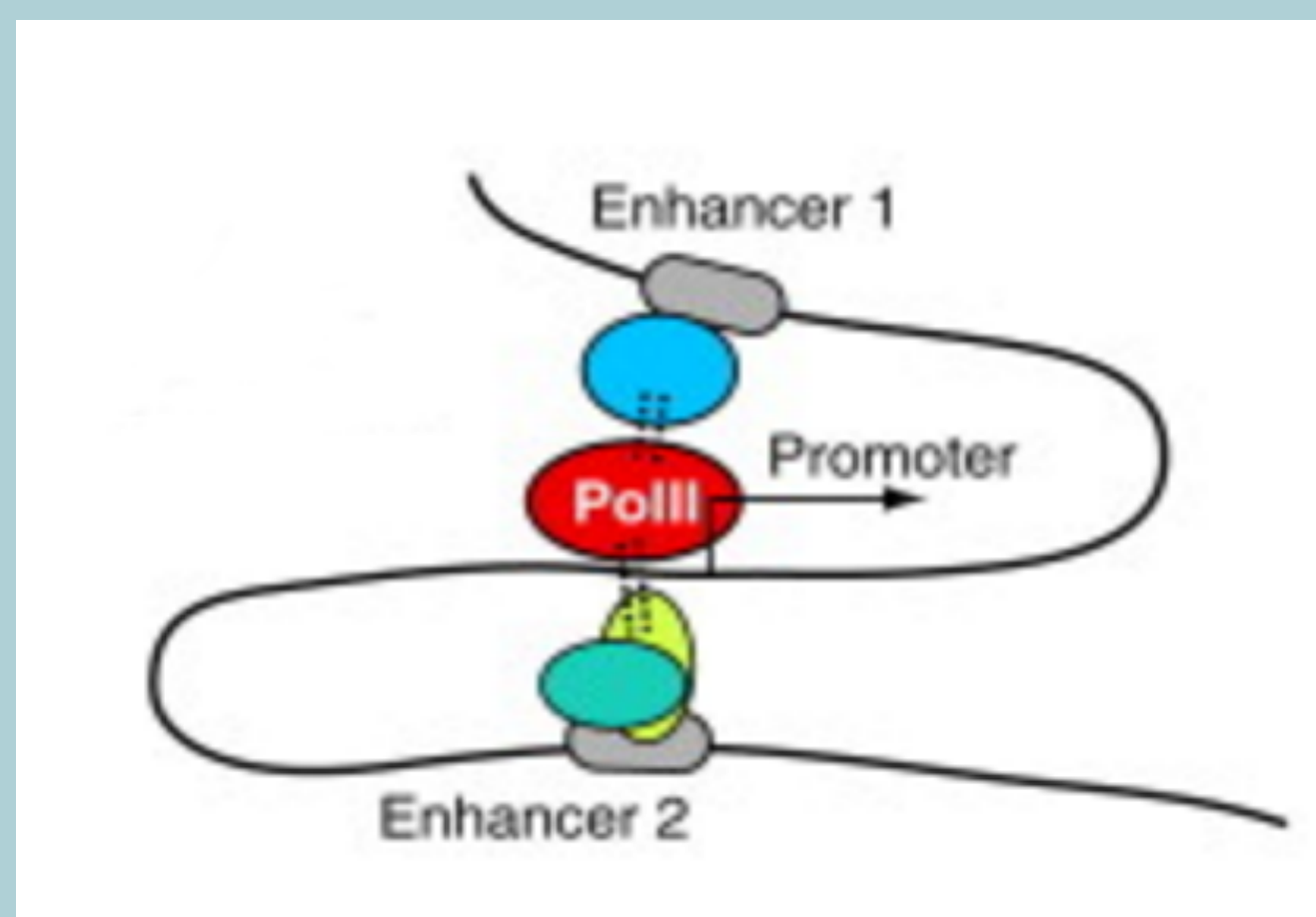


# ALTRE: An Analysis Workflow for Identifying ALTered Regulatory Elements in the Cancer Epigenome Using Chromatin Accessibility Data

Elizabeth Baskin, Rick Farouni, and Ewy Mathé

## Background

- Regulatory elements are regions of non-coding DNA which regulate the transcription of target genes. The two most common types are promoters and enhancers.
- Whereas a promoter initiates transcription and is located near a gene's transcription start site, an enhancer modulates transcription by interacting with a promoter and can be located further away.
- Genes are expressed when chromatin opens up allowing the formation of a DNA loop in which enhancers and promoters interact.



Transcriptional Regulation

- DNA accessibility is regulated by epigenomic signals such as histone modifications, ncRNA expression, DNA methylation, and transcription factors.

## The Cancer Epigenome:

- Cancer cells exhibit an unexpected abundance of mutations both in the genomic sequence of regulatory elements and in epigenetic modifier genes, defined as genes whose products alter the structure of chromatin.
- Differences in the state of chromatin (i.e. open or closed) between cancer cells and normal cells helps us map the aberrant epigenetic landscape of cancer cells.
- Genome-wide chromatin accessibility assays such as DNase-seq, FAIRE-seq, and ATAC-seq map open regulatory regions in different cell types and help decode the regulatory landscape of cancer cells.

## Project's Goal

To develop an R package that implements a user-friendly workflow and an accompanying visual interface (a Shiny web app) for the downstream analysis of chromatin accessibility data (e.g. DNase-seq and ATAC-seq) with the purpose of identifying regulatory elements altered in cancer cells.

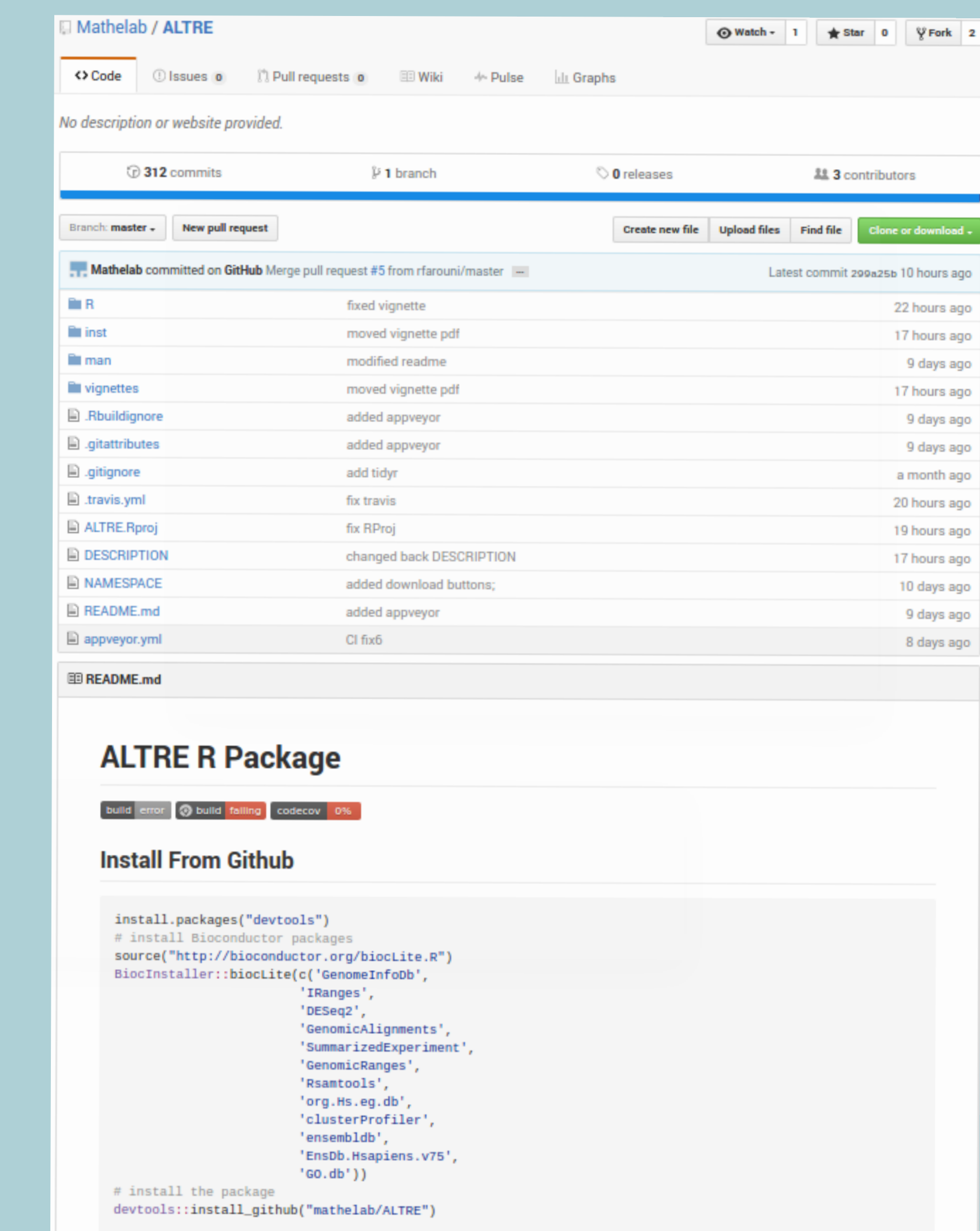
### The Workflow:

ALTRE's analysis workflow consists of eight steps, which can be run either using the R console or the Shiny web app. The pipeline proceeds as follows:

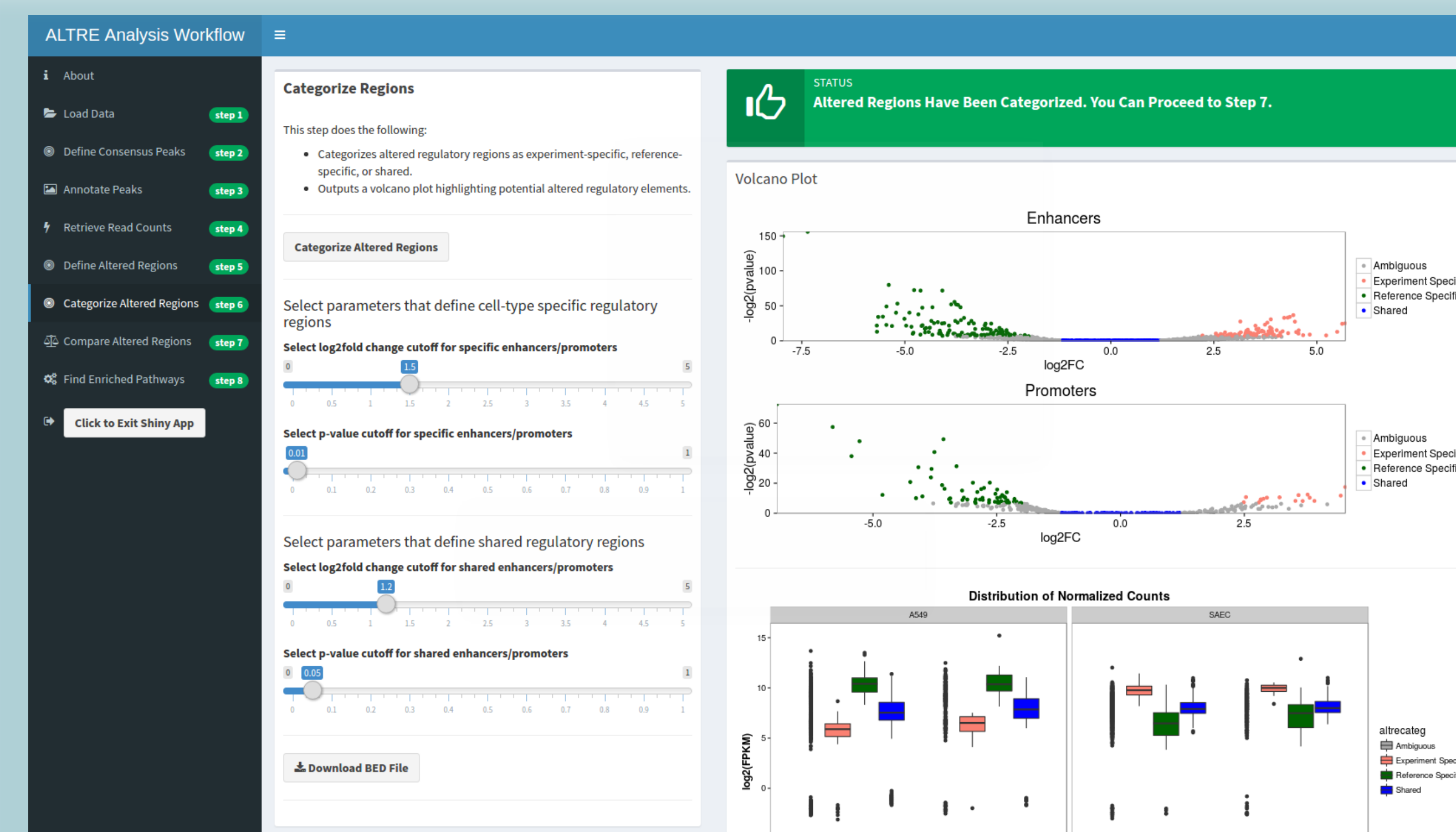
1. Loads the BED and BAM files for sample types of interest
2. Generates consensus peaks across replicates
3. Annotates the peaks as candidate promoters or enhancers
4. Counts the number of reads in each regulatory region of each sample type
5. Identifies which open regulatory regions are differentially altered between cancer cells and normal cells
6. Categorizes the open regulatory regions as cancer cell-specific, normal cell-specific, or shared
7. Compares two methods of identifying open regulatory regions, one based on peak intensity and the other on peak presence
8. Determines which pathways are overrepresented in altered promoters and enhancers as returned by GO Enrichment Analysis

## Package Availability

- ALTRE is an open source R package released under a copyleft GNU GPL v2.0 license.
- A journal paper describing the package has been submitted and is under review<sup>[1]</sup>.
- The package repository is hosted on Github. The R package can be downloaded from <https://github.com/Mathelab/ALTRE>. Instructions on how to install the package are provided on the welcome page.



ALTRE Repository's Github Welcome Page



Step 6 of the Workflow in the Shiny Web App

## References

- [1] Baskin, E., Farouni, R., & Mathe, E. (2016). ALTRE: workflow for defining ALTered Regulatory Elements using chromatin accessibility data. Submitted.

## Acknowledgements:

This software is being developed at Dr. Mathé's lab in the Department of Biomedical Informatics. Dr. Mathé's lab is supported by the Translational Data Analytics Initiative at the Ohio State University.