Questions Albers et al (2010) *Genome Research* 21:961-973.

Dindel: Accurate indel calls from short-read data

1. What are some of the challenges in calling insertion-deletion polymorphisms from short read data using mapping based approaches?

2. One of the steps in the proposed algorithm for the analysis of indels is to ensure that candidate events are left aligned. Why is that? Can you propose a strategy for left aligning candidate events?

3. The authors state that describing hypothesis as a set of candidate haplotypes makes it easier to model various sources of error… What sources of error are discussed? Why is it easier to model these in the context of candidate haplotypes?

4. When trying to discover insertion deletion polymorphisms, the authors state that it is important to also model haplotypes of nearby SNPs. Why is that?

5. What quality filters to the authors propose to further improve the quality of the list of insertion deletion polymorphisms they generate?

6. The authors model sequencing error as a function of homopolymer length. This allows for increased error rates near simple sequence repeats. What features of the sequencing error model might be missing from the proposed implementation?

7. When aligning reads to a candidate haplotype, the authors use the Viterbi algorithm to calculate the probability of the optimal alignment of each read to a candidate haplotype. What are some advantages and disadvantages of this algorithm?

8. What are the key features of figure 6?

9. What struck you most about the paper?