Genomic SVD

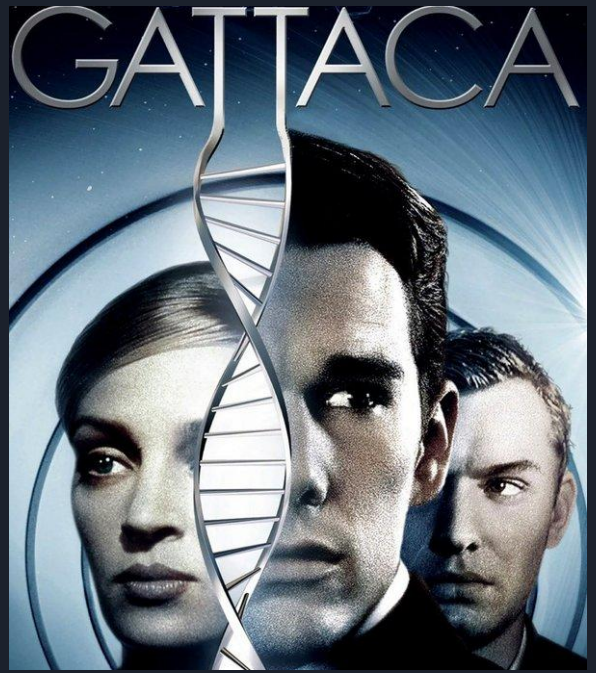
https://github.com/kstisser/GeneSVD

By Karissa Stisser

Genomic Applications

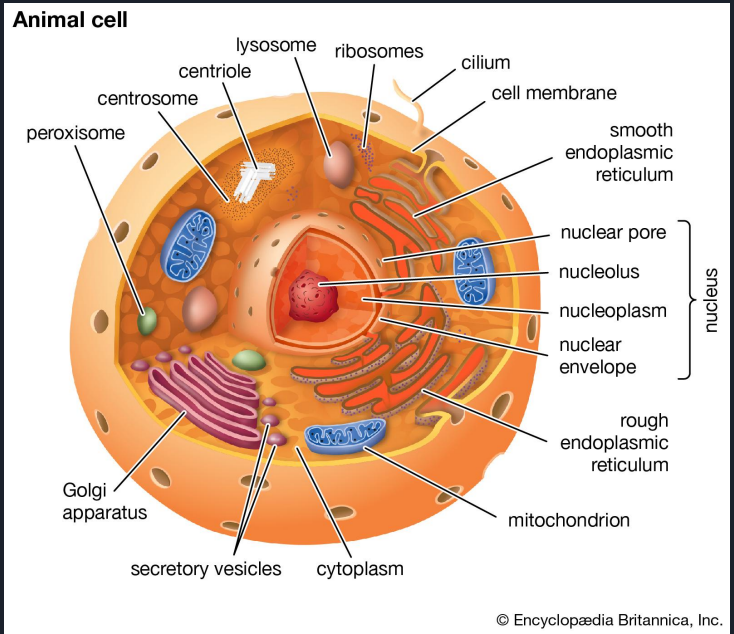
- Precision Medicine
- Diagnosis of genetic disorders
- Vaccine development
- Genetically modified produce
- Differences in humans & animals
- Crispr

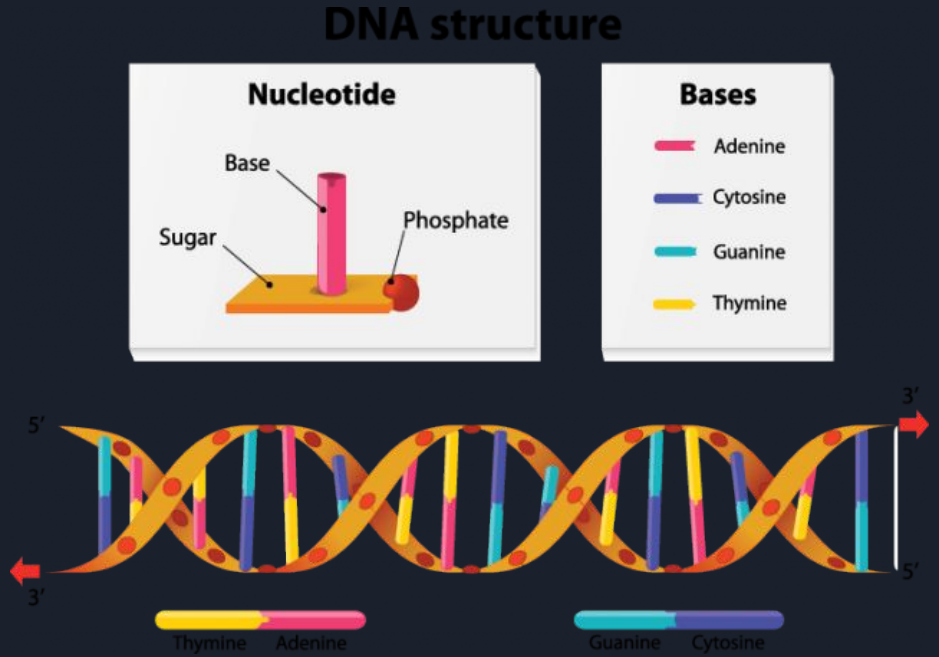
https://github.com/kstisser/GeneSVD



https://github.com/kstisser/GeneSVD

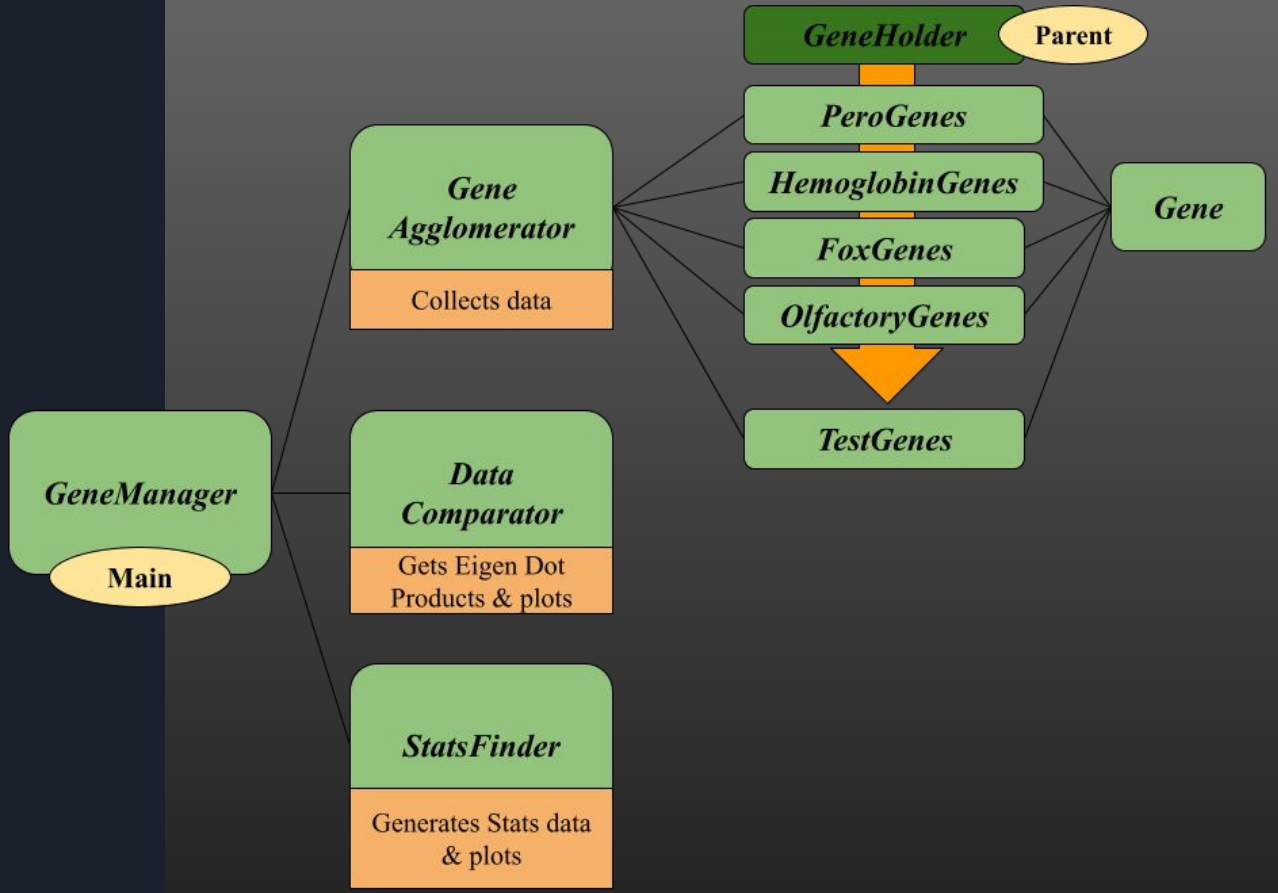
Remembering Biology



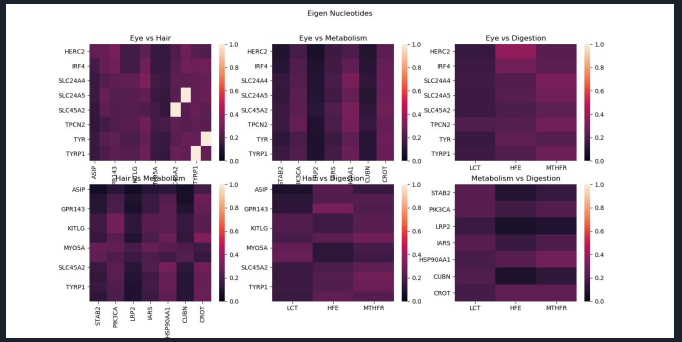


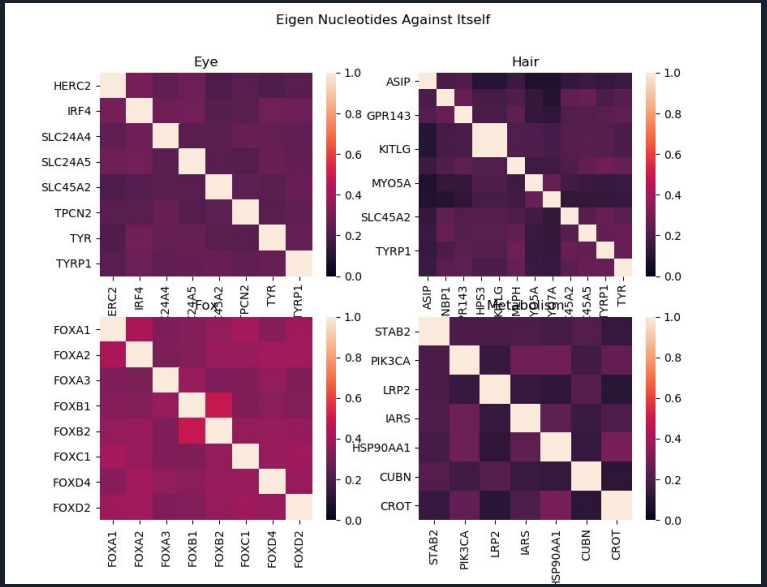
Project Outline

- Collects all data
  - One hot encoding (OHE)
    - SVD(gene)
    - SVD(family)
  - Numeric
    - SVD(family)
- Statistics plots
  - ATCG ratios
  - SVD ratios vs seq length
  - Singular value comp.
- Dot Product Plots
  - Sanity test check
  - Family vs Family (genes)
  - Family vs Itself
  - Family vs Family (combined)



Gene Families





Peroxeridoxen

- Antioxidant genes
- (6)- 3 classes

Hemoglobin

- Blood genes
- (3)- 2

Fox

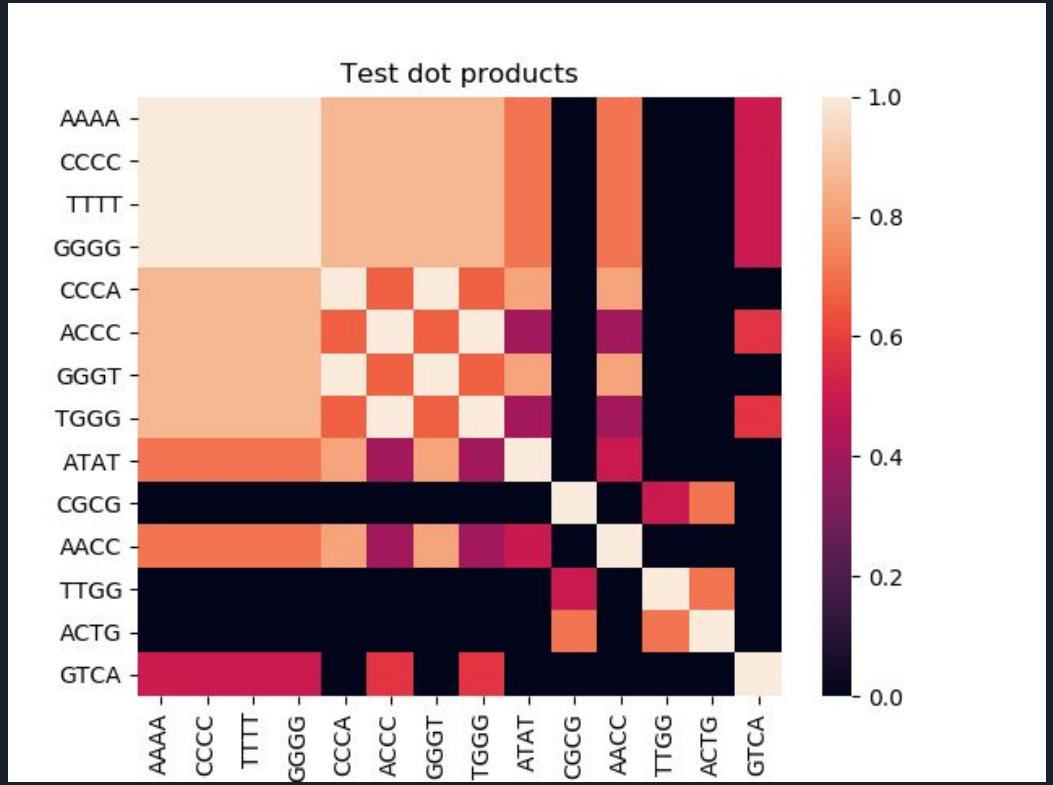
- Regulates gene expression
- (8)

Olfactory

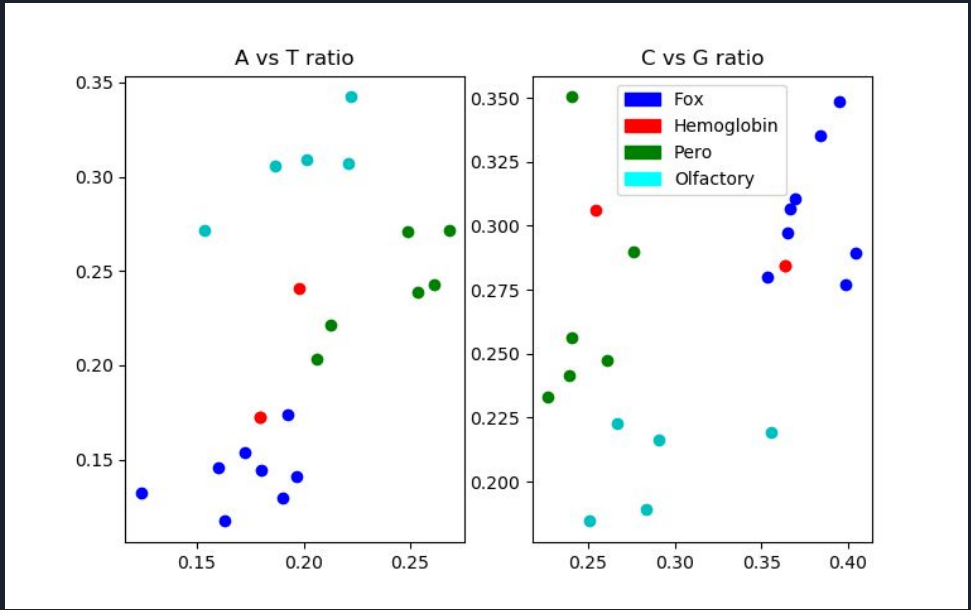
- Smell genes
- (5)- 2 smells

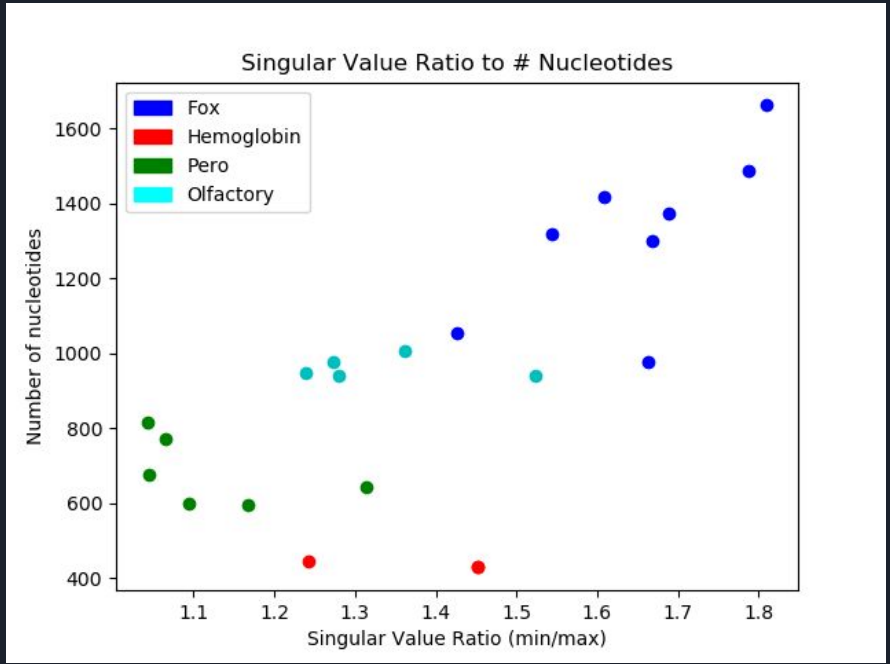
Sanity Test Results

- Captures patterns
- Doesn't care as much about content matching
- Example: GGGT=CCCA

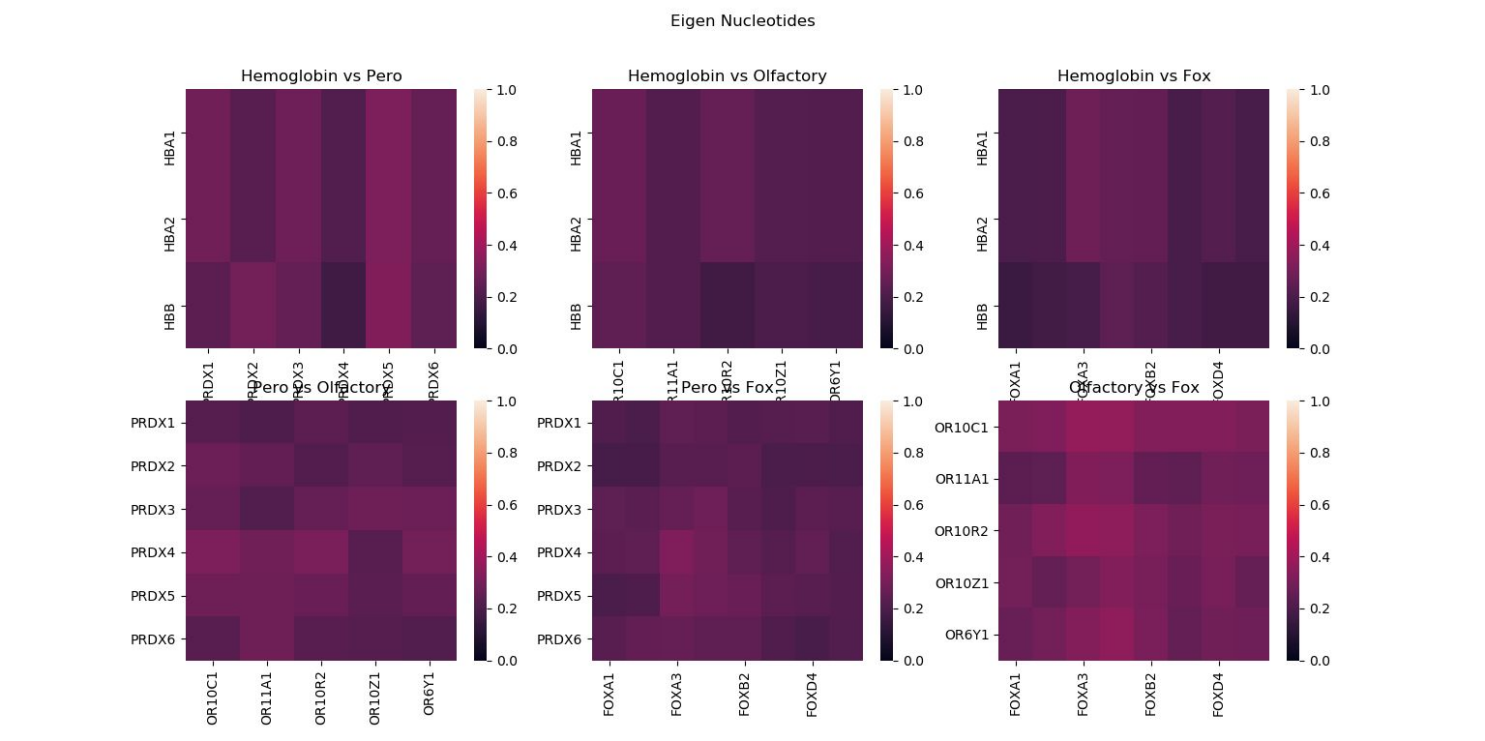


Statistics & Trends

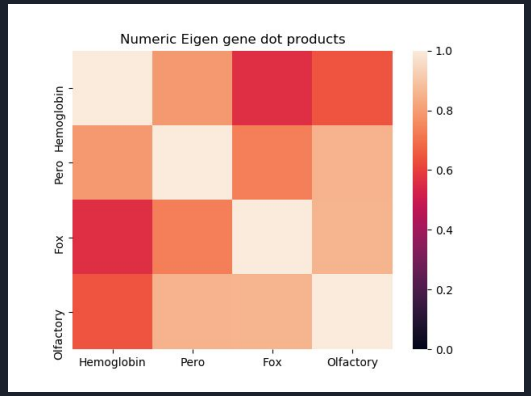


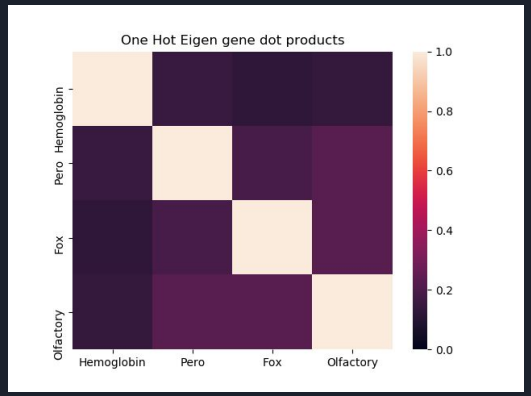


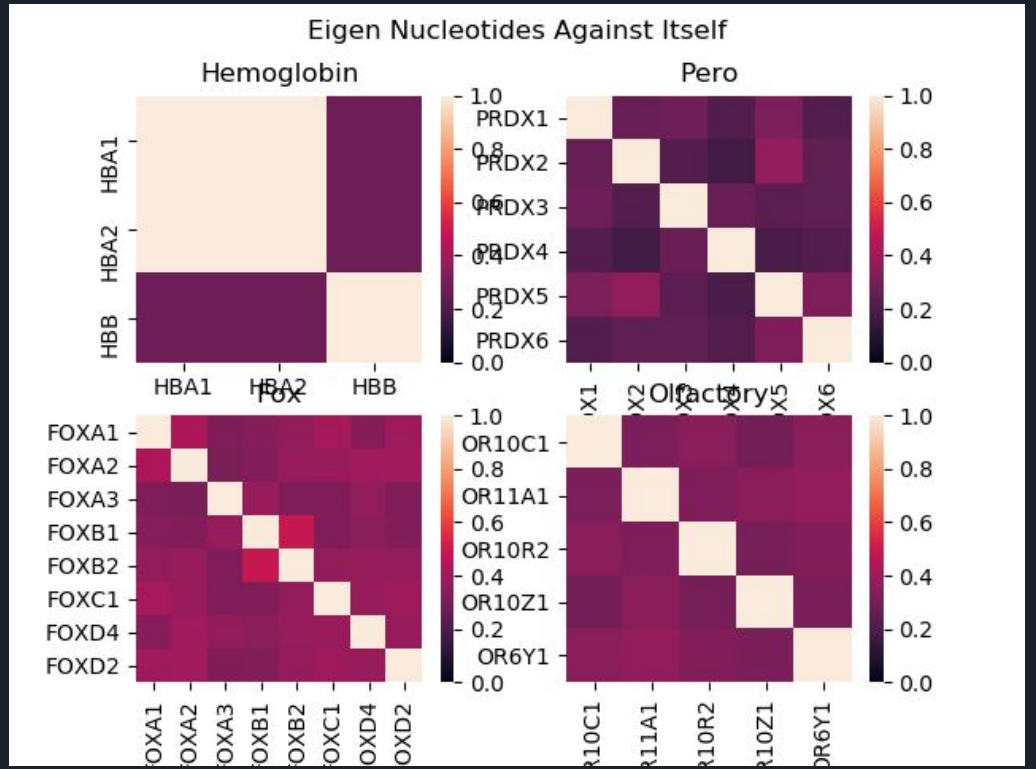
Family vs Family EigenGenes Dot Product



Fam vs Fam (Numeric & 1 hot) & Fam vs Itself







https://github.com/kstisser/GeneSVD

Singular Values

```
##################################################
('We have ', 22, ' total genes to analyze')
##################################################
('We have ', 26, ' total test genes')
('We have ', 8, ' total Fox genes')
('We have ', 3, ' total Hemoglobin genes')
('We have ', 6, ' total Pero genes')
('We have ', 5, ' total Olfactory genes')
##################################################
**********************************************************
Normalized Singular values of all data:
[8.96016033e+01 3.11707789e+01 1.96865756e+01 1.42475760e+01
 1.24091415e+01 1.12388673e+01 1.07229177e+01 9.99065572e+00
 9.71815886e+00 9.61244487e+00 9.40612756e+00 8.65868885e+00
 8.45112137e+00 8.40583552e+00 8.04384751e+00 7.78282335e+00
 7.55029734e+00 7.45568953e+00 7.08111018e+00 6.33352075e+00
 6.17620169e+00 9.41781545e-15]
**********************************************************
```