1158-92-283 **H. Tomas Rube*** (trube@ucmerced.edu), UC Merced, School of Engineering, 5200 N. Lake Road, Merced, CA 95343, and **Chaitanya Rastogi**, **Judith F. Kribelbauer**, **Xiaoting Li** and **Harmen J. Bussemaker**. *Quantifying sequence readout by transcription factors through principled analysis of high-throughput SELEX data.* Preliminary report.

Transcription factors (TFs) are central to cellular information processing and control gene expression by binding specific DNA sequences in the genome. In recent years, the sequence specificity of hundreds of TFs have been characterized using high-throughput in vitro assays that utilize deep sequencing to quantify DNA binding. Variations of these assays can characterize binding by a single TF, multiple TFs forming a complex, RNA-binding proteins, or binding to chemically modified DNA. However, no unified mathematical method for analyzing all these data yet exists. Challenges in developing such a method include both the diversity and complexity of the experimental designs, and the sparsity of the data, where each DNA sequence is observed at most once. In this talk I will introduce a principled modeling framework, ProBound, that overcomes these challenges and infers quantitative models of TF binding specificity. The flexibility of ProBound allows it to work on all currently available data. ProBound uses multi-task learning to infer robust binding models from multiple experiments, a feature especially useful for analyzing the binding specificity of TF complexes. I also will discuss similarities and differences between ProBound and deep learning. (Received March 03, 2020)