FILER: a framework for harmonizing and querying large-scale functional genomics knowledge

PNGC is pleased to announce the release of the FILER, a functional genomics repository developed by NIAGADS, now published in NAR Genomics and Bioinformatics.

Functional genomic and annotation data such as tissue-specific regulatory/enhancer elements, transcription factor binding, chromatin states and interactions are widely used in systems biology, genetic and genomic studies, e.g., to interpret non-coding genome-wide association study signals or to characterize the experimentally identified genomic regions.

We built FILER to provide a scalable, unified, high-throughput and robust access to massive, heterogeneous functional genomic (FG) data collections (> 59,000 datasets) across > 1,000 tissue/cell types curated, harmonized, and integrated from > 20 data sources including ENCODE, GTEx, FANTOM5, NIH Roadmap Epigenomics and other large-scale projects.

All data in FILER can be queried by tissue/cell type, biological sample, assay, genomic feature type and other data attributes. Importantly, genomic queries by intervals/regions of interest are supported with high efficiency thanks to the FILER genomic indexing and search engine. In addition to uniquely providing harmonized FG and annotation data in uniform, consistent data formats, FILER provides pre-processed data per tissue/cell type, to allow users to customize which tissues/cell types to include depending on their research questions.
FILER is also available as a stand-alone version for offline, batch processing in cloud or high-performance computing (HPC) environments. For example, FILER FG and annotation data can be integrated together with the investigation/user specific experimental data and used within custom / user high-throughput genetic and genomic analysis workflows.

New Datasets available at https://www.niagads.org/datasets

NG00115 - Similar Genetic Architecture of Alzheimer's Disease and Differential APOE Effect Between Sexes-Wang et al. 2021

NG00119 - Health and Retirement Study Genotype Data 2006-2012