PBL 7: Discovering Permutations; Modeling A Genome

Pevzner suggests (p. 178ff) that the order of genes in an organism is represented by a permutation, presumably of an originally defined order that can be represented as the n-tuple $(1, 2, 3, \ldots, n)$. The current order could be written $(\pi_1, \pi_2, \ldots, \pi_n)$, where π_i is in $\{1, 2, \ldots, n\}$ and $\pi_i \neq \pi_j$ if $i \neq j$.

He defines a particular kind of permutation on n things, which he calls a *reversal*, and is also known as an *inversion* in the realm of biology, but not in the realm of mathematics. A reversal changes the order of the elements of an n-tuple by writing $(\pi_1, \pi_2, \ldots, \pi_i, \pi_{i+1} \ldots, \pi_{j-1}, \pi_j, \ldots, \pi_n)$ as $(\pi_1, \pi_2, \ldots, \pi_j, \pi_{j-1} \ldots, \pi_{i+1}, \pi_i, \ldots, \pi_n)$, where $i < j$.

To study the effects of reversals and the model of a genome as a permutation, you should investigate the following questions.

1. How many permutations of n things are there? How many that leave one element of an n-tuple fixed?

2. The set of permutations on n things is called the symmetric group of order n, and denoted S_n. Show how to construct all the elements of S_n given the elements of S_{n-1}. Use this scheme to write out the elements of S_2, S_3, S_4, S_5 as n-tuples, given that $S_1 = \{(1)\}$.

3. What does it mean to say that S_n is a group?

4. What does it mean to say that S_n is generated by a set of permutations?

5. Which permutations of S_4 are obtainable using a single reversal? Can all permutations of S_4 be obtained using a sequence of reversals? How about S_5? S_n?

6. What biological mechanism supports the idea that reversals are an important concept to model?

7. Using the breakpoint graph technique, find the reversal distance and a corresponding sequence of reversals necessary to transform the sequence of genes in the third chromosome of the Olympic strain of *Drosophila pseudoobscura* into that of the Standard strain (see Dobzhansky and Sturtevant, 1938).