

**Faculty Name:**

David Knowles

**Faculty Email:**

dak2173@columbia.edu

**Lab:**

Knowles Lab

**Project Title:**

Unraveling the genetic basis of neurological disorders with machine learning

**Description:**

RNA splicing, the cellular process by which "junk" intronic regions are removed from initially transcribed RNA, is tightly regulated in healthy human development but frequently dysregulated in disease. This complex process involves hundreds of proteins and non-coding RNAs, so that predicting the effects of mutations on outcome is beyond current biophysical models. We instead take a data-driven approach: leveraging large-scale RNA-seq and massively parallel reporter assay data to train deep neural networks to predict splicing directly from gene sequence. This project will extend and apply such models to better understand the genetic basis of ALS, Alzheimer's and/or autism using large-scale whole-genome and RNA sequencing data we have access to through collaborators at the New York Genome Center.

**Location of Research:**

Hybrid (both Remote and On Site)

**# of hrs/week:**

40

**Department/Program:**

Computer Science

**Eligibility:**

MS

**To apply, please contact:**

David Knowles

dak2173@columbia.edu