Lecture 6 – Feb 8 – Markov chains continued

AA×BB to get an F1 with a genotype of AB. Recombination in the production of the F2 allows you to learn the map.

Let π_i be the equilibrium relative frequency of state i. So, if the chain is irreducible, $\pi_i = \mathbb{P}(s_\infty = i)$ where s_∞ is the state at some infinitely large time span.

$$\begin{aligned} \pi_i &= \sum_{j \in \mathcal{S}} \pi_j \mathbb{P}(s_{t+1} = i \mid s_t = j) && (1) \\ &= \sum_{j \in \mathcal{S}} \pi_j p_{ji} && (2) \end{aligned}$$

If our data is the genotypes along a chromosome, then our state space is $\mathcal{S} = \{AA, AB, BB\}$. So we can say, that if we choose very tight spacing of markers, then the probability of a recombination event in one chromosome, r, is very small so that we can ignore 2 recombination events in one interval:

$$\begin{aligned} \mathbb{P}(s_{t+1} = BB \mid s_t = BB) = \mathbb{P}(s_{t+1} = AA \mid s_t = AA) &= 1 - 2r && (3) \\ \mathbb{P}(s_{t+1} = AB \mid s_t = BB) = \mathbb{P}(s_{t+1} = AB \mid s_t = AA) &= 2r && (4) \\ \mathbb{P}(s_{t+1} = AB \mid s_t = AB) &= 1 - 2r && (5) \\ \mathbb{P}(s_{t+1} = BB \mid s_t = AB) = \mathbb{P}(s_{t+1} = AA \mid s_t = AB) &= r && (6) \end{aligned}$$

This assumes no segratation distortion. By using a Markov chain, we assume no interference.

$$\begin{aligned} \pi_{AA} &= \sum_{j \in \mathcal{S}} \pi_j p_{ji} && (7) \\ &= \pi_{AA}(1 - 2r) + \pi_{AB} r && (8) \\ \pi_{AB} &= \pi_{AA} 2r + \pi_{AB}(1 - 2r) + \pi_{BB} 2r && (9) \\ \pi_{BB} &= \pi_{BB}(1 - 2r) + \pi_{AB} r && (10) \\ & && (11) \end{aligned}$$

By algebra:

$$\begin{aligned} \pi_{AA} &= \pi_{AA}(1 - 2r) + \pi_{AB} r && (12) \\ 2r\pi_{AA} &= r\pi_{AB} && (13) \\ 2\pi_{AA} &= \pi_{AB} && (14) \end{aligned}$$

So, $\pi_{AA} = \pi_{BB} = 0.25$ and $\pi_{AB} = 0.5$