calculation

Calculation about mendel gene inherit

1. Calculating genotype frequency

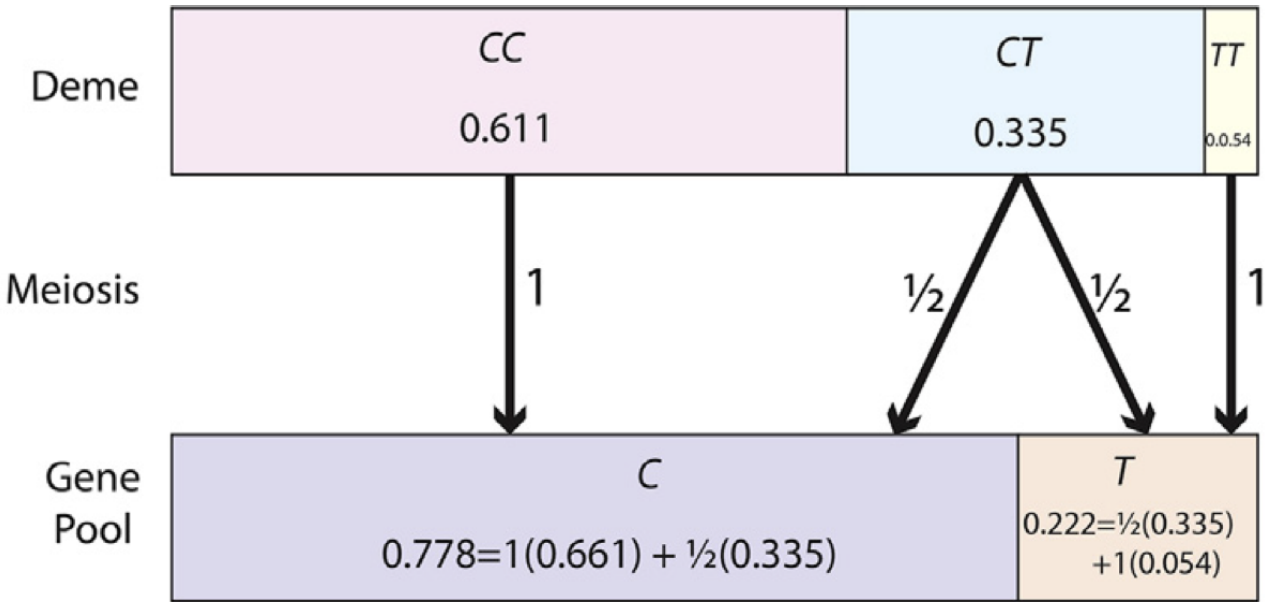
Number with genotype CC: 102
Number with genotype CT: 56
Number with genotype TT: 9
These genotype numbers are converted into genotype frequencies simply by dividing the observed number of each genotype by the total sample size; that is:

2. Gene pool

defined by SNP
Number with genotype CC: 102
Number with genotype CT: 56
Number with genotype TT: 9

3. Gamete

- Another definition of gene pool



$$g_i = \sum_{j=1}^{n} G_j t_{j \to i}$$

-
- gi 是潜在配子群体中配子类型 i 的频率，
- Gj 是群体中基因型 j 的频率，
- n 是基因型的数量，
- tj→i 是基因型 j 通过减数分裂产生配子类型 i 的过渡概率。

4. meiotic transition probabilities

假设我们有一个二倍体个体，其基因型为AaAa。在理想情况下，这个个体通过减数分裂产生四种配子类型的概率是相等的。然而，如果两个位点之间存在连锁不平衡（也就是说，两个位点之间的重组频率小于50%），那么这个个体产生AA和aa配子的概率就会高于产生Aa和aA配子的概率。

我们可以用过渡概率来描述这种情况。例如，我们可以设定以下的过渡概率：

- tAaAa→AA=0.3
- tAaAa→aa=0.3
- tAaAa→Aa=0.2
- tAaAa→aA=0.2

5. Genotype frequencies

- can predict genotype frequency from allele frequency

6. Application of HWE (注意运用极限的思想)

AR disease(allele recessive)

- When the disease-causing allele frequency q approaches 0, and the normal allele frequency p approaches 1
    - carrier frequency 约等于 2q

AD disease(allele recessive)

- Q2: The frequency of individuals affected with osteogenesis imperfect (an AD disease) is 1 in 10000 in Denmark. Calculate the disease-causing allele frequency.

Testing whether the condition fits HWE

- Q3: Genotype frequencies of MN blood group are based on a sample of 1066 Chukchi individuals, a native people of eastern Siberia (Roychoudhury and Nei 1988). Test whether this population is under the null hypotheses of HWE.

Genotype	Observed
MM	165
MN	562
NN	339

7. Heterozygous advantage

- 已知relative fitness, 需要求达到平衡的时候的配子频率

Genotype	Relative fitness
AR/AR	0.65
ST/ST	0.50
AR/ST	1.00

8. Mutation balance

- 有时候需要计算突变平衡的情况下的mutation rate
  Q5: Duchenne muscular dystrophy: an X-linked recessive trait. Lethal in hemizygous males and recessive homozygous females. At equilibrium for a population, what is the rate of new mutations that alter the disease-causing allele?

Mutation-selection Balance

- At equilibrium, the frequency of a mutant allele in a population is a balance of the intensity of natural selection against the allele
- the rate of elimination of deleterious alleles by natural selection (s)
- the rate at which new mutant alleles are generated (μ).
- 如果是在这样的条件下达到了突变平衡，那么在隐性条件下（之所以是2u是由于平衡条件不仅仅是要求alleles平衡，还必须有phenotype平衡。所以2u代表的是突变成为隐形的个体）

$$2\mu = 2sq^2$$

- 在显性突变中

$$2\mu = 2sp$$

Recombination frequency

- recombination frequency (or fraction) is θ

- From 0 to 0.5
  - 0: no recombination at all, the distance is too close
  - 0.5: distance high

Linkage disequilibrium

> is the non-random association of alleles at closely linked loci in a given population. 两个基因不完全独立地遗传，就会出现某种程度的连锁。

- 如果两个基因是完全独立遗传，那么

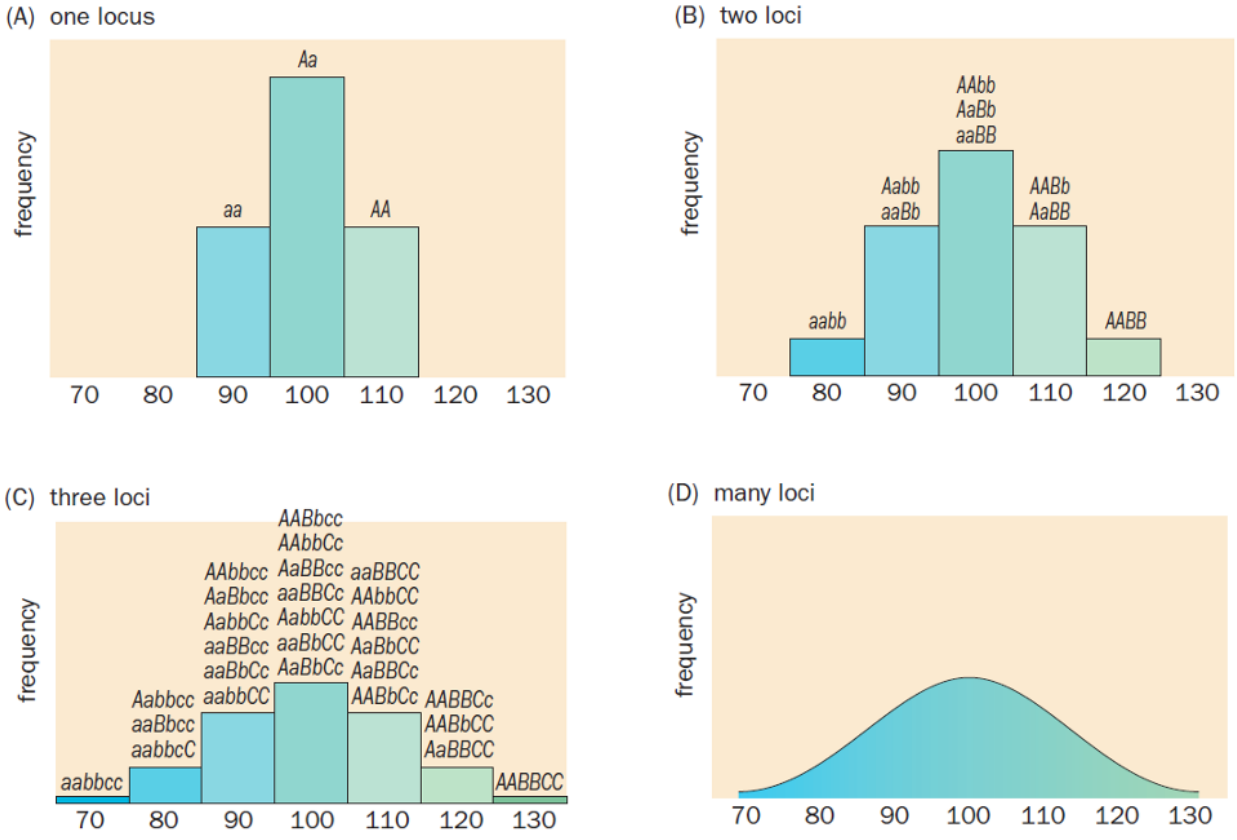
$$P(A) \times P(B)! = P(AB)$$

- 但是如果两者不等于，就可以推测两个基因应该是有某种程度的连锁的

calculation about multifactorial inheritance

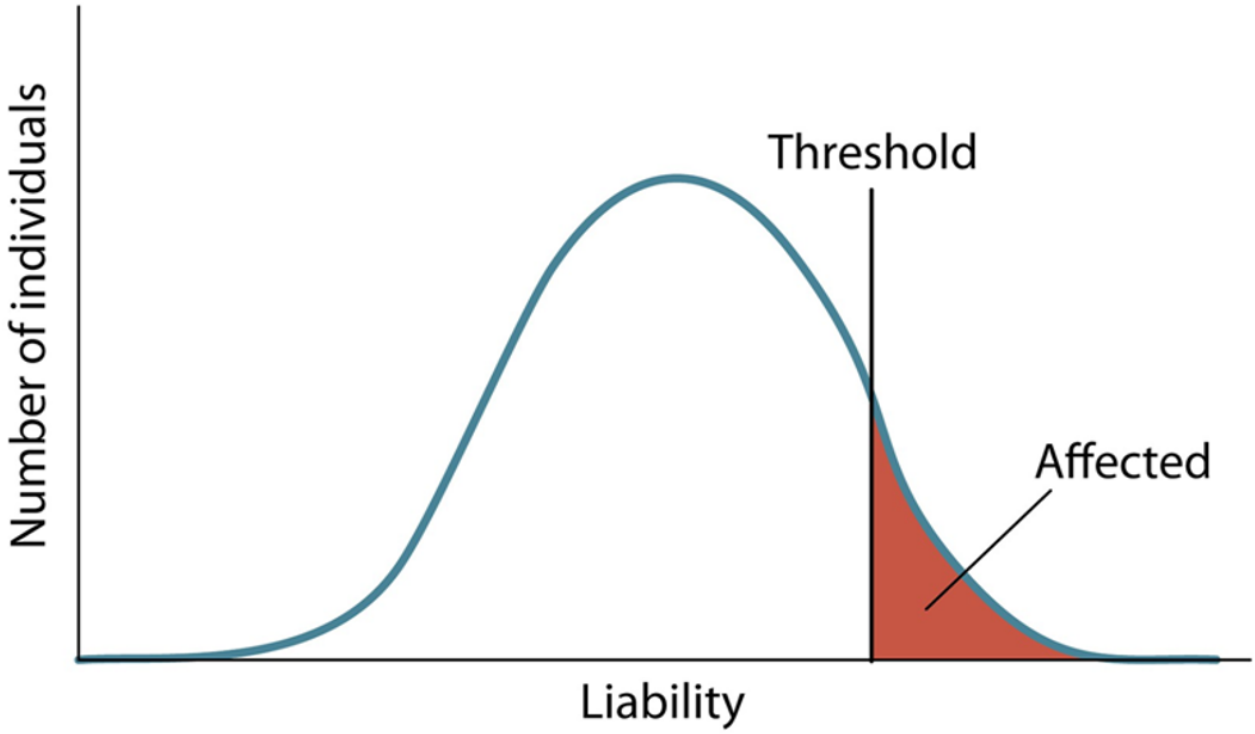
Simple additive polygenic model

- 每个基因型都不是独立作用的，而是都有加性



Liability threshold model

- Follows the principle of simple additive polygenic model
- But qualitative, set a threshold
  - Lower than the threshold, nothing happens
  - Higher than the threshold, disease



Risk ratio

$$\lambda_r = \frac{\text{the frequency of the trait in relatives}}{\text{the frequency of the general population}}$$

Variance in a phenotype

$$V_P = \overbrace{V_A + V_D}^{\text{Genetic variance}} + \overbrace{V_E + V_I}^{\text{Environmental variance}} + \text{Cov}_{GE} + \overbrace{V_M}^{\text{Measurement variance}}$$

V_A	= additive genetic variance
V_D	= deviation due to dominance and epistasis
V_E	= environmental variance
V_I	= interaction variance
V_G	= $V_A + V_D$
Cov_{GE}	= covariance of genetics and environment

Heritability

- 描述一个给定的性状的变异中有多少可以归因于基因的变异（基因factor的权重）

$$h^2 = V_G/V_p$$

V_G代表基因，V_p是这个表型总的variance

After Twin studies

- the disease concordance can be used as the importance of genetic factor(占比)
- $$h^2 = 2 \times (C_{MZ} - C_{DZ})$$

	Concordance Rate		
Trait or Disease	MZ twins	DZ twins	Heritability
Alcoholism	0.6	0.3	0.6
Autism	0.92	0.2	>1
Cleft lip/palate	0.38	0.08	0.6
Diabetes, type 1	0.35-0.5	0.05-0.1	0.6-0.8
Diabetes, type 2	0.7-0.9	0.25-0.4	0.9-1.0
Measles	0.95	0.87	0.16
Schizophrenia	0.47	0.12	0.7

Odds ratio

	患病	不患病
有基因A	a	b
无基因A	c	d

- odds ratio
  - $$= \frac{(a/c)}{b/d} = \frac{ad}{bc}$$
    - if odds ratio > 1: the genetic can improve the disease(positive relation)
    - If odds ratio < 1: the genetic can protect the disease(protective)