600.438 Computational Genomics
Project Proposal

Sachi Sanghavi

April 28, 2016

1 Motivation

Aging is a universal biological process; and is known to be a major risk factor in diseases like type-2 diabetes, cancer, Alzheimer's disease, and more. However, the underlying molecular mechanisms of aging are still not well understood. This leaves the question "Is aging genetically programmed?" open for debate. Studying the association between aging and gene expression might give us some insights which could lead to novel treatments for treating age-related diseases.

We propose to study aging – gene expression association in humans [in subcutaneous adipose tissue and whole blood tissue], and its link to disease.

2 Background

The study is based on Yang, *et al.* (2015). The major difference is that we have *more* samples for the same tissues. As as a result, it'd be interesting to compare our results.

3 Aims and Methods

1. Find age – gene expression association and report it for top 100 genes for subcutaneous adipose tissue and whole blood tissue
    - Use linear regression (correct for gender and top 3 genotype PCs, and ensure gene expression PCs have p-value > 0.05), bootstrapping (remove 20% low expressed genes), and permutation (estimate false positives).
2. Functional annotation of aging genes [up- and down-regulated separately]

- Use David tools.

3. Tissue specific link between aging genes and complex disease genes.
   - Disease gene data compiled from NIH GWAS and OHIM using clustering.
   - Fisher's test for aging – diseases genes link.
4. Compare results with Yang, *et al.* (2015).

4 Data

Gene expression data is taken from GTEx project v6 (PC values for top 3 genotypes are available; labeled C1, C2, and C3 in figure 1).

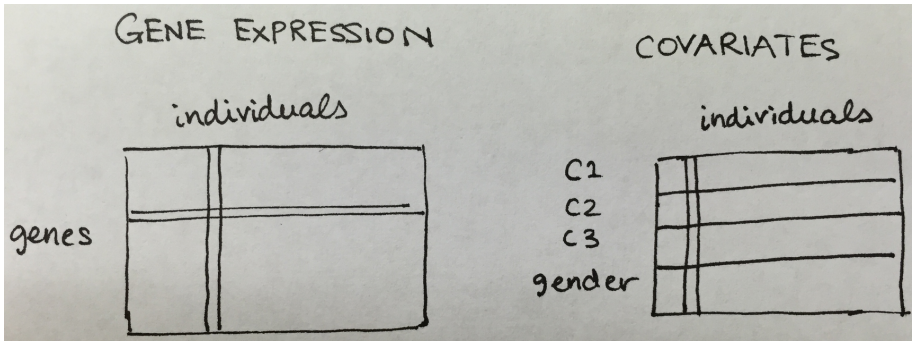

Figure 1: Format of the two main data files from GTEx