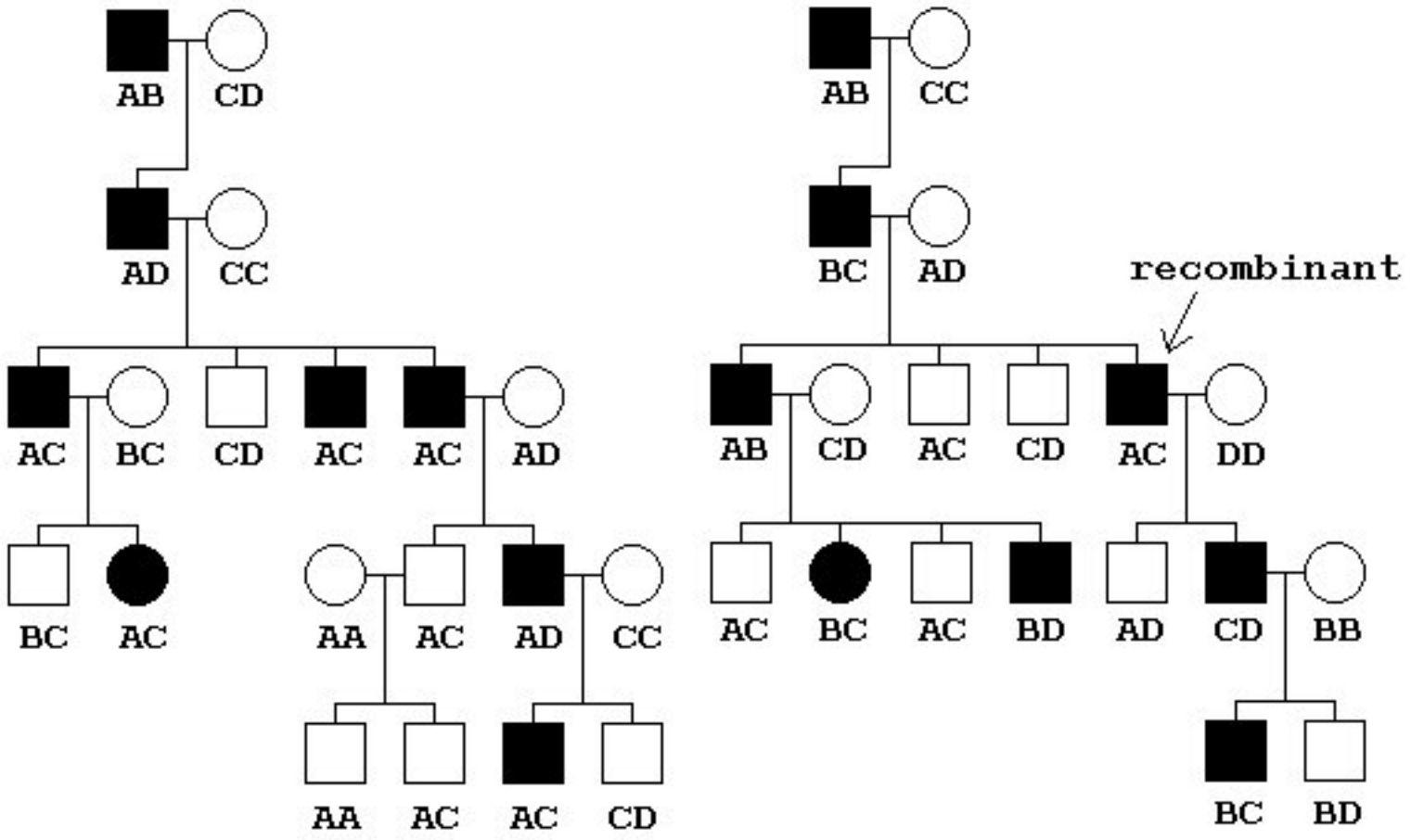
Lecture 17A — Pedigree-based Gene Mapping

- How to read a pedigree
- Transmission probabilities
- Lod scores

Thanks to Mary Kuhner for most slides

Reading a pedigree



- Squares are males, circles are females
- Shaded symbols are affected, Half-shaded are carriers

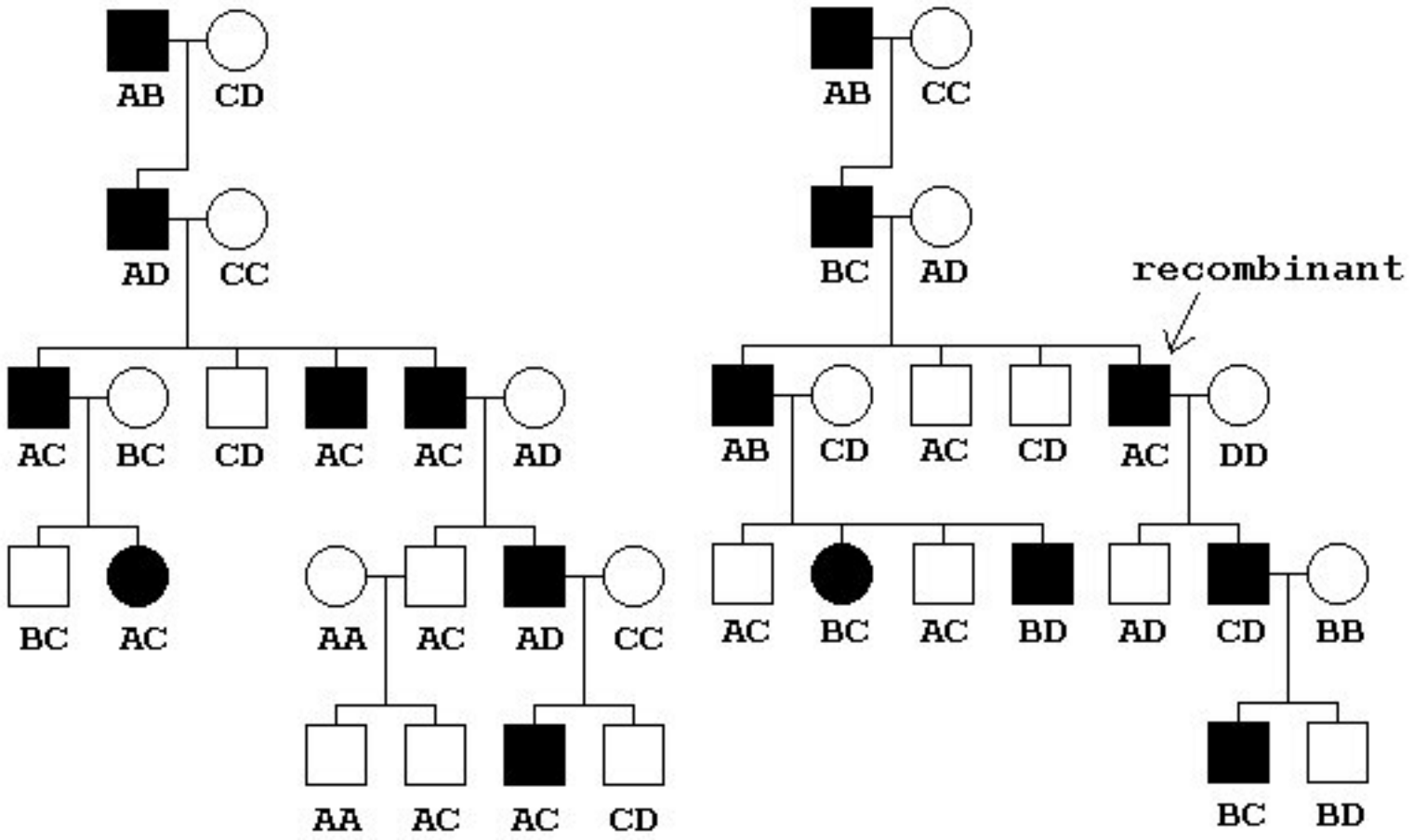
Modes of inheritance

- Dominant–one gene copy leads to trait
- Recessive–two gene copies lead to trait
- Intermediate/Codominant–heterozygote is distinct

Modes of inheritance

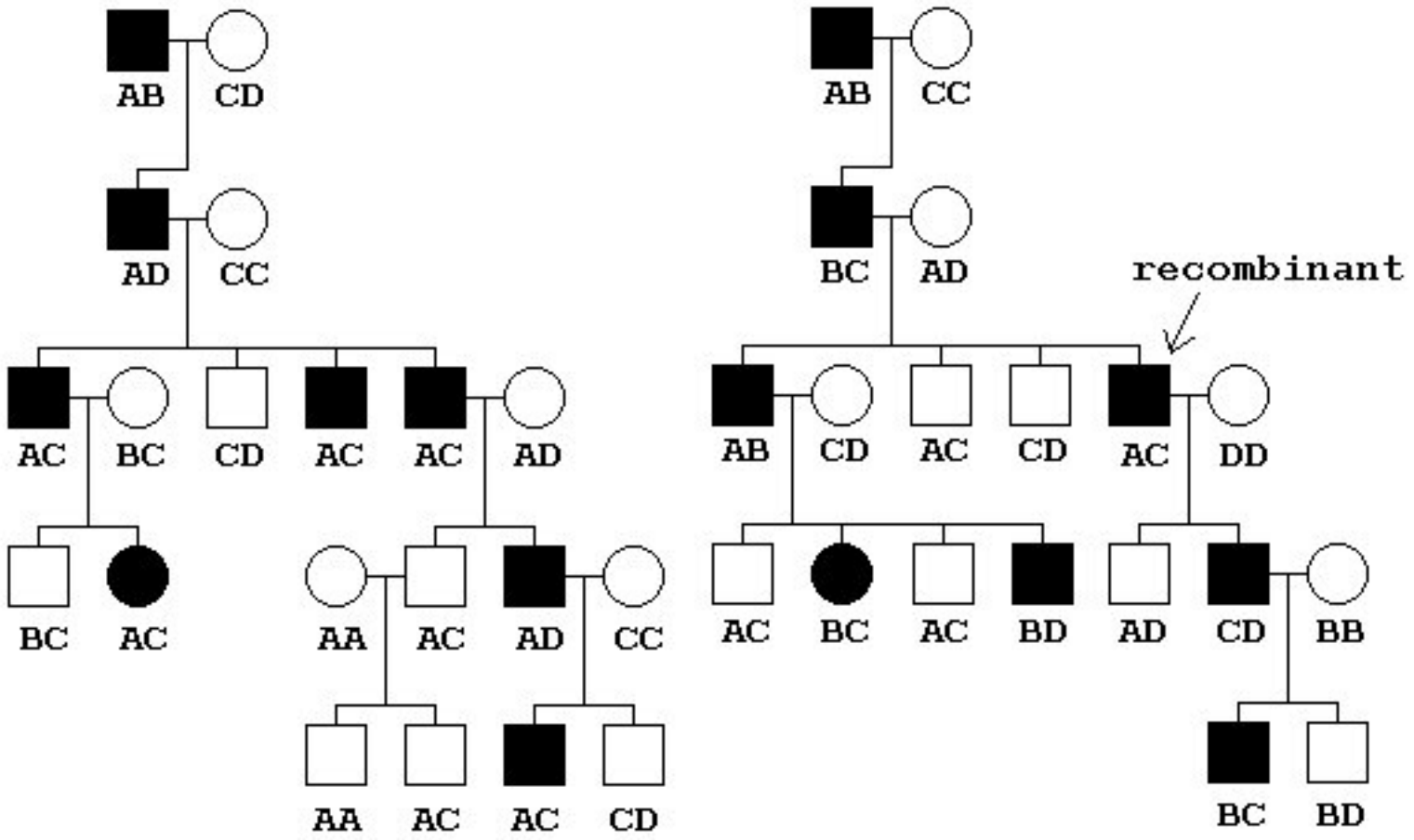
- Recessive trait:
  - Skips generations
  - Shows up in both sides of the family tree
  - Two affected individuals have only affected offspring
- Dominant trait:
  - Does not skip generations
  - Often in only one side of family tree
  - Two affected individuals may have unaffected offspring

Analyzing a pedigree with marker data



- Try to identify the chromosome carrying the disease trait
- Trace it through the pedigree

Recombinants



- In the left pedigree, disease assorts with A throughout
- In the right pedigree, there has been a recombination

Recombination frequency

- Written as θ
- Percentage of transmissions in which a (newly) recombinant chromosome was transmitted
- $\theta = 0$ is perfect linkage
- $\theta = 0.5$ is no linkage

Lod score

- Lod==”Log of Odds”
- Lod score measures probability of pedigree under linkage versus no linkage hypotheses
- Normally computed using log_{10} (base 10 log)

$$Lod = \log_{10} \frac{P(\text{data} \mid \theta)}{P(\text{data} \mid \theta = 0.5)}$$

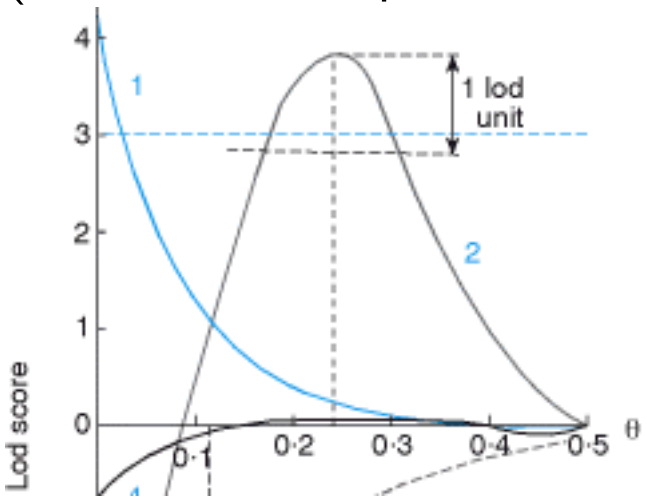
$$Lod = \log_{10} \frac{(1-\theta)^{NR} \times \theta^{R}}{0.5^{(NR+R)}}$$

Lod score

- Lod scores can be added across families
- Value greater than 3.0 considered to show linkage
- (This is a 1 in 1000 chance–conservative but allows for multiple tests)
- Value less than -2.0 shows non-linkage (100:1 against)

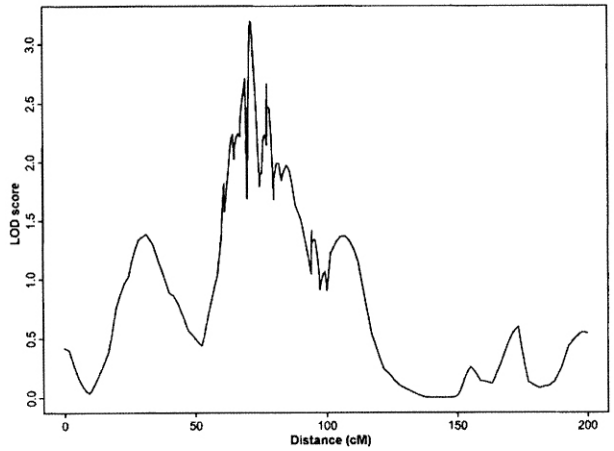
Complex Pedigrees

- When individuals are ambiguous, can sum over possibilities
- MCMC (Markov chain Monte Carlo) can be used here
- Compute Lod score for different values of θ
- Plot as a curve: maximum is most likely recombination distance. (Another example of *maximum likelihood parameter estimation.)*



from Strachan & Read Human Molecular Genetics 2 http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=hmg

Lod score curve



- More than one marker makes a better map
- Multiple densely placed markers give the most accurate map

Family linkage studies

- Advantages:
  - Reduced chance of disease heterogeneity within a family
  - Clear observation of recombinations
- Disadvantages:
  - Suitable large families rare
  - Seldom locates gene more closely than 5 cM
  - Very difficult for late-onset diseases