Study Questions for "Gene set enrichment analysis for genome-wide DNA methylation data" by Maksimovic et al.

1. What are the main challenges in linking differentially methylated CpGs to functional effects via genes?

2. How do GOmeth and GOregion adapt existing enrichment tests to account for these challenges?

3. How is GOmeth different from GOregion? Why might the pattern in Figure 5A be less strong than the similar trend in Figure 1B?

4. How do the authors argue that their methods are better than others? Are you convinced? Would you use this method?

5. Do you think selecting random sets of individual CpGs to be differentially methylated in the creation of null simulations is an acceptable simplification? Why or why not?

6. How is the truth defined for the evaluations based on analyses of real data? What are the pros and cons of this approach?

7. **Challenge question:** Should we be concerned that the GOmeth approaches applied to data from null simulations are all below 0.05 on the y-axis in Figure 3A? Why or why not?