



St. Jude BioHackathon

Title

An interactive application to perform and interrogate peak calling performance with various parameters

Category

GUI Tool Development

Challenge

Peak calling is common for molecular profiling assays that generate signal buildup across the genome associated with a specific molecular feature of interest. These are assays like ChIP-seq, Cut and run, or ATAC-seq.

The issue with peak calling is that these signal profiles differ very wildly depending on what's being assayed. For example, profiling a TF by cut and run typically generates very sharp, distinct peaks with almost no background. On the other hand, ChIP-seq tends to have higher background, and both of these assays are subject to antibody quality.

In addition, you can have much more broad, large peaks that span several kb, such as for some histone marks. This means that peak calling is very much not one-size fits all and can lead to significant effort and frustration to determine the optimal peak caller and parameters for a given assay.

Tweaking parameters and re-running peakcallers is a slow and frustrating process, so ultimately the question is how can we make that process more convenient and quicker.

Benefit

Solving this would allow us to optimize peak calling for different antibodies and assays much more quickly and streamline the analysis process.

Helpful Tools, Packages, or Software

Singularity or Docker to containerize the peak callers (MACS v1.4, MACS2, MACS3, SEACR, etc). R shiny, dash python, or python QT for GUI generation. jbrowse or similar for interactive track views.

Test Data

ENCODE, pretty much any ChIP-seq, ATAC-seq, Cutnrun datasets from GEO.