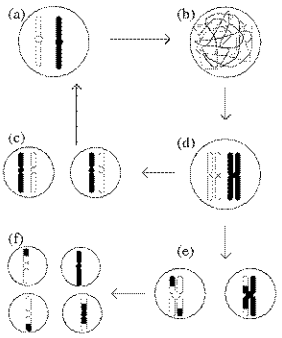
Chapter 4: Multilocus Data

4.1 Meiosis, Genetic Maps, Interference
4.2 Multilocus Recombination and Linkage
4.3 Big Pedigrees: computations on graphs
4.4 Monte Carlo methods on pedigrees

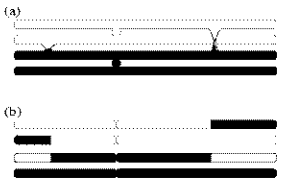
4.1 INHERITANCE AT LINKED LOCI
4.1.1 The processes of mitosis and meiosis



- (a) shows a pair of homologous chromosomes just after cell division.
- (b) shows the chromosomes in metaphase, at this phase we have 2m (yes!) of DNA in each cell nucleus.
- (d) shows the reconcentrated chromosomes prior to cell division. In phase (b) chomosomes duplicate, forming two *chromatids*, held together at the centromere.
- (c) Shows the result of *mitosis* – normal somatic cell growth and division. Each chromosome divides, with one chromatid going to each daughter cell, giving rise to new cells with identical original diploid genome content.
- NOTE: all DNA is double-stranded: at stage (d) there are 8 DNA strands.
- For meiosis; (e) & (f), see next page.

The meiosis process and outcomes.

- In the first meiotic division, (e) of previous page, one chromosome goes to each daughter cell. In order to insure each cell gets correct haploid genome content, the homologues must align tightly.
- At this stage, they can exchange material at chiasmata points as shown. (Singular is chiasma.)
- In (f), after the second meiotic division, there are four gamete cells, each with full haploid genome complement ....
- but, often with large segments (approx. 10^8 base pairs) from each of the parental homologues.
- For sperm cells all become sperm, but for egg cells just one becomes an egg and the other three are discarded.



The outcomes of the process of meiosis, shown for a single pair of homologous chromosomes in the nucleus of a cell of a diploid organism:

(a) Corresponds to stage (e): *chiasmata* are shown

(b) Corresponds to final gametes (f): *crossovers* are shown.

4.1.2 Genetic **distance** and Mather's formula

- The set of four chromatids (potential gamete chromosomes) is the tetrad.
- Each chiasma involves 2 of the 4 (i.e. 1/2) the potential gametes. In these gametes, a chiasma results in a crossover. The genetic distance d in Morgans is the expected number of crossovers between the loci on a given gamete.
- Hence:
- 1. The expected number of chiasmata between the loci on the tetrad is 2d.
  2. Genetic distance is ALWAYS additive, since expectations are additive. (Note we usually we measure genetic distance in centimorgans:
  100 cM = 1 Morgan.)
  3. Genetic distance has little to do with physical distance; but 1cM≈10^6 bp.
- Mather's formula (1938): No chromatid interference means each chiasma results in a crossover in a given gamete, independently with probability 1/2. Assuming this, in a given chromosome interval of genetic length d, suppose there are N(d) chiasmata. (N(d) can have any probability dsn.)
- Then: If N(d)=0, there are no chiasmata, no crossovers, and hence no recombination. If N(d)= n > 0, the probability of an odd number of crossovers is 1/2. (See homework). Thus we have Mather's formula:
  ρ(d) = (1/2) P(N(d) > 0) = (1/2) (1 - P(N(d) =0)).
- The only assumption here is the absence of chromatid interference: in this case ρ(d) is an increasing function of d, and is bounded above by 1/2.

4.1.3 Map functions: Haldane's map function

- ρ(d), as a function of genetic distance d is the *map function*.
- In Haldane's model (1919), crossovers are assumed to occur as a Poisson process, rate 1 (per Morgan). Thus there is no *interference*.
- The number of crossovers C(d) is Poisson with mean d, the numbers of crossovers in disjoint intervals are independent, and, conditionally on the number occurring, their locations are uniformly and independently distributed.
- Under Haldane's model, ρ(d) is the probability that a Poisson random variable with mean d is odd:
  ρ(d) = Σ_{k odd} exp(-d) d^k / k!
  = (1/2) exp(-d) Σ_{k=0}^∞ ((d^k / k!) – ((-d)^k / k!))
  = (1/2) (1 - exp(-2d)).
- Note that, under this model, ρ(d) is an increasing function of d, ρ(d) → 1/2 as d → ∞, and ρ(d) ≈ d as d → 0.
- Note also that under Haldane's model, number of chiasmata N(d) is Poisson mean 2d. Then P(N(d) =0) = exp(-2d); Mather's formula applies.

4.1.4 Interference and other map functions

- In fact, interference exists, mainly in chiasmata inhibiting the nearby presence of others, and hence also of nearby crossovers.
- Also, for reliable meiosis there needs to be at least one chiasma on every chromosome pair. This means that even the smallest chromosomes (in bp) have genetic length 50 cM: one half the gametes have a crossover.
- There are lots of models -- the model determines the map function. The reverse is not true.
- Historically, people would use their favorite map function to transform ρ to additive genetic distance d. Another important classical map function is the Kosambi map function: many older published genetic maps used this map function.
- However, if d is small, ρ(d) ≈ d, and nowadays the distance between adjacent markers in small.
- The important point is how multilocus computations are done, not what map function is used; almost all multilocus computations assume no interference.