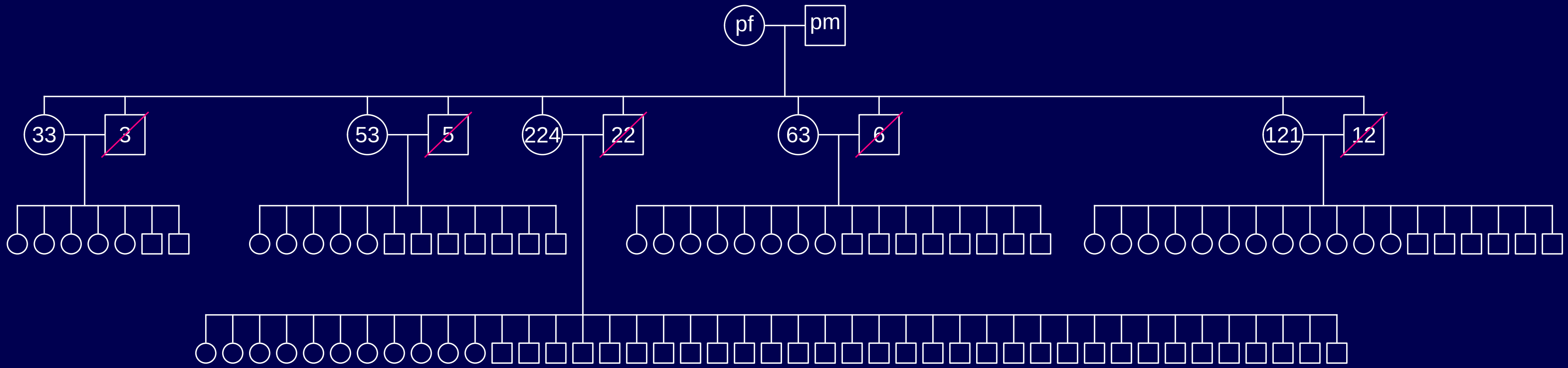
Example of two markers

	99481-HaeIII			
114467-MboI	AC	AD	BC	BC
AC	8	3	1	0
AD	0	10	0	1
BC	2	0	11	1
BD	0	1	0	9

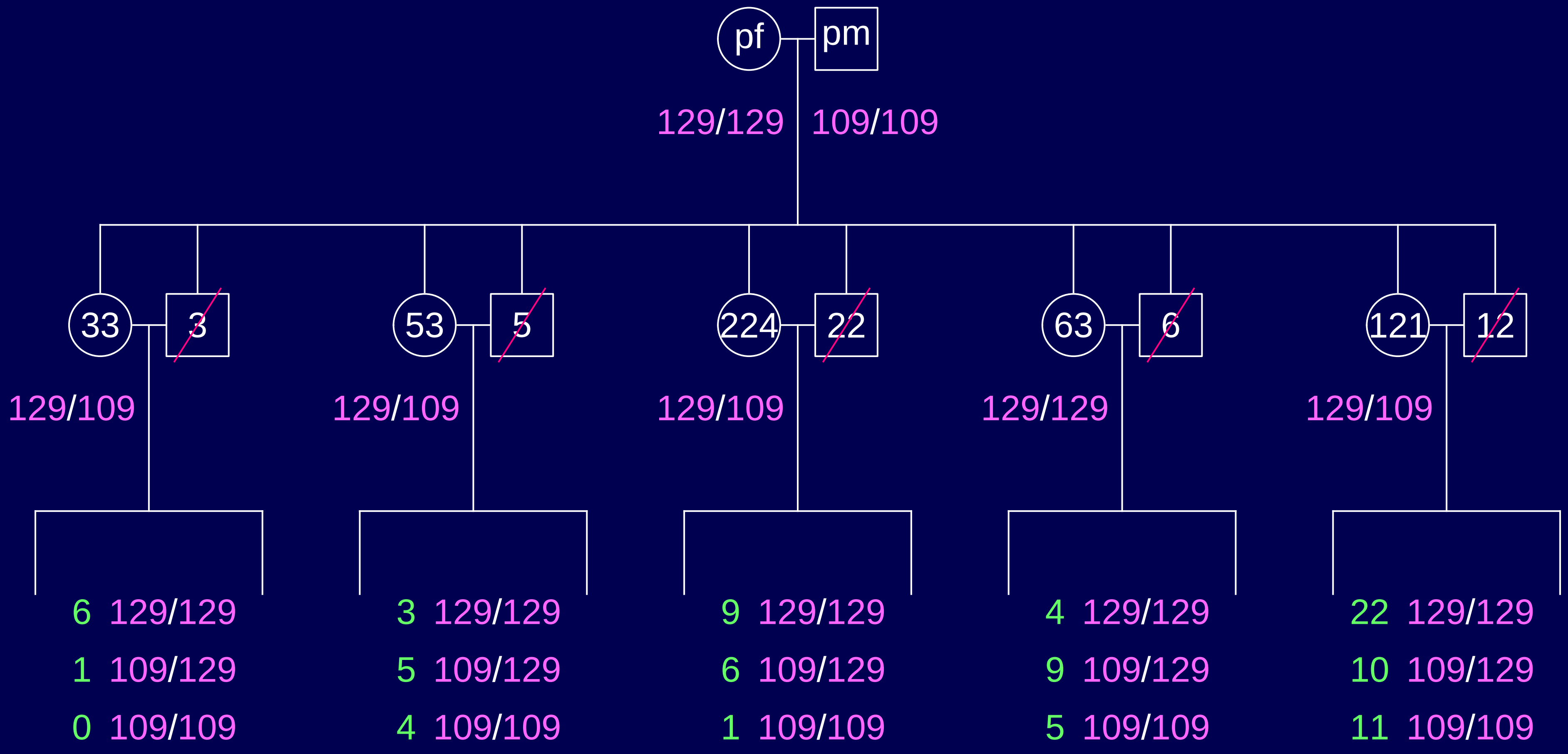
C. savignyi LG 5



Culex tarsalis pedigree



Marker TB1



pf pm

129/129 109/109

33 ~~3~~
129/109
6 129/129
1 109/129
0 109/109

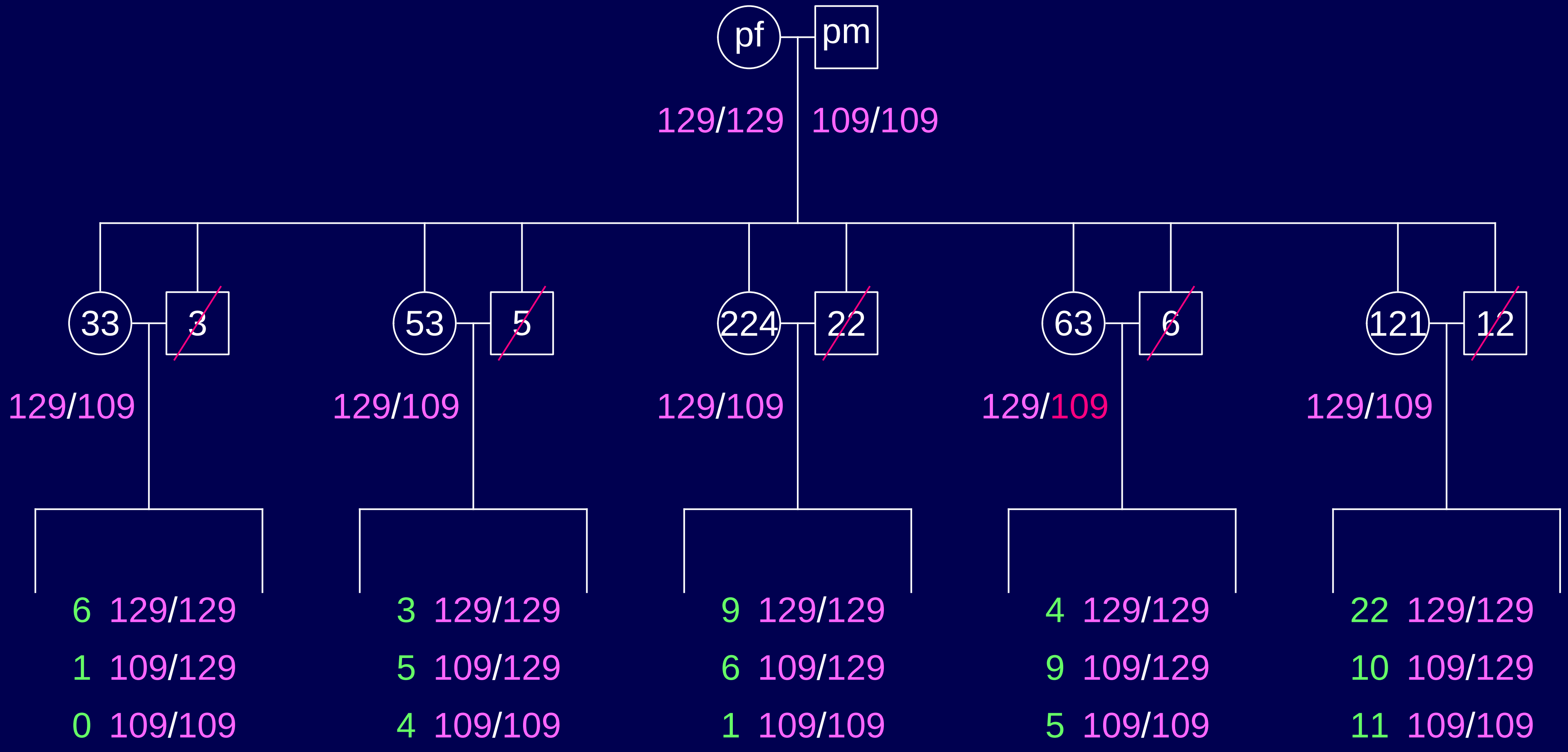
53 ~~5~~
129/109
3 129/129
5 109/129
4 109/109

224 ~~22~~
129/109
9 129/129
6 109/129
1 109/109

63 ~~6~~
129/129
4 129/129
9 109/129
5 109/109

121 ~~12~~
129/109
22 129/129
10 109/129
11 109/109

Marker TB1



pf pm

129/129 109/109

33 ~~3~~

129/109

6 129/129
1 109/129
0 109/109

53 ~~5~~

129/109

3 129/129
5 109/129
4 109/109

224 ~~22~~

129/109

9 129/129
6 109/129
1 109/109

63 ~~6~~

129/109

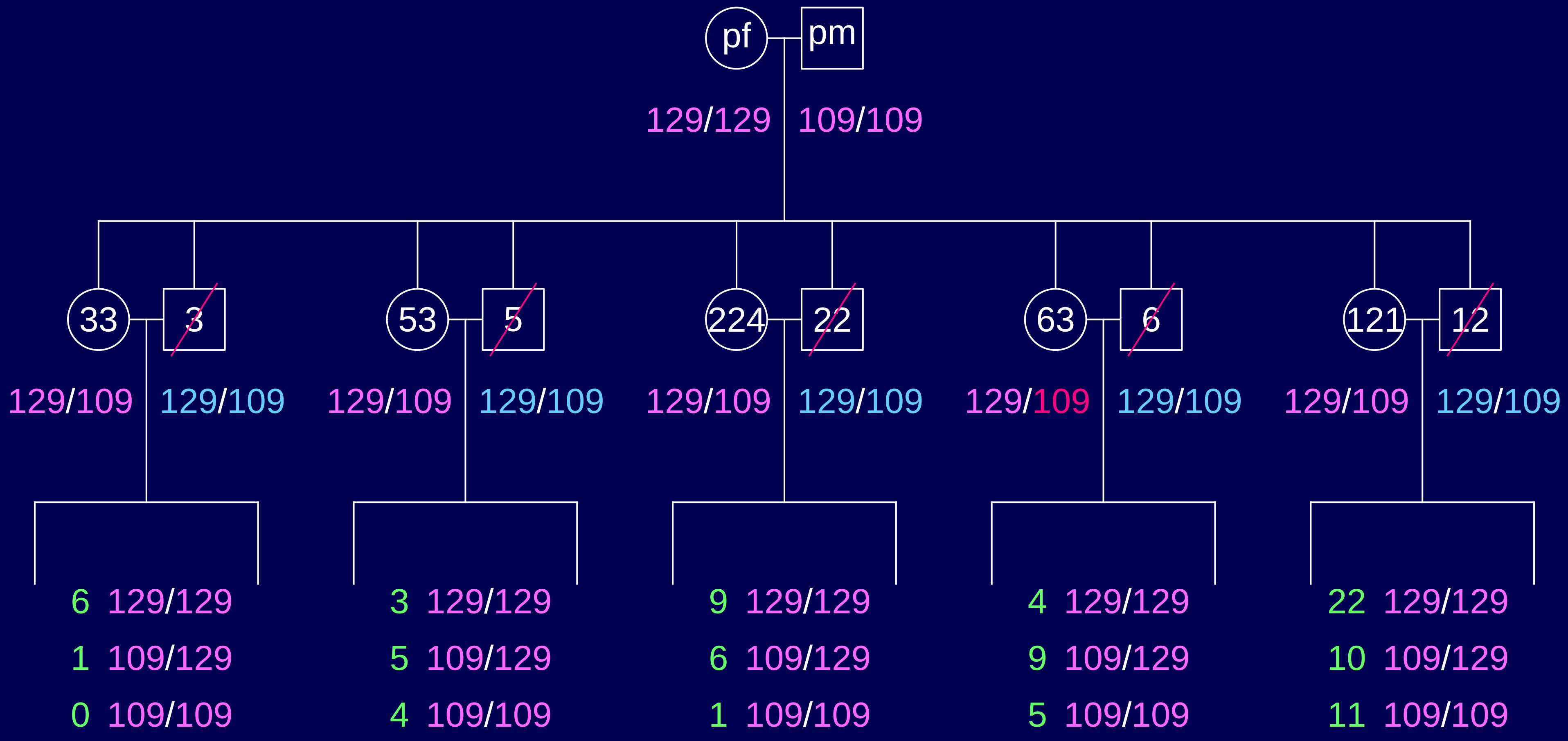
4 129/129
9 109/129
5 109/109

121 ~~12~~

129/109

22 129/129
10 109/129
11 109/109

Marker TB1



pf pm

129/129 109/109

33 ~~3~~
129/109 129/109
6 129/129
1 109/129
0 109/109

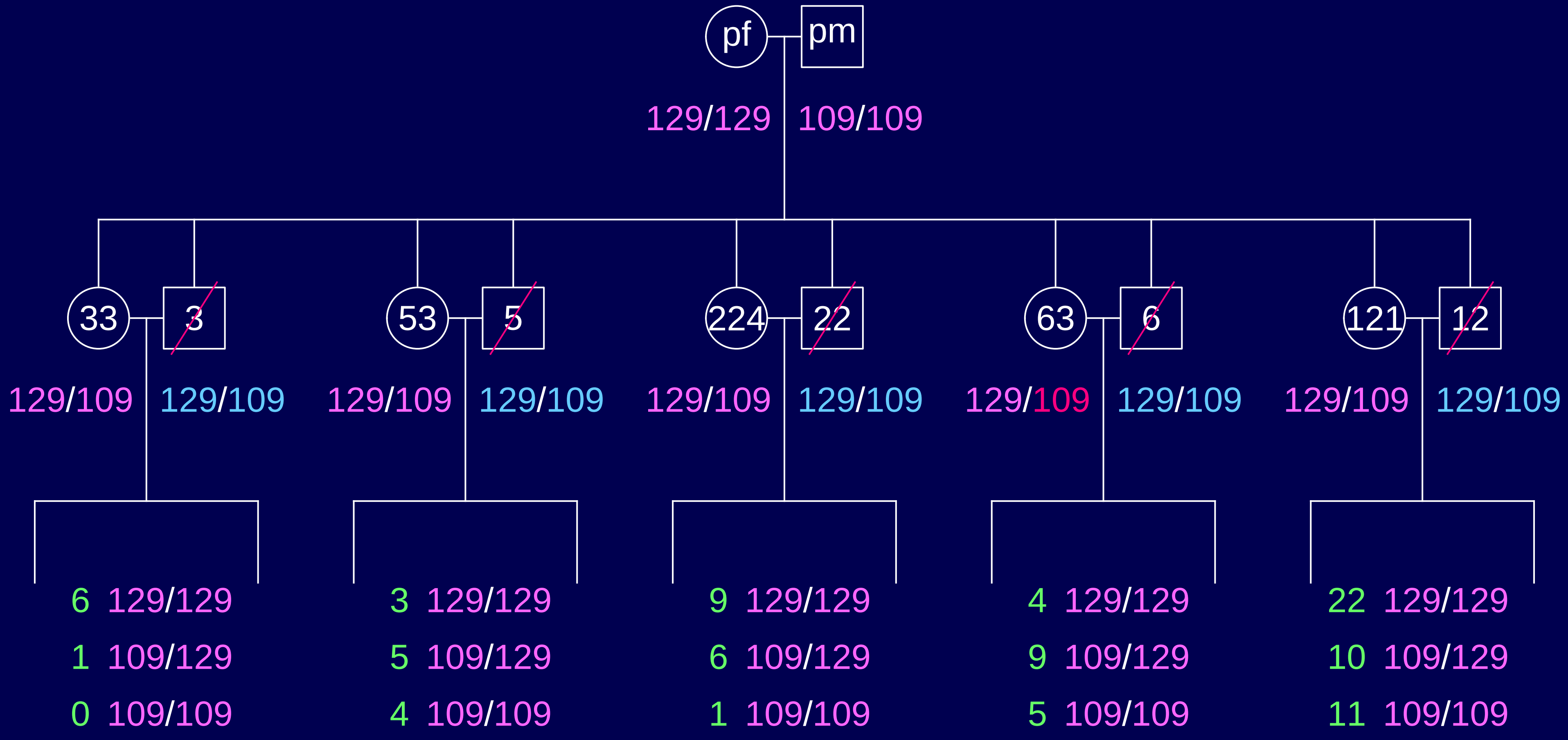
53 ~~5~~
129/109 129/109
3 129/129
5 109/129
4 109/109

224 ~~22~~
129/109 129/109
9 129/129
6 109/129
1 109/109

63 ~~6~~
129/109 129/109
4 129/129
9 109/129
5 109/109

121 ~~12~~
129/109 129/109
22 129/129
10 109/129
11 109/109

Marker TB1



pf pm

129/129 109/109

33 ~~3~~
129/109 129/109
6 129/129
1 109/129
0 109/109

53 ~~5~~
129/109 129/109
3 129/129
5 109/129
4 109/109

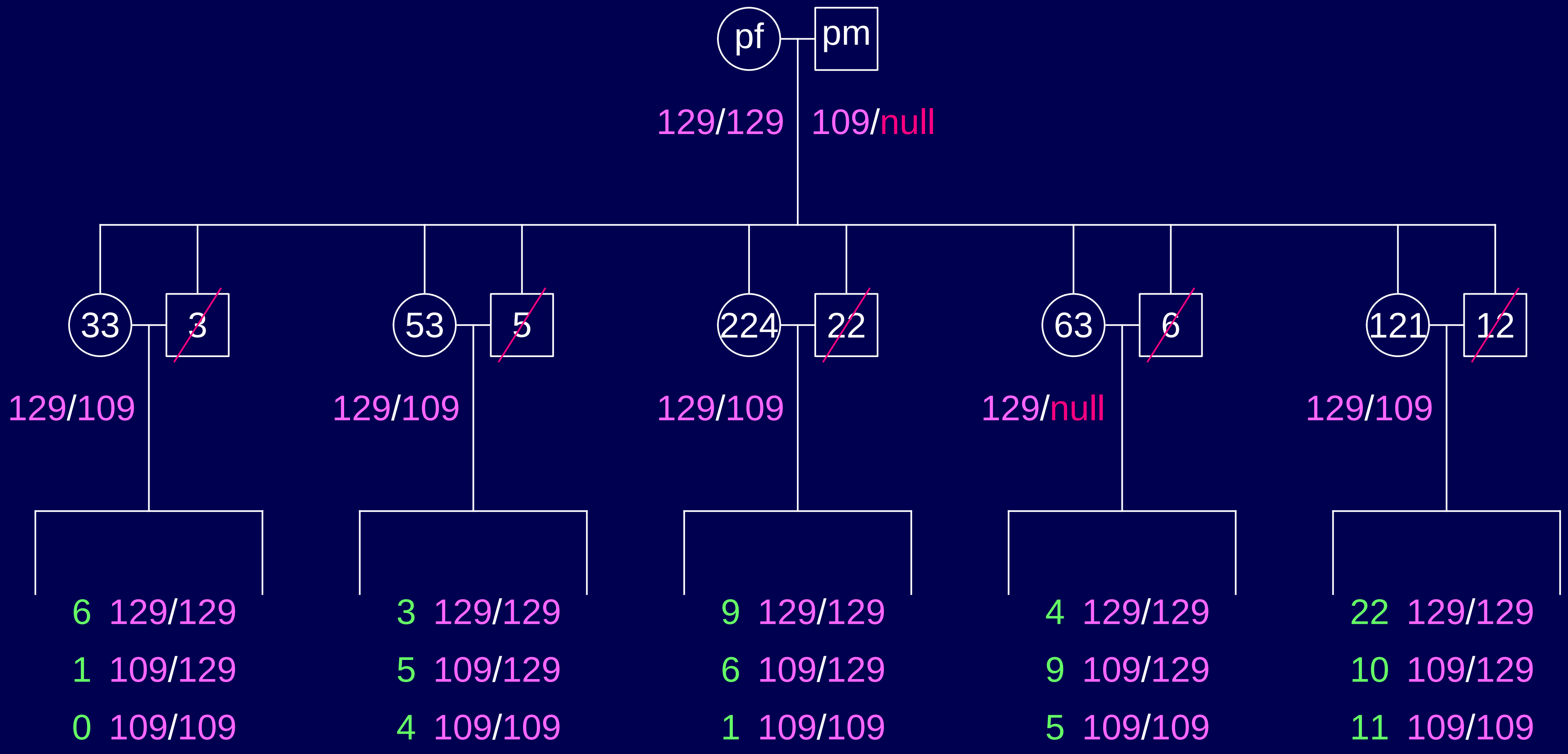
224 ~~22~~
129/109 129/109
9 129/129
6 109/129
1 109/109

63 ~~6~~
129/109 129/109
4 129/129
9 109/129
5 109/109

121 ~~12~~
129/109 129/109
22 129/129
10 109/129
11 109/109

Overall:
44 129/129
31 109/129
21 109/109

Marker TB1



pf pm

129/129 109/null

33 3 53 5 224 22 63 6 121 12

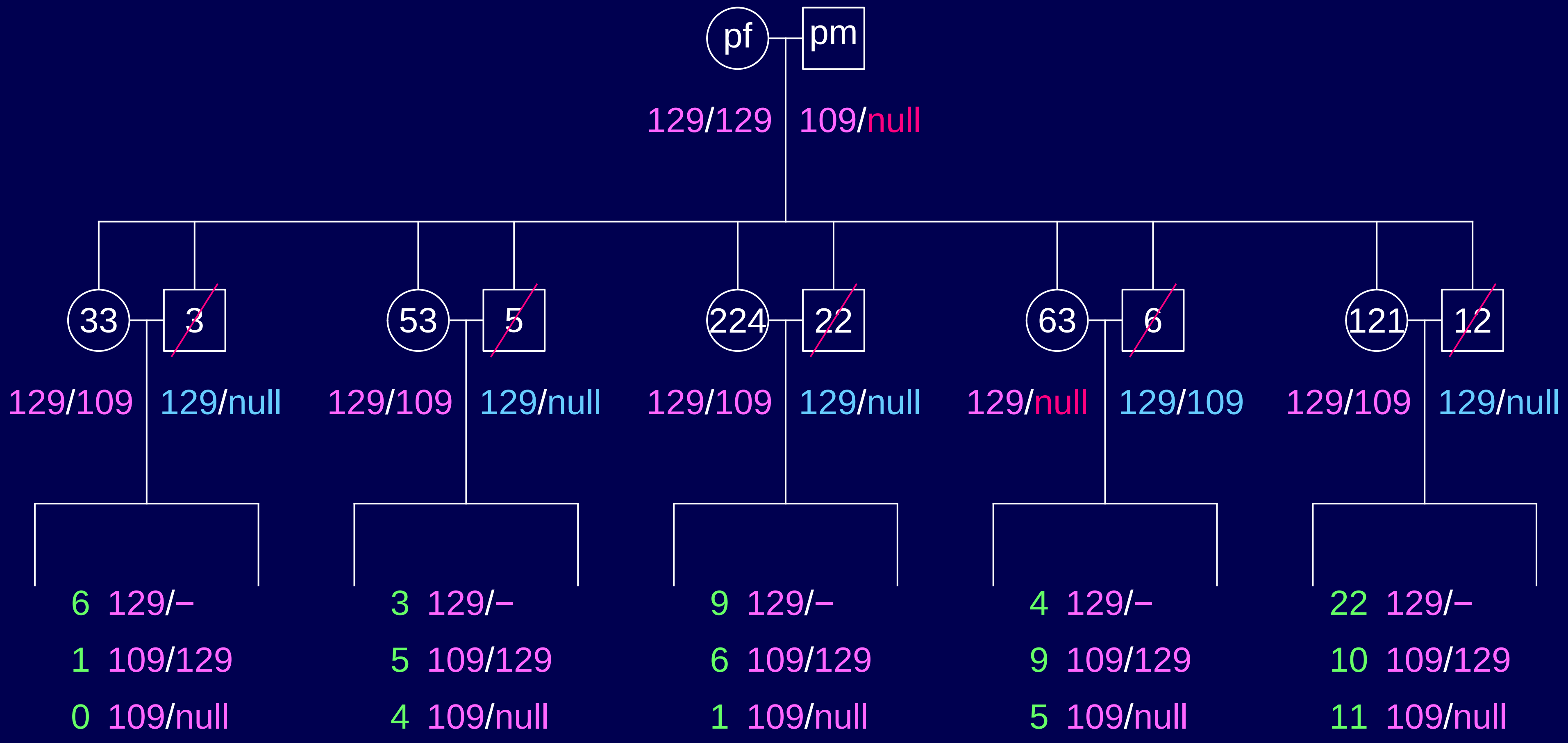
129/109 129/109 129/109 129/null 129/109

6 129/129 3 129/129 9 129/129 4 129/129 22 129/129

1 109/129 5 109/129 6 109/129 9 109/129 10 109/129

0 109/109 4 109/109 1 109/109 5 109/109 11 109/109

Marker TB1



pf pm

129/129 109/null

33 ~~3~~

129/109 129/null

6 129/−
1 109/129
0 109/null

53 ~~5~~

129/109 129/null

3 129/−
5 109/129
4 109/null

224 ~~22~~

129/109 129/null

9 129/−
6 109/129
1 109/null

63 ~~6~~

129/null 129/109

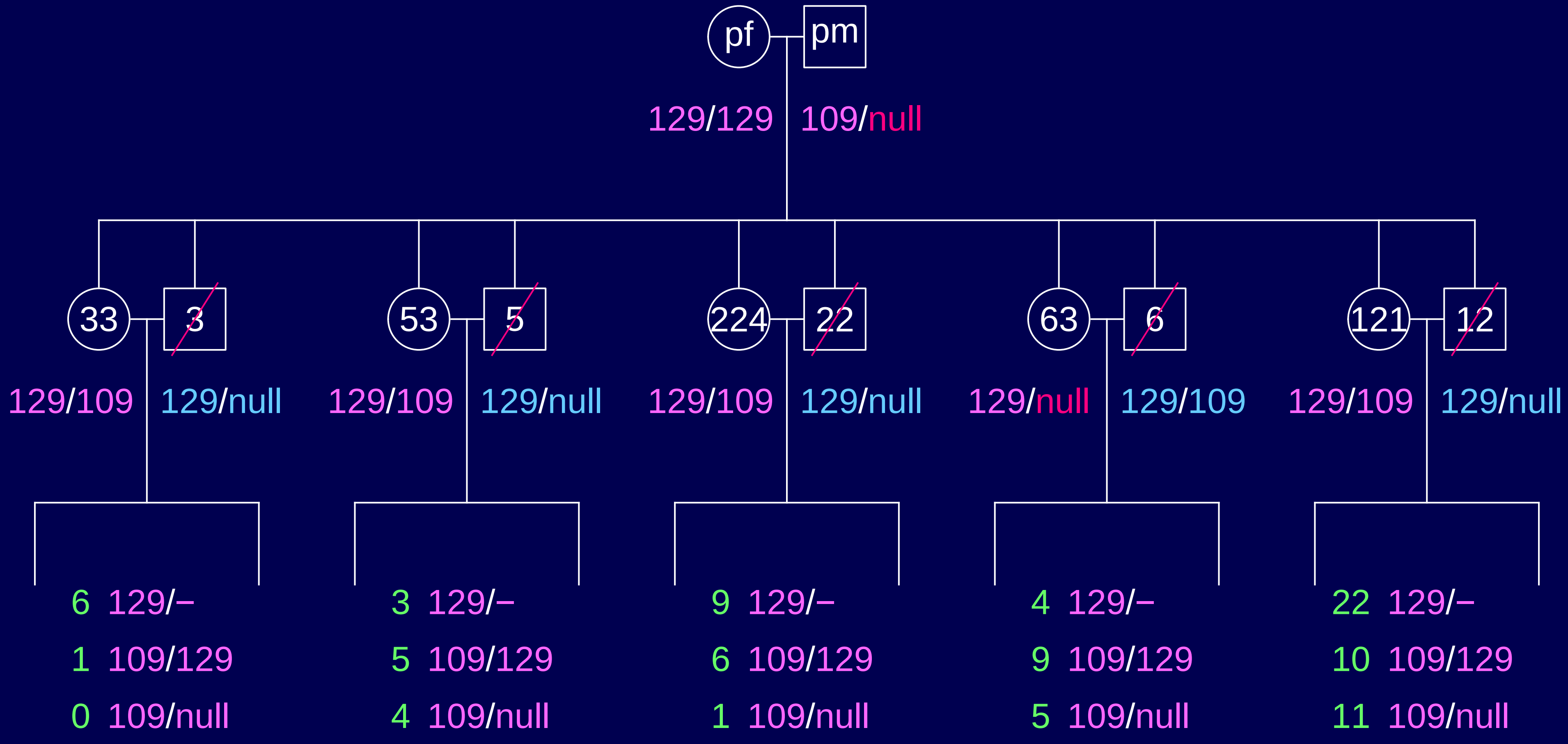
4 129/−
9 109/129
5 109/null

121 ~~12~~

129/109 129/null

22 129/−
10 109/129
11 109/null

Marker TB1



pf pm

129/129 109/null

33 ~~3~~
129/109 129/null

6 129/−
1 109/129
0 109/null

53 ~~5~~
129/109 129/null

3 129/−
5 109/129
4 109/null

224 ~~22~~
129/109 129/null

9 129/−
6 109/129
1 109/null

63 ~~6~~
129/null 129/109

4 129/−
9 109/129
5 109/null

121 ~~12~~
129/109 129/null

22 129/−
10 109/129
11 109/null

Overall:

44 129/−
31 109/129
21 109/null

Morals

- Genetic maps continue to be useful
- Be careful about automated map construction
- Careful, tedious work is often necessary
- The simplest things can have the greatest impact
- Artifacts can be more interesting than anything else
- Don’t give a sex-averaged map of the X chromosome
- Genotype the F_1 parents
- Null alleles are a pain

Acknowledgments

Terry Speed, Univ. California, Berkeley

Jim Weber, PreventionGenetics

Mark Neff, Univ. California, Davis

Lucy Rowe, The Jackson Lab

Matt Hill and Arend Sidow, Stanford

Meera Venkatesan and Jason Rasgon, Johns Hopkins