$$(p+q)^2 = \underset{AA}{p^2} + \underset{Aa}{2pq} + \underset{aa}{q^2} = 1.0$$

Deviation from Hardy Weinberg Equilibrium: Wahlund Effect

Real data: Navajo populations

- MN blood type in Navajo
  - MM: 305
  - MN: 52
  - NN: 4

Real data: Navajo populations

- MN blood type in Navajo
  - MM: 305 0.845 p(M) = 0.917
  - MN: 52 0.144 q(N) = 0.083
  - NN: 4 0.011

  TOTAL: 361

<u>HW Predicted</u>:

MM - 0.841

MN - 0.152 Very close to HW

NN - 0.007

Real data: Aborigine populations

- MN blood type in Aborigine
  - MM: 22
  - MN: 216
  - NN: 492

Real data: Aborigine populations

- MN blood type in Aborigine
  - MM: 22 0.030
  - MN: 216 0.296
  - NN: 492 0.674

  TOTAL: 730

p(M) = 0.178

q(N) = 0.822

<u>HW Predicted</u>:

MM - 0.031

MN - 0.293 Very close to HW

NN - 0.676

Real data: Mixed population

- MN blood type
  - MM: 327
  - MN: 268
  - NN: 496

Real data: Mixed population

- MN blood type
  - MM: 327 0.300
  - MN: 268 0.246
  - NN: 496 0.454

  TOTAL: 1091

p(M) = 0.423

q(N) = 0.577

HW Predicted:

MM - 0.179

MN - 0.488 **NOT** close to HW

NN - 0.333

Random mating assumption

- Any two individuals as likely to breed as any other two individuals
- Is a Navajo lady **as likely to breed with** an Aborigine as she is with another Navajo???
  - NO!

Hardy Weinberg assumption VIOLATED!

- Hardy Weinberg assumption not rejected:
  - Within Navajo
  - Within Aborigine
- Deviation from HW observed in combination
  - Result of “non-random mating”
- Observations:
  - Too few heterozygotes
    - Expected: 0.488
    - Observed: 0.246

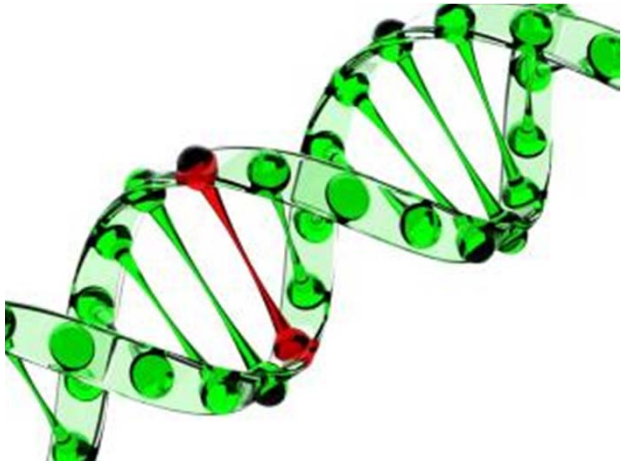
Hardy Weinberg assumption VIOLATED!

- Hardy Weinberg assumption not rejected:
  - Within Navajo
  - Within Aborigine
- Deviation from HW observed in combination
  - Result of "non-random mating"
- Observations:
  - Too few heterozygotes
    - Expected: 0.488
    - Observed: 0.246

Wahlund effect- sampling across populations gives underrepresentation of heterozygotes relative to HW

Why does it matter???

- First step in genome-wide association studies of genetic diseases is usually to **test for HW**!
  - Why?



GWAS assumes Hardy-Weinberg (or nearly so)

- Assumes "LD" detected between marker alleles and disease alleles is caused by **close proximity/ lack of recombination**
  - 20% of those with "AA" genotype have disease
  - 5% of those with "aa" genotype have disease
    - Association between "A" marker genotype and disease

- Being in "different populations" (ie, not random mating) *also* creates LD...

Example:

- Imagine Population 1 all “AA”
- Imagine Population 2 all “aa”
- If disease is more abundant in Population 1 than Population 2
  - Would say “AA” more likely to have disease than “aa”
  - “Fake” LD between disease and gene A
  - Disease gene may be on different chromosome than A, or disease maybe not even influenced by genetics at all!

Punchline

- If there are allele frequency differences between "populations" at a SNP
- If there are disease incidence differences between those "populations" too
- THEN, it'll (erroneously) look like a gene near the SNP causes/ contributes to the disease

- Testing for HW helps you avoid this error
  - Identifies if looking at one or more "populations"

Very important to test for HW, BUT OFTEN NOT DONE!

- *Am J Epidemiology* 2006- Exclusion of studies in which HW was violated changed conclusions and changed statistical significance of gene-disease associations

- *Eur J Human Genetics* 2005- "testing and reporting for HWE is often neglected and deviations are rarely admitted in the published reports."

Real example where HW test WAS done but interpreted incorrectly…

- 2000 study of BRCA2 variants

  Newborn males from hospital in United Kingdom

  –AA: 644

  –Aa: 435

  –aa: 116

  Calculate HW expectations.

  Speculate on fit or lack of fit.

Real example where HW test WAS done but interpreted incorrectly…

- 2000 study of BRCA2 variants

  Newborn males from hospital in United Kingdom

  –AA: 644 0.54 A: 0.72, a: 0.28

  –Aa: 435 **0.36** **HW Expected**

  –aa: 116 0.10 AA: 0.52

  TOTAL: 1195 Aa: **0.40**

  Deviation from HW aa: 0.08

  statistically significant… too few Aa observed.

  Authors inferred Aa are less healthy than AA or aa.

Ironic tidbit...

Quote from Hardy (1940) in
A Mathematician's Apology :

I have never done anything 'useful'. No discovery of mine has made, or is likely to make, directly or indirectly, for good or ill, the least difference to the amenity of the world.

He was very wrong on this.

Written when he was 62 and lamenting
the waning of his mathematic ability…

Image Credits, Unit 10-3

- Aborigines, © 2007 Mombas, CC by-SA 3.0, en.wikipedia.org
- Cupid, © Anna Velichkovsky, all rights reserved, www.photoxpress.com
- Green DNA w/ red mutation, © appler, all rights reserved, www.photoxpress.com
- Microarray, © 2006 Schutz, CC by-SA 3.0, en.wikipedia.org.
- Crowd, © 2006 SchuminWeb, CC by-SA 2.5, en.wikipedia.org
- United Kingdom map, © 2006 Captain Blood, CC by-SA 3.0, en.wikipedia.org
- Pink ribbon, © 2007 MesserWoland, CC by-SA 3.0, en.wikipedia.org.