Software and Datasets

We develop and maintain a variety of software, genomic annotations and web services.

Annotation Datasets

- **Roulette** is a mutation rate model identifying the mutagenic effect of Polymerase III transcription at the basepair resolution.
- **$S_{\text{het}}$$S_{\text{het}}$$S_{\text{het}}$$S_{\text{het}}$ are gene-based estimates of selection strength.

Tools for Assessment of Variants

- **PolyPhen-2** is a tool for predicting the effect of an amino acid substitution on protein structure and function, based on comparative genomics and experimentally determined protein structures. It is available as a web service, and can also be downloaded as a standalone application.
- **SNPtrack** is a simple interface for mutation mapping and identifying causal mutations from whole-genome sequencing studies. It is available as a web service.

Tools for Mass Spectrometry and Proteomics

- **MS-BLAST** is a tool for searching protein sequences identified with tandem mass spectrometry against databases of protein sequences. It is available as a web service and as a standalone software.

Tools for Statistical Genetics

- **Joint Likelihood Mapping (JLIM)** is a tool to test for shared genetic effect between two genetic association data, for example, a disease GWAS study and gene expression QTL (eQTL) study.
- **Joint Likelihood Mapping 2 (JLIM_2.0)** is a version of JLIM which supports meta-analysis across more than one cohort of matching ancestry.
- **Joint Likelihood Mapping (JLIM) 2.5** is a new version of JLIM based on summary statistics.
• **NPS** is a tool for polygenic risk scoring based on partitioning-based non-parametric shrinkage algorithm.

• **RVTT** is a novel statistical test of trend that assesses the relationship of the frequency of qualifying rare variants in a pathway with dichotomous disease phenotypes leveraging the Cochran-Armitage test statistic.

### Tools for Cancer Genomics

• **MutPanning** is designed to detect rare cancer driver genes from aggregated whole-exome sequencing data.

• **CBaSE** enables cancer type and gene-specific estimation of the strength of negative and positive selection. It is available as a browser-based tool as well as for download as a standalone package.

### Tools for Population Genetics

• **simDoSe** is a fast and flexible Wright-Fisher simulator for arbitrary diploid selection evolving through realistic human demography.