**Systems Genetics Research Plan**

**Li Luo**

System genetics is totally a new domain of study for me. It’s now a popular topic as it takes statistical analysis into studying the relationship between phenotypes and genotypes. After reading the materials provided, I think I can divide this project into 2 parts: i) exploring the correlation between different genotypes; ii) building a model that can link phenotypes and genotypes.

**Plan**

* Data cleaning

Change the data into the format and structure that is suitable to do further study.

* Exploring the correlation between different genotypes (the results are supposed to be: the closer two markers are physically to each other, the higher is their correlation)

Correlation between genotypes in the same chromosome

Correlation between genotypes in different chromosome

Relationship between genotypes and phenotypes

* Building a model that can link phenotypes and genotypes (pay special attention to multicollinearity)

Try different models from the simple ones to more complex ones such as PCA, tree methods etc.