

**Statistics For Genomics – Reading Assignment 2**  
**Due: May 2<sup>nd</sup> At the Beginning of Class**

**Read the papers (available from the class website):**

- (1) “Automated gating of flow cytometry data via robust model based clustering”
- (2) “Per-channel basis normalization methods for flow cytometry data”

**Briefly Answer Each of These Questions:**

- (1) What are the advantages of automated gating versus manual gating?
- (2) In the first paper, why do the authors use a mixture of t-distributions rather than a mixture of normal distributions for clustering? What is the difference between the t-distribution and the normal distribution?
- (3) In the second paper, how do the authors identify landmarks? How do they register peaks?
- (4) Can you think of some potential problems with the landmark registration and alignment algorithm?