**GMQL tool for the analysis of (epi)genomic data features:**

Ex.1 – Activator and Repressive epigenetic signals

Among the different epigenetic signals, *H3K27ac* and *H3K27me3* are generally associated with active and repressed chromatin regions, respectively. Considering *H3K4me1* in cell line *A549* and the aforementioned signals (*H3K27ac* and *H3K27me3*) in *broadPeak* format, under *ethanol treatment (EtOH)*:

1. Using GMQL, select the required ENCODE *ChIP-seq* data for the reference human genome *hg19* (originally from the @UCSC website), i.e. the *H3K4me1*, *H3K27ac* and *H3K27me3* under *ethanol treatment (EtOH)* in *broadPeak* format, and the *promoter* region annotation for the same reference human genome, by writing the required GMQL statements.
2. Then, compute the following:
   1. Active *H3K4me1* regions (i.e. overlapping with *H3K27ac* regions)
   2. Repressed *H3K4me1* regions (i.e. overlapping with *H3K27me3* regions)
   3. Poised *H3K4me1* regions (i.e. being simultaneously active and repressed)
   4. Active *H3K4me1* regions in promoters
   5. For each region in (d) find the closest *H3K4me1* region farther than 10 kb
   6. Store the last result (e)
   7. Run the created GMQL query and report: running time, obtained number of samples and number of regions in each sample. Can the result include replicated regions? Why?

Ex.2 – Differential binding

Consider the JUN Transcription Factor (TF) in cell line K562 in two different conditions: treatment with interferon alpha 30 minutes (IFNa30) and untreated.

Considering the untreated sample as the baseline:

1. Using GMQL, select the required ENCODE *ChIP-seq* data for the reference human genome *hg19* (originally from the @UCSC website), i.e. for JUN antibody\_target with *none*, or *IFNa30* treatment in *narrowPeak* format, and the *RefSeq* promoter region annotation for the same reference human genome, by writing the required GMQL statements. Combine replicas if needed.
2. Then, compute the following:
3. DNA regions in common with the baseline and the treatment
4. Considering the regions identified in (a), find the promoters in which they are present
5. Store the last result (b)