St. Jude Children's®
Research Hospital

St. Jude BioHackathon

Title

A web app for fine-scale population/ethnicity identification and visualization

Category

GUI Tool Development

Challenge

Population/ethnicity identification is important in analysis of germline variants or associations with phenotypes because the allele frequencies can very significantly among different populations or ethnicities. Identification of population/ethnicity usually takes many steps of processing. Interactive visualization of population/ethnicity is lacking and will be useful for illustrating the fine-scale population structure.

Benefit

The tool can be used by anyone that needs to identify the population/ethnicity and visualize them, for example, in clinical/research evaluation of rare germline variants of patients

Helpful Tools, Packages, or Software

Hail, python, R, UMAP, TSNE, RShiny or related for a web-based app.

Test Data

gnomAD, 1000 genomes/HGDP WES/WGS data matrix if we need to build our own population/ethnicity classifier.