KG Lab ENCODE Submission Protocol

Version 1

1. Submit **EXPERIMENT form** containing pre-database-defined information such as “Award” (grant or funding information), “lab” (lab information), and etc. Will get an **”EXPERIMENT”** key from the database for later use. Each **EXPERIMENT** represents a group of samples either are biological replicates or technical replicates and the same target. For examples, ChIP-seq data “e10.5\_heart\_H3K4me3” and “e10.5\_limb\_H3K4me3” will be two experiments.
2. Submit **LIBRARY form** containing pre-database-defined information and get a **”LIBRARY”** key from the database. Each **LIBRARY** represents an individual/sample from one **EXPERIMENT**. For examples, ChIP-seq data “e10.5\_heart\_H3K4me3” might have 2 samples and will generate 2 library records such as “e10.5\_heart\_H3K4me3\_YS001” and “e10.5\_heart\_H3K4me3\_YS002”.
3. Submit **REPLICATE form** containing an user-generated unique **“ALIASES: ARRAY”** key which connect the **”EXPERIMENT”** key from step 1 and **”LIBRARY”** key from step 2 together. Each **REPLICATE** represents a single sample, which will link to one file at the last submission step. Each biological or technical replicates will have it’s own replicate recode.
4. Submit **FILES form** containing a unique 1-to-1 **“ALIASES: ARRAY”** key for each file and get a **“FILE”** key from the database. The **“FILE”** key will be used for the files submission to ENCODE.
5. Using **“FILE”** key to submit files to AWS.
6. Pre-defined information in ENCODE:
   1. <https://www.encodeproject.org/help/getting-started/#organization>
   2. <https://www.encodeproject.org/awards/>
   3. <https://www.encodeproject.org/labs/>
   4. <https://www.encodeproject.org/datasets/>
   5. <https://www.encodeproject.org/biosamples/>
   6. <https://www.encodeproject.org/targets/>
   7. <https://www.encodeproject.org/documents/>
   8. <https://www.encodeproject.org/projects/>
7. Pipelines
   1. <https://www.encodeproject.org/pipelines/>
   2. <https://www.encodeproject.org/atac-seq/#overview>
   3. <https://github.com/kundajelab/atac_dnase_pipelines>