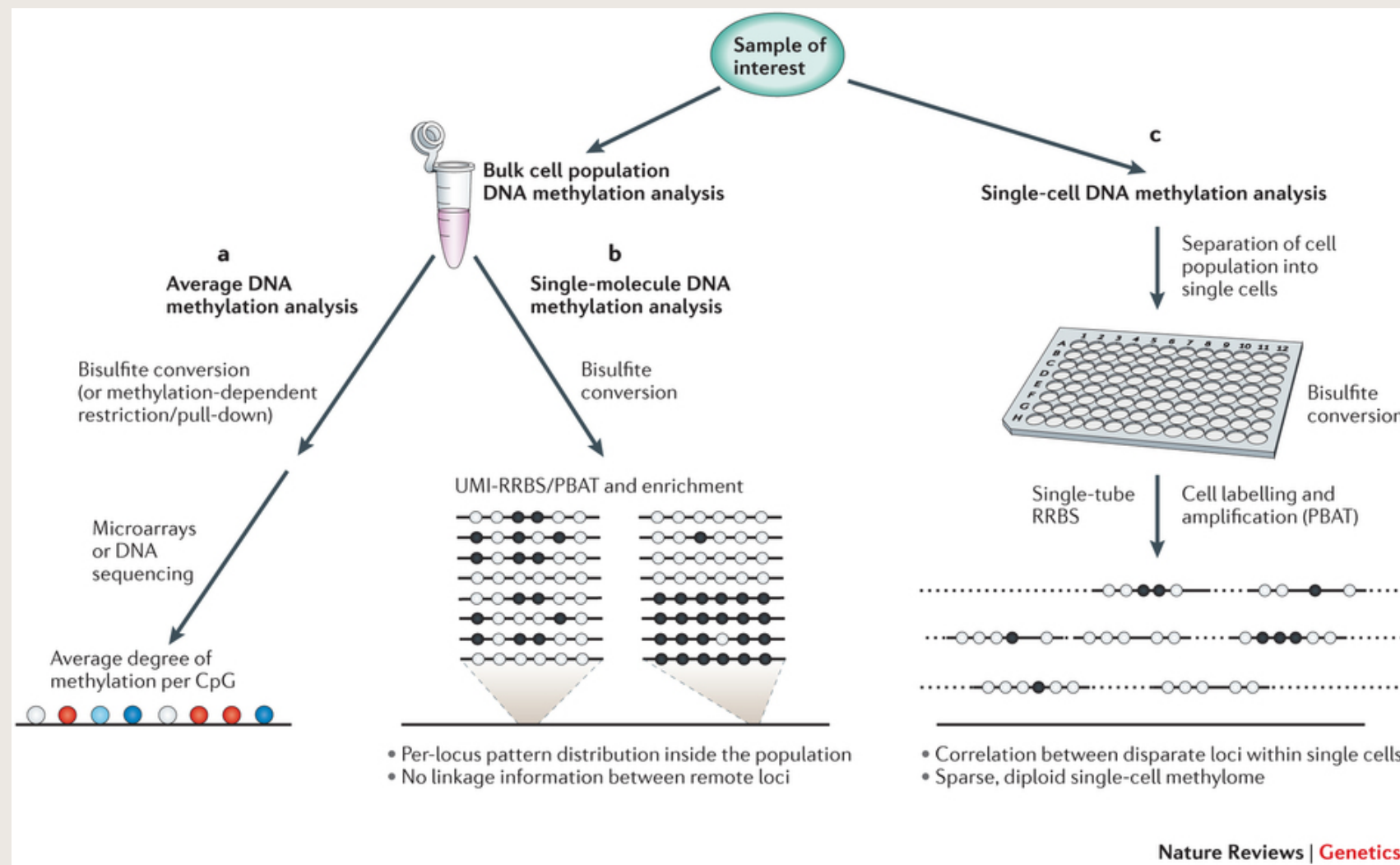


# Screw: tools for building reproducible single-cell epigenomics workflows

Kieran O'Neill, Chelsey Fang, Benjamin Decato, Azhar Khandekar, Alexander Goncareenco, Ben Busby, Aly Karsan  
Genome Sciences Centre, BC Cancer Agency, Vancouver, BC, Canada

## Background: Single-cell DNA Methylation Sequencing

DNA methylation is a heritable epigenetic mark that shows a strong correlation with transcriptional activity. The gold standard for detecting DNA methylation is whole genome bisulfite sequencing (WGBS). Recently, WGBS has been performed successfully on single cells (SC-WGBS) [Schwartzman2015]. The resulting data represents a fundamental shift in the capacity to measure and interpret DNA methylation, especially in rare cell types and contexts where subtle cell-to-cell heterogeneity is crucial, such as in stem cells or cancer.



**DNA Methylation, both bulk and single cell** Schwartzman et al (2015) Nature Reviews in Genetics (used by permission)

## Reproducible Research with CWL and Docker

Reproducible research means completely reproducing a given bioinformatic analysis. This requires having the exact **data**, **code** and **software** that was used.

- Open data in bioinformatics is a fairly solved problem.
- Code is getting there with RMarkdown/Jupyter, but could be better.
- Software versions (and accompanying OS/ecosystem) are a big problem.

CWL aims to provide a uniform and fully reproducible way of representing bioinformatics workflows. Docker aims to provide the exact environment in which an analysis was run. Together, they promise to help bioinformaticians to publish fully reproducible research.

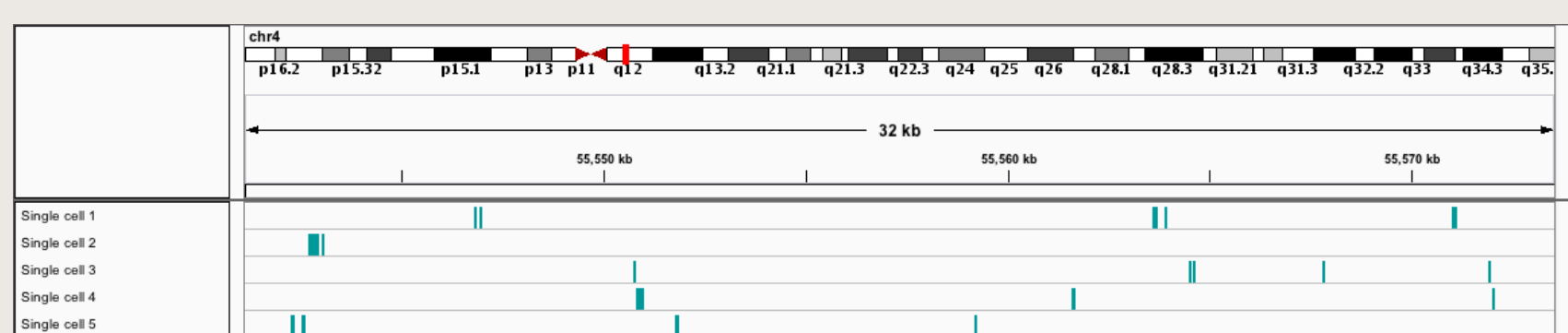
As a side effect, Dockstore enables easy sharing of workflow components to help build new workflows.

## Enter Screw

Screw stands for Single Cell Reproducible Epigenomics Workflow

Screw aims to provide a series of CWL+Dockerised mini-workflows and workflow components for creating fully reproducible single-cell DNA methylation analysis.

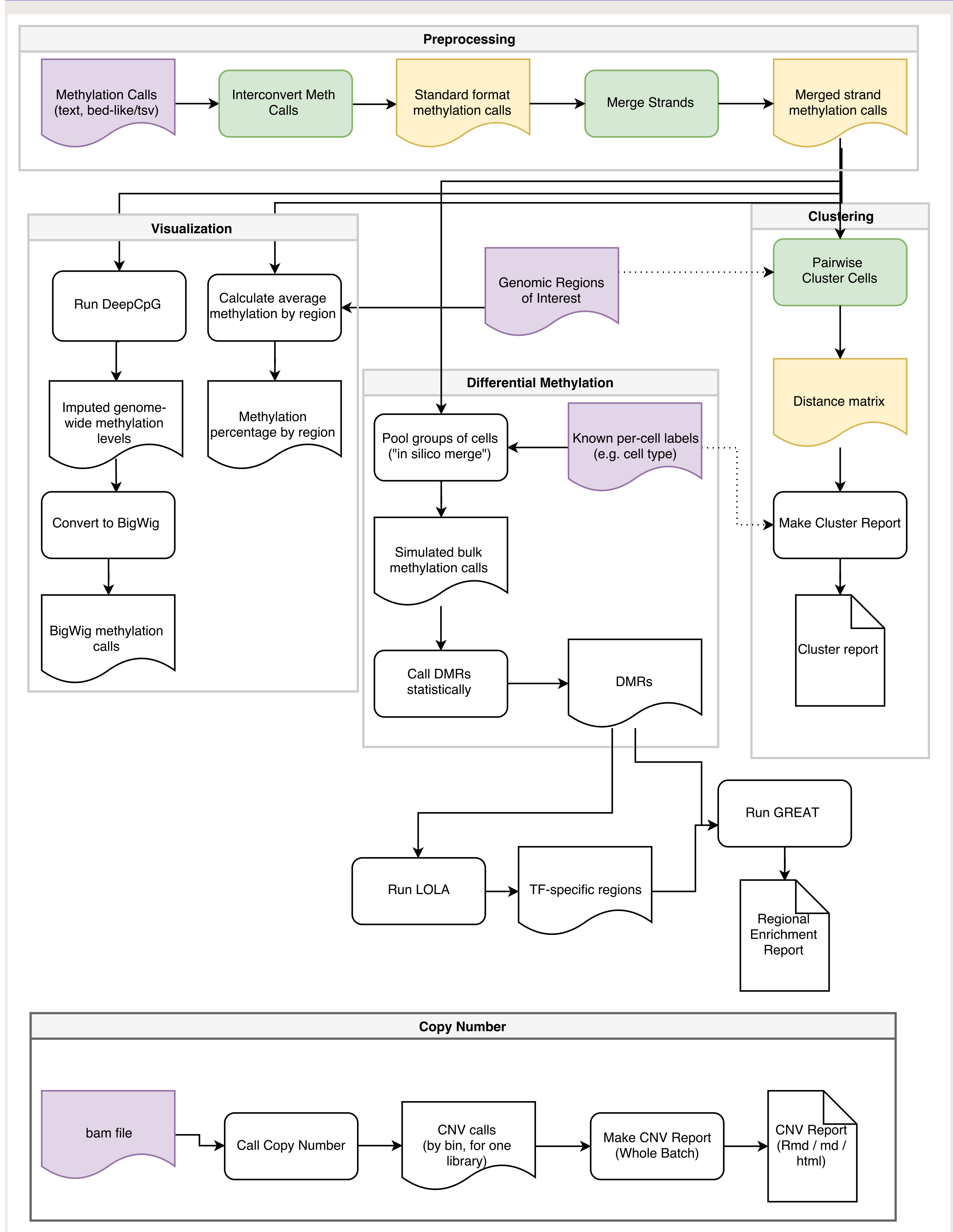
## What SC-WGBS Data Looks Like



## Preprocessing

## Clustering

### Screw Workflow



## CWL Stumbling Blocks

## Future Plans

tSNE

## How did Screw come about?

Screw was initially created during the NCBI Genomics Hackathon in March 2017, organised by Ben Busby.



## Contact/GitHub

**GitHub:** <https://github.com/Epigenomics-Screw>

**Kieran O'Neill:** [koneill@bcgsc.ca](mailto:koneill@bcgsc.ca)