Office hours
3-4pm Wednesdays
304A Stanley Hall

QUIZ: Nov. 20, 21, 24
Covers material through lecture Nov. 17

More realistic situation: in dad, phase of alleles unknown

\[
\begin{align*}
\text{Dad phase unknown} \\
b & = 0.5 \\
\text{odds ratio } & = \frac{1}{2} r^n + \frac{1}{2} (1-r^n) \\
\frac{0.5 \text{(total # meioses)}}{}
\end{align*}
\]

What single r value best explains the data?
For this, you need to search r's.

Modern genetic scans

What does the “max” in “max LOD score” refer to?
A. The strongest-linking marker
B. The most probable recombination fraction
C. The most severe phenotype

Modern genetic scans

Remember?

Max LOD score is the one from the best r value
Modern genetic scans

What is the simplest explanation for so many tall black lines around Chr 13?
A. Multiple markers in the region, which makes LOD higher
B. Multiple markers are all linked to a single disease mutation
C. Multiple mutations on Chr 13 cause the disease
D. Higher LOD is counted by the number of linking markers

Modern genetic scans

(Smooth curve = inferred genotype at positions between markers)

Modern genetic scans

Modern genetic scans
Modern genetic scans

Why would an experiment fail to observe linkage?

Marker density matters

But…

No Major Schizophrenia Locus
Detected on Chromosome 1q in a Large Multicenter Sample

Douglas J. L. Martinian, Peter A. Holmes, Grahame Leckman, Brian Miller, Jean R. Kurz, Yuka M. Kubo, 
Mark C. Schiavo, Keith M. Williams, Michael J. Owen, 
Peter A. Holmes', Kurt R. Sanders, Gerald Nestadt, 
John E. Martinian, and Michael J. Owen. 

But if the only marker you test is >50 cM away, will get no linkage.

Try to minimize genotyping cost.
Number of families matters

If low number of patients, no statistical significance.

Improper statistics

Can make noise look like a fabulously significant linkage peak.

Locus heterogeneity

(a) Complementation: mutations in two different genes

P Aa bb x aabb

F1 Aa Bb
Genetic mechanism of complementation

(b) Noncomplementation: mutations in the same gene

P Aa bb x Aa bb

F1
Genetic mechanism of noncomplementation

Fig. 3.16
Age of onset in breast cancer

Table 1. Recombination frequency

<table>
<thead>
<tr>
<th>Family</th>
<th>N</th>
<th>Recombination frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>25</td>
<td>0.005</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>0.005</td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>0.005</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>0.005</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>0.005</td>
</tr>
<tr>
<td>6</td>
<td>25</td>
<td>0.005</td>
</tr>
<tr>
<td>7</td>
<td>25</td>
<td>0.005</td>
</tr>
<tr>
<td>8</td>
<td>25</td>
<td>0.005</td>
</tr>
<tr>
<td>9</td>
<td>25</td>
<td>0.005</td>
</tr>
<tr>
<td>10</td>
<td>25</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Linkage of Early-Onset Familial Breast Cancer

Age of onset in breast cancer

Table 1. Recombination frequency

<table>
<thead>
<tr>
<th>Family</th>
<th>N</th>
<th>Recombination frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>25</td>
<td>0.005</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>0.005</td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>0.005</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>0.005</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>0.005</td>
</tr>
<tr>
<td>6</td>
<td>25</td>
<td>0.005</td>
</tr>
<tr>
<td>7</td>
<td>25</td>
<td>0.005</td>
</tr>
<tr>
<td>8</td>
<td>25</td>
<td>0.005</td>
</tr>
<tr>
<td>9</td>
<td>25</td>
<td>0.005</td>
</tr>
<tr>
<td>10</td>
<td>25</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Familial breast cancer is heterogeneous.

Only early-onset families show linkage.

Familial breast cancer is heterogeneous.
Locus heterogeneity

A landmark: BRCA1

More breast cancer FYI (see lecture 9/15)

BRCA1 and 2 FYI

<table>
<thead>
<tr>
<th>Gene</th>
<th>Normal Function of Gene (Gene’s), or Disease Syndromes Resulting from Mutation</th>
<th>Function of Normal Protein Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATM</td>
<td>Cerk1: Rb, Mdm2, p53, and BRCA1</td>
<td>Transcription factor</td>
</tr>
<tr>
<td>AR</td>
<td>cerk1: Rb, 11-14, and IGF-1</td>
<td>Transcription factor</td>
</tr>
<tr>
<td>TP53</td>
<td>Cerk1: Rb, 11-14, and IGF-1</td>
<td>Transcription factor</td>
</tr>
<tr>
<td>BRCA1</td>
<td>Repair of DNA breaks</td>
<td>Repair of DNA breaks</td>
</tr>
</tbody>
</table>

Many tumor suppressor genes have been associated with specific functions in the cell cycle necessary for accuracy of cell division.
BRCA1 and 2 FYI

• Only ~10% of breast cancers are hereditary

• Different from sporadic: age, histology, sex

• BRCA1 and BRCA2 found from linkage analysis of families with multiple affecteds (1990, 1994)

• BRCA1 or 2 mutation = ~80% likely to get disease
BRCA1 and 2 FYI

- Only ~10% of breast cancers are hereditary
- Different from sporadic: age, histology, sex
- BRCA1 and BRCA2 found from linkage analysis of families with multiple affecteds (1990, 1994)
- BRCA1 or 2 mutation = ~80% likely to get disease

Even familial form is more than just BRCA1 and 2

Multiple causes = hard to find any one cause

In the limit of studying a single family with severe disease, more likely to find one strong locus.

But hard to find such families, and segregating allele may not be relevant for chronic/common disease.

Significance cutoff
Rule of thumb: don’t believe linkage unless odds > 1000. Why?

LOD scores

$r =$ genetic distance between marker and disease locus

$$\text{Odds} = \frac{P(\text{pedigree} \mid r)}{P(\text{pedigree} \mid r = 0.5)}$$

$$= \frac{(1-r)^n \cdot r^k}{0.5^{(\text{total # meioses})}}$$

<table>
<thead>
<tr>
<th>$r$</th>
<th>Odds</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>12.244</td>
</tr>
<tr>
<td>0.2</td>
<td>10.737</td>
</tr>
<tr>
<td>0.3</td>
<td>6.325</td>
</tr>
<tr>
<td>0.4</td>
<td>2.867</td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
</tr>
</tbody>
</table>

Coins

$r =$ intrinsic probability of coming up heads (bias)

$$\text{Odds} = \frac{P(\text{your flips} \mid r)}{P(\text{your flips} \mid r = 0.5)}$$

$$= \frac{(1-r)^n \cdot r^k}{0.5^{(\text{total # flips})}}$$
Odds = \frac{P(\text{your flips | } r)}{P(\text{your flips | } r = 0.5)} = \frac{(1-r)^k \cdot r^n}{0.5^{\text{total # flips}}}

Unknown we seek is "fairness" of a coin (analogous to recombination fraction)

Odds ratio of model that coin is biased, relative to null

If you do 10,000 flips and 7,000 are heads, what do you expect for r?
A. 0
B. 0.7
C. 0.5
D. 1
Coins

Take out a coin and flip 4 times.

How many heads?

Want to find intrinsic prob of heads (analogous to recombination fraction).

With only 4 data points, can’t use $\chi^2$ (analogous to a small family).

$r = \text{intrinsic probability of coming up heads (bias)}$

\[
\text{Odds} = \frac{(1-r)^n \cdot r^k}{0.5^{\text{total # flips}}}
\]

Odds ratio of model that coin is biased, relative to null

<table>
<thead>
<tr>
<th>Odds</th>
<th>2 heads</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1296</td>
<td>0.9</td>
</tr>
<tr>
<td>0.4096</td>
<td>0.8</td>
</tr>
<tr>
<td>0.7056</td>
<td>0.7</td>
</tr>
<tr>
<td>0.9216</td>
<td>0.6</td>
</tr>
<tr>
<td>0.92</td>
<td>1</td>
</tr>
<tr>
<td>0.7056</td>
<td>0.4</td>
</tr>
<tr>
<td>0.4096</td>
<td>0.3</td>
</tr>
<tr>
<td>0.1296</td>
<td>0.2</td>
</tr>
<tr>
<td>0.0</td>
<td>0.1</td>
</tr>
<tr>
<td>0.0</td>
<td>0</td>
</tr>
</tbody>
</table>

Odds ratio of model that coin is biased, relative to null

<table>
<thead>
<tr>
<th>Odds</th>
<th>2 heads</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1296</td>
<td>0.9</td>
</tr>
<tr>
<td>0.4096</td>
<td>0.8</td>
</tr>
<tr>
<td>0.7056</td>
<td>0.7</td>
</tr>
<tr>
<td>0.9216</td>
<td>0.6</td>
</tr>
<tr>
<td>0.92</td>
<td>1</td>
</tr>
<tr>
<td>0.7056</td>
<td>0.4</td>
</tr>
<tr>
<td>0.4096</td>
<td>0.3</td>
</tr>
<tr>
<td>0.1296</td>
<td>0.2</td>
</tr>
<tr>
<td>0.0</td>
<td>0.1</td>
</tr>
<tr>
<td>0.0</td>
<td>0</td>
</tr>
</tbody>
</table>

$r = \text{intrinsic probability of coming up heads (bias)}$
**Coins**

$r = \text{intrinsic probability of coming up heads (bias)}$

<table>
<thead>
<tr>
<th>Heads</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Odds</td>
<td>0</td>
<td>0.1</td>
<td>0.4</td>
<td>1</td>
<td>16</td>
</tr>
<tr>
<td>0.1</td>
<td>0</td>
<td>0.0144</td>
<td>0.1024</td>
<td>0.6144</td>
<td>0.4096</td>
</tr>
<tr>
<td>0.2</td>
<td>0</td>
<td>0.0288</td>
<td>0.2048</td>
<td>1.2288</td>
<td>0.8192</td>
</tr>
<tr>
<td>0.3</td>
<td>0</td>
<td>0.0432</td>
<td>0.2832</td>
<td>1.8528</td>
<td>1.2288</td>
</tr>
<tr>
<td>0.4</td>
<td>0</td>
<td>0.0576</td>
<td>0.3648</td>
<td>2.4832</td>
<td>2.0568</td>
</tr>
<tr>
<td>0.5</td>
<td>0</td>
<td>0.0720</td>
<td>0.4464</td>
<td>3.1152</td>
<td>2.8896</td>
</tr>
<tr>
<td>0.6</td>
<td>0</td>
<td>0.0864</td>
<td>0.5280</td>
<td>3.7472</td>
<td>3.7216</td>
</tr>
<tr>
<td>0.7</td>
<td>0</td>
<td>0.0998</td>
<td>0.6096</td>
<td>4.3800</td>
<td>4.5552</td>
</tr>
<tr>
<td>0.8</td>
<td>0</td>
<td>0.1132</td>
<td>0.6912</td>
<td>5.0132</td>
<td>5.3888</td>
</tr>
<tr>
<td>0.9</td>
<td>0</td>
<td>0.1266</td>
<td>0.7728</td>
<td>5.6464</td>
<td>6.2224</td>
</tr>
</tbody>
</table>

**Observed Rate**

- The observed rate is the best numerical solution.

**Notes**

- $r = \text{intrinsic probability of coming up heads (bias)}$
- $\text{Odds} = \frac{(1-r)^n}{r^n} \cdot \frac{1}{0.5^{\text{total # flips}}}$

---

**Coins**

$r = \text{intrinsic probability of coming up heads (bias)}$

<table>
<thead>
<tr>
<th>Heads</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Odds</td>
<td>0</td>
<td>0.1</td>
<td>0.4</td>
<td>1</td>
<td>16</td>
</tr>
<tr>
<td>0.1</td>
<td>0</td>
<td>0.0144</td>
<td>0.1024</td>
<td>0.6144</td>
<td>0.4096</td>
</tr>
<tr>
<td>0.2</td>
<td>0</td>
<td>0.0288</td>
<td>0.2048</td>
<td>1.2288</td>
<td>0.8192</td>
</tr>
<tr>
<td>0.3</td>
<td>0</td>
<td>0.0432</td>
<td>0.2832</td>
<td>1.8528</td>
<td>1.2288</td>
</tr>
<tr>
<td>0.4</td>
<td>0</td>
<td>0.0576</td>
<td>0.3648</td>
<td>2.4832</td>
<td>2.0568</td>
</tr>
<tr>
<td>0.5</td>
<td>0</td>
<td>0.0720</td>
<td>0.4464</td>
<td>3.1152</td>
<td>2.8896</td>
</tr>
<tr>
<td>0.6</td>
<td>0</td>
<td>0.0864</td>
<td>0.5280</td>
<td>3.7472</td>
<td>3.7216</td>
</tr>
<tr>
<td>0.7</td>
<td>0</td>
<td>0.0998</td>
<td>0.6096</td>
<td>4.3800</td>
<td>4.5552</td>
</tr>
<tr>
<td>0.8</td>
<td>0</td>
<td>0.1132</td>
<td>0.6912</td>
<td>5.0132</td>
<td>5.3888</td>
</tr>
<tr>
<td>0.9</td>
<td>0</td>
<td>0.1266</td>
<td>0.7728</td>
<td>5.6464</td>
<td>6.2224</td>
</tr>
</tbody>
</table>

---

**Coins**

$r = \text{intrinsic probability of coming up heads (bias)}$

<table>
<thead>
<tr>
<th>Heads</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Odds</td>
<td>0</td>
<td>0.1</td>
<td>0.4</td>
<td>1</td>
<td>16</td>
</tr>
<tr>
<td>0.1</td>
<td>0</td>
<td>0.0144</td>
<td>0.1024</td>
<td>0.6144</td>
<td>0.4096</td>
</tr>
<tr>
<td>0.2</td>
<td>0</td>
<td>0.0288</td>
<td>0.2048</td>
<td>1.2288</td>
<td>0.8192</td>
</tr>
<tr>
<td>0.3</td>
<td>0</td>
<td>0.0432</td>
<td>0.2832</td>
<td>1.8528</td>
<td>1.2288</td>
</tr>
<tr>
<td>0.4</td>
<td>0</td>
<td>0.0576</td>
<td>0.3648</td>
<td>2.4832</td>
<td>2.0568</td>
</tr>
<tr>
<td>0.5</td>
<td>0</td>
<td>0.0720</td>
<td>0.4464</td>
<td>3.1152</td>
<td>2.8896</td>
</tr>
<tr>
<td>0.6</td>
<td>0</td>
<td>0.0864</td>
<td>0.5280</td>
<td>3.7472</td>
<td>3.7216</td>
</tr>
<tr>
<td>0.7</td>
<td>0</td>
<td>0.0998</td>
<td>0.6096</td>
<td>4.3800</td>
<td>4.5552</td>
</tr>
<tr>
<td>0.8</td>
<td>0</td>
<td>0.1132</td>
<td>0.6912</td>
<td>5.0132</td>
<td>5.3888</td>
</tr>
<tr>
<td>0.9</td>
<td>0</td>
<td>0.1266</td>
<td>0.7728</td>
<td>5.6464</td>
<td>6.2224</td>
</tr>
</tbody>
</table>
By chance, can get good LOD score for just about anything.

The more students you have flipping coins, the more likely you are to see this “unlikely” combination.

The multiple testing problem
Multiple testing in genetics

Testing lots of markers for linkage to a trait is analogous to having lots of students, each flipping a coin.

Can get spurious high LOD to an unlinked marker, just by chance.

Don’t let this happen to you!

Significance cutoff

Using LOD=3 as cutoff more or less eliminates this problem. We’ll see why on Friday.