

CoBRA-Containerized Bioinformatics in ChIP/ATAC-seq

Ava James*

Department of Medical Oncology, Harvard Medical School, USA

*Corresponding author: Ava James, Department of Medical Oncology, Harvard Medical School, USA, Email- avajames@acd.edu

Received date: October 03, 2021; Accepted date: October 17, 2021; Published date: October 24, 2021

Copyright: © 2021 James A. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Commentary

Chromatin immunoprecipitation sequencing (ChIP-seq) and the Assay for Transposase-Accessible Chromatin with high-throughput sequencing (ATAC-seq) have grown to be crucial technology to successfully degree protein-DNA interactions and chromatin accessibility. However, there's a want for a scalable and reproducible pipeline that carries right normalization among samples, correction of reproduction quantity variations, and integration of recent downstream evaluation equipment.

Containerized Bioinformatics workflow for Reproducible ChIP/ATAC-seq Analysis (CoBRA), a modularized computational workflow which quantifies ChIP-seq and ATAC-seq height areas and plays unsupervised and supervised analyses [1]. CoBRA gives a complete modern-day ChIP-seq and ATAC-seq evaluation pipeline that may be utilized by scientists with confined computational revel in. This allows researchers to advantage speedy perception into protein-DNA interactions and chromatin accessibility thru pattern clustering, differential height calling, motif enrichment, assessment of web web sites to a reference database, and pathway evaluation.

ChIP-seq is a high-throughput generation that offers specific insights into protein feature through mapping genome-extensive binding web web sites of DNA-related proteins. ATAC-seq is a high-throughput generation this is vital withinside the evaluation of genome-extensive chromatin accessibility. While several pipelines for reading ChIP-seq and ATAC-seq facts.

There stays a sturdy want for pipelines that may be run through customers who've confined revel in in using computational biology equipment. Comparisons among ChIP-seq and ATAC-seq experiments can offer perception into variations in protein occupancy, histone marks, and chromatin accessibility.

However, evaluation pipelines presently to be had lack beneficial additives important for such analyses. For example, there's a want for higher normalization among samples, adjusting reproduction quantity variations (CNVs), making use of newly advanced downstream annotation equipment which includes Cistrome DB Toolkit and integrating epigenetic facts with RNA-seq facts [2].

CoBRA can quantify ChIP-seq and ATAC-seq height areas, and carry out unsupervised and supervised analyses. It gives pattern clustering, differential height calling, motif enrichment and clustering, annotation of differential web web sites through a reference database, and pathway evaluation [3].

CoBRA makes use of Snake make, a workflow control machine to create the computational pipeline. Using the Snake make machine allows the reproducibility and scalability of CoBRA. This framework additionally lets in for the addition or substitute of evaluation equipment, in addition to for the parallelization of computationally extensive processes. To make CoBRA portable, the workflow and its software program dependencies are to be had as a Docker field, which may be used on any device with Docker installed. This consists of neighbourhood servers, high-overall performance clusters, and cloud-primarily based totally machines. Docker robotically downloads all required software program dependencies due to the fact the field encapsulates all the helping software program and libraries, putting off the opportunity of conflicting dependencies [4].

CoBRA robotically plays a de novo motif evaluation on every of the 3 clusters of available web web sites to discover motifs of ability transcriptional regulators enriched in differentially available chromatin elements. This evaluation diagnosed many transcription component binding motifs enriched in every cluster.

Some strategies fail to normalize ChIP-seq records correctly in calling differential peaks while the FRiP rating is impacted with the aid of using perturbations. CoBRA reduces fake positives and identifies greater authentic differential peaks with the aid of using correctly normalizing ChIP-seq records in step with sequencing depth.

Snakemake framework lets in for fast integration of latest strategies or substitute of present tools. Modules may be introduced actually with the aid of using including a brand new Snake make rule and including a flag with inside the config file. The framework lets in for high-quality flexibility due to the fact every module may be evaluated in its very own surroundings the use of unique tools [5].

References

1. Baca SC, Takeda DY, Seo JH, Hwang J, Ku SY, et al (2021). Reprogramming of the FOXA1 cistrome in treatment-emergent neuroendocrine prostate cancer. *Nat Commun* 12: 1-12.
2. Cejas P, Xie Y, Font-Tello A, Lim K, Syamala S, et al (2021). Subtype heterogeneity and epigenetic convergence in neuroendocrine prostate cancer. *Nat commun* 12: 1-11.
3. Min S, Lee B, Yoon S (2017). Deep learning in bioinformatics. *Brief Bioinformatics* 18: 851-869.
4. Stein L (2002). Creating a bioinformatics nation. *Nature* 417: 119-120.
5. Larranaga P, Calvo B, Santana R, Bielza C, Galdiano J, et al (2006). Machine learning in bioinformatics. *Brief bioinformatics*, 7: 86-112.