GGSB PRELIM QUESTION # 10

A major challenge in computational genomics is predicting the functional consequences of variant alleles, especially for non-coding variants. The most advanced approaches should in principle integrate information from comparative genomics and functional genomics. Yet, combining these disparate sources of data is challenging. One recent approach to this problem is embodied in the software LINSIGHT, described by Huang et al in *Nature Genetics* (2017, 49:618-624).

a. Describe at least two previous attempts to address this same broad problem and briefly describe the conceptual and computational frameworks each used.

b. For LINSIGHT, draw the graph for the model used in the paper on the board (Figure 1b) and explain the various components.

c. Now that you've described the model, describe the inference machinery developed to estimate the relevant parameters that are used for prediction.

d. Provide your evaluation of the model in terms of its biological realism.

e. Suggest a line of research that would allow you to further evaluate the efficacy of a tool like LINSIGHT.

Huang et al in *Nature Genetics* (2017, 49:618-624).