Announcements 3/6

- Read Angermueller et al. (2016)
  - Deep learning for computational biology
  - Can skim the image analysis sections
- Optionally read Ward and Kellis (2012)
  - Interpreting noncoding genetic variants
- Optionally read Zhou and Troyanskaya (2015)
  - DeepSEA
- HW2 due Thursday
  - Track your late days, see policy on website

Midterm study guide

- Midterm on Tuesday, March 13 in class
- No printed materials allowed
- No calculations
- Covers material through Thursday's lecture (epigenetics, Gaussian process, PIQ, and ROC curves but not non-coding variants and deep learning)

Midterm study guide

- Content includes (but is not limited to) material from lecture and required reading:
  - Bailey and Elkan (1995)
  - Lawrence et al. (1993)
  - Elemento et al. (2007)
  - Storey and Tibshirani (2003)
  - Sherwood et al. (2014)
  - Lever et al. (2016)
- Optional reading will also be helpful, provides background, details, and motivation not written in lecture slides

Midterm study guide

- Focus on terms, concepts, strengths/weaknesses, algorithmic strategies
  - Why or in what cases would we use one method or experimental technology instead of another?
  - How does method X compare to method Y and what are the unique advantages of each?
  - What is important concept Z (e.g. mutual information)?
  - What does a method optimize? How does it converge?
  - What assumptions does a method make?
  - How does a method implement important concepts and strategies?
  - What do we need to consider when analyzing a particular type of data?

Midterm study guide

- Example questions
  - What different assumptions to the MEME OOPS and ZOOPS models make?
  - Which motif-finding algorithm is best if the user wants to account for the probability a motif was generated by the background distribution?
  - What is FWER and what does it measure?
  - Draw an example of a dependency among two continuous variables that could be detected with mutual information but not Pearson's correlation.
  - Why is DNA methylation status valuable when interpreting gene regulation?