Office hours
Wednesday 3-4pm
304A Stanley Hall

Association mapping (qualitative)

Association scan, qualitative

<table>
<thead>
<tr>
<th></th>
<th>Case</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>χ² test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C's</td>
<td>141</td>
<td>797</td>
</tr>
<tr>
<td>G's</td>
<td>47</td>
<td>433</td>
</tr>
</tbody>
</table>

Association scan, quantitative

Association vs. linkage

Unrelated individuals (usually)

Related individuals
Association vs. linkage

Unrelated individuals (usually)

Extreme of linkage study is one large family; less likely that phenotype has multiple genetic causes (locus heterogeneity).

Related individuals

Strong, easy to detect

Association vs. linkage

Unrelated individuals

Related individuals

Strong, easy to detect, but rare in population

Unrelated individuals

Related individuals

Strong, easy to detect, but rare in population; may not be reflective of common disease.

Also, hard to collect family data.

Association vs. linkage

Unrelated individuals

Related individuals

Strong, easy to detect, but rare in population; may not be reflective of common disease.

Also, hard to collect family data.

(e.g. BRCA1)
Association vs. linkage

Unrelated individuals
Common but weak effects

Strong, easy to detect, but rare in population; may not be reflective of common disease. Also, hard to collect family data.

Related individuals
Common but weak effects; need 1000’s of samples to detect. If no common cause, can fail.

Another key feature of association mapping: resolution

many recombinations have happened since common ancestor; shared region is small, no co-inheritance between distant markers

Small number of generations; individuals share big chunks of genome; can get co-inheritance between distant markers
Association vs. linkage

Small number of generations; individuals share big chunks of genome; can get co-inheritance between distant markers

So you need very high density of markers to get signal in an association study, but you get very high spatial resolution.

In the “old days” of sparse markers, linkage analysis was the best strategy.

Causative variant very close

But there is a pitfall of association tests: “population structure”

Diabetes in Native Americans

Summary

The prevalence of diabetes mellitus among the Pima Indians, who live in a hot desert environment in Arizona, U.S.A., has been determined by means of ophthalmologic screening tests. Using prospective observational panels of diabetes-free and diabetes-prone members of the Pima Indian population, a prevalence of 1.8% for diabetes was found in the white population normal and hyperglycemic groups may be logi-

cally separated on the basis of the binomial of the frequency distributions of two-hour post-load glucose levels.

(1971)
Diabetes in Native Americans

Family studies indicate it is at least partly genetic, not environmental.

Association mapping causal loci

Typed IgG heavy chains with protein assay. Phenotypes can serve as markers too...

(Multiple proteins from chr 14 region: haplotype)

Table 1: Association Between Diabetes and the Haplotype

<table>
<thead>
<tr>
<th>Gm</th>
<th>diabetes</th>
<th>control</th>
</tr>
</thead>
<tbody>
<tr>
<td>no Gm</td>
<td>1343</td>
<td>3284</td>
</tr>
<tr>
<td>Gm</td>
<td>23</td>
<td>270</td>
</tr>
</tbody>
</table>

Note: Diabetes was not correctly associated with the haplotype (Gm14:12)² = 1.96, P = 0.05. Phenotype + 0.27 95% confidence interval: 0.08-0.46.
Association mapping causal loci

“Gm is protective against diabetes?”

<table>
<thead>
<tr>
<th></th>
<th>diabetes</th>
<th>control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gm</td>
<td>23</td>
<td>270</td>
</tr>
<tr>
<td>no Gm</td>
<td>1343</td>
<td>3284</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gm haplotype and Type 3 Diabetes Mellitus: An Association in American Indians with Genetic Admixture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-identified heritage</td>
</tr>
<tr>
<td>Most “full heritage” members don’t have the haplotype</td>
</tr>
<tr>
<td>The few without N.A. heritage are much more likely to have thehaplotype</td>
</tr>
<tr>
<td>Gm haplotype is very rare in self-identified 100% Pima members.</td>
</tr>
</tbody>
</table>
Gm haplotype is very rare in self-identified 100% Pima members. Gm is a marker for Caucasian ancestry.
Caucasian ancestry is associated with Gm haplotype.
Caucasian ancestry is associated with lower diabetes risk.

But Gm is not associated with lower diabetes risk!

Controls are enriched for Caucasians

N.A. and Caucasians are different at many loci

At any one of these loci, Caucasian-like allele will be enriched in control samples.
Association and admixture

Don’t believe any one locus is causative!

Microarrays and genotyping

DNA microarrays

Post-genome era: the sequence and location of the oligos are known

DNA microarrays

Genotyping by array
Genotyping by array

Fig. 11.10

oligo (human genome fragment) index

middle nucleotide on chip oligo
Genotyping by array

Sample DNA is labeled, allowed to hybridize

Fig. 11.10

Genotyping by array

middle nucleotide on chip oligo

2450
2460

2450
2460

readout of sample

Genotyping by array

Fabrication technology allows millions of oligos (each present in millions of copies) on a single slide

Genotyping by single-base extension

High density of markers necessary for association