· Project Summary/Abstract: 30 line limit

· Project Narrative: 2-3 sentence maximum

· References/Bibliography – List all authors. No et als.

[Project Summary/Abstract:](#_jofmliumgnd0)

[Project Narrative](#_h8k7twug9wrt)

[References/Bibliography](#_ismsyit6ric3)

# Project Summary/Abstract:

Cancer genomics has revealed that there are often millions mutations per genome[HY1] but only a small fraction of them are in coding regions. Yet, almost all of the known driver mutations in cancer are in coding regions. Many tools have been developed to identify drivers and prioritize non-coding mutations. But currently, the relative impact of coding and non-coding variants, especially how to prioritize them in a uniform context, is largely unknown. In this proposal, it include three aims. The core components of the first aim will define mathematical model that employ various features are related to the molecular the in the impact on molecular activity or molecular endotype and the phenotype of both coding and non-coding variance these include things like where they sit in the network where they break TF motif and so forth. And it also combines information about recurrence in large cohorts. We will provide an innovative scheme to update our model in a Bayesian framework using large-scale experimental data. The update and the validation will lead to a more accurate and usable model. The second aim, high-throughput molecular experiments: eSTARR-seq and InPOINT, will be used to quantify and measure the effect of non-coding and coding variants respectively. The output of the activity will be fed back for paramter tuning of the model. The last aim, we will focus this on a particular subsystem related to prostate cancer but is that related to p53 and RB. We will use CRISPR to build a cellular variants impact evaluation model and introduce targeted mutation in coding and noncoding regions from normal prostate cell lines, which will grow in prostate organoid to investigate tumor progression effect.

# Project Narrative

We will firstly develop a mathematical model that encompasses a wide range of genomic features combines information from both the molecular, nucleotide-level scale (biochemical/biophysical, evolutionary, and network) with information about recurrence and whole-organism disease phenotype. Then we will use eSTARR-seq and InPOINT technology approach to quantify molecular activity for noncoding and coding variants respectively, the outcome of which can be used for bayesian parmeter tuning of our mathematical model. Last, we will focus on variants in targeted genes and regulatory elements and use CRISPR to introduce celluar mutation and then grow the cell in prostate organoid to investigate tumor progression effect.

# References/Bibliography