CATCH your ChIP profiles!

Fiona G Nielsen[1,2], Kasper Markus[3], Henk Stunnenberg[2], Lene Monrad Favrholdt[3], Martijn Huynen[1]

[1]Department of Molecular Biology, Nijmegen Centre for Molecular Life Sciences, the Netherlands; [2]University of Nijmegen Medical Centre, the Netherlands; [3]University of Southern Denmark, Odense, Denmark.

In the field of epigenetics the regulation of genes is examined not at the level of DNA (genomics) but at the level of proteins associated to and interacting with the DNA. The distinct patterns of bound proteins are indicators of gene activity, as well as clues to decyphering the complex combinatorial regulatory machinery of the cell. Characterising epigenetic profiles from ChIP profiling experiments in a genome-wide scale is an overwhelming task that calls for automated data-mining techniques to aid analysis.

We present the program CATCH to perform unsupervised clustering of ChIP profiles. CATCH implements a hierachical clustering algorithm with an exhaustive all-pairs comparison at each clustering step. As part of the comparison CATCH simultaneously aligns the ChIP signal profiles, so the profiles are repeatedly clustered by their most informative alignment.

With CATCH we have analysed the ENCODE ChIP-on-chip data set from Heintzman et al [1]. Next to recovering the average promoter and enhancer profiles as described in [1], we discovered several other distinct epigenetic patterns. Our results include specific profile patterns for bidirectional promoters, as well as a pattern specific for promoters of the heavily repressed genes from the olfactory receptor gene family.

CATCH automates the clustering of ChIP profiles and enables easy browsing and export of results. This ease of analysis is paramount to manage the increasing volume of experimental ChIP profiling data.

CATCH is open-source and freely available from the authors homepage: www.cmbi.ru.nl/~fnielsen/CATCH

[1] Heintzman et al, Nature Genetics vol 39, 3, March 2007.