Investigating the effect of environmental toxicants on germ cells via bisulfite sequencing

Medhini Sosale

Allard Laboratory, BIG Summer Program, Institute for Quantitative and Computational Biosciences, UCLA

msosale@gmu.edu

UCLA QCBio

Abstract

Bisphenol A (BPA) and arsenic are hazardous chemicals that cause reproductive dysfunctions over multiple generations. This project applies bisulfite sequencing (BS-seq) to investigate the epigenetic mechanisms underlying their transgenerational inheritance. BS-seq can comprehensively identify the cytosines that have been methylated and how these methylation events can be altered by a treatment. We compared the effectiveness of two BS-seq pipelines, wg-blimp and BSBolt, in detecting differentially methylated regions (DMRs) from untreated, BPA-exposed, and arsenic-exposed *in vitro* mouse primordial germ cells. Specifically, we used whole genome BS-seq (WGBS) to analyze the effects of BPA, and reduced representation BS-seq (RRBS) to analyze the effects of arsenic. While wg-blimp is designed to run a complete WGBS pipeline, BSBolt requires steps to be performed individually. I found that wg-blimp is harder to troubleshoot but that BSBolt is more time consuming. Using BSBolt, I obtained data for FastQC analysis and aligned the WGBS genomes to a reference. I was also able to complete RRBS and began to analyze the results for DMRs.

Background

- **Bisphenol A** (BPA) is an **environmental toxicant** that can cause **reproductive dysfunction**[2]
- **Arsenic** is a naturally-occurring **chemical element** and a highly toxic **carcinogen**
- Exposure to substances such as BPA and arsenic has **epigenetic** effects[3]
- Whole-genome bisulfite sequencing (**WGBS**) and reduced representation bisulfite sequencing (**RRBS**) can be used to analyze these epigenetic effects
- Identify **differentially methylated regions (DMRs)** by looking for cytosines in the genome that have been methylated by a treatment
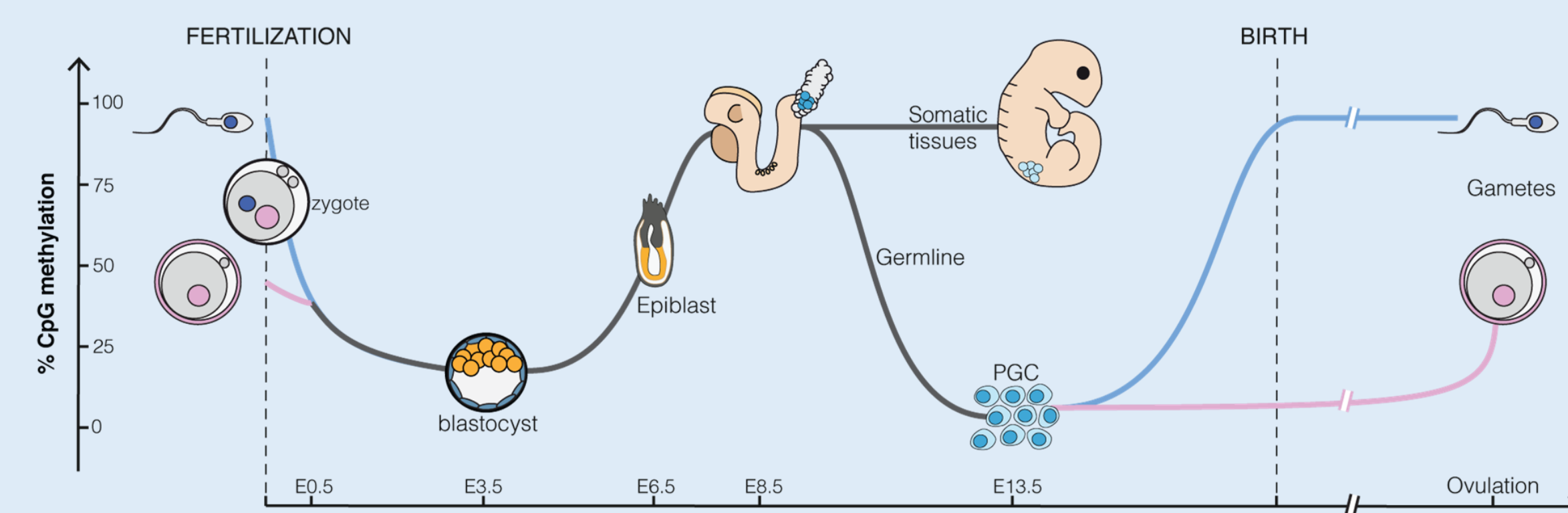- Using data from *in vitro* **mouse primordial germ cells**[2]

Figure 1: Development stages of germ cells (provided by Dr. Patrick Allard)

Methods

Figure 2: Diagram of wg-blimp workflow[4]

- **wg-blimp** is a **WGBS** pipeline[4]
- Completes all steps end-to-end with just **one command**
- Intended to run **faster** and use **less digital memory** than existing WGBS pipelines
- Includes visualization in Integrative Genomics Viewer (**IGV**) and graphing in **R Shiny**

- **BSBolt** can be used for RRBS or WGBS[1]
- Developed at **UCLA**
- Requires steps to be completed **individually** but allows for **more customization**
- Takes **more time** than wg-blimp but is **easier to troubleshoot**

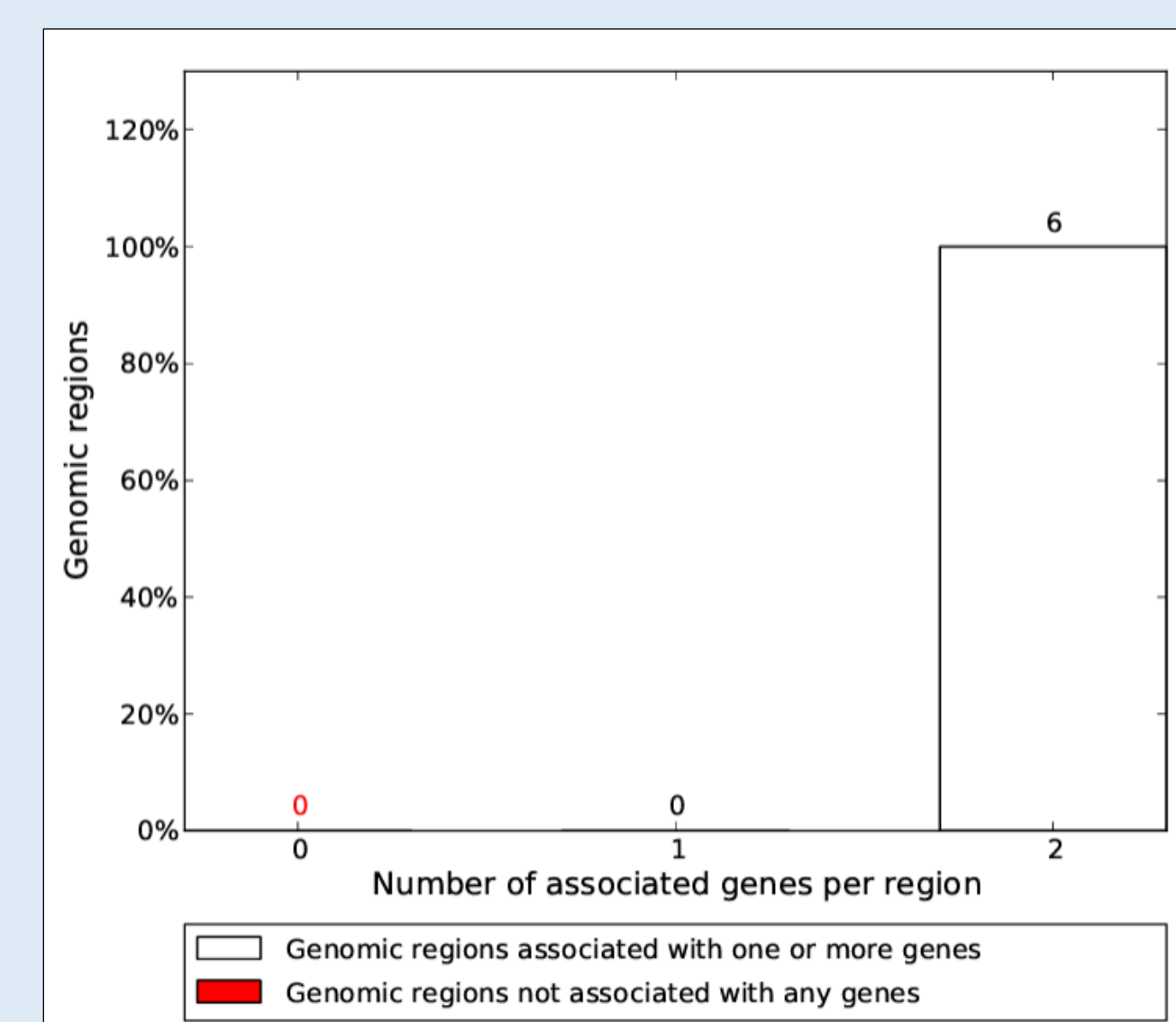Figure 3: Diagram of BSBolt workflow[1]

Results

Figure 4: Associated genes per differentially methylated region

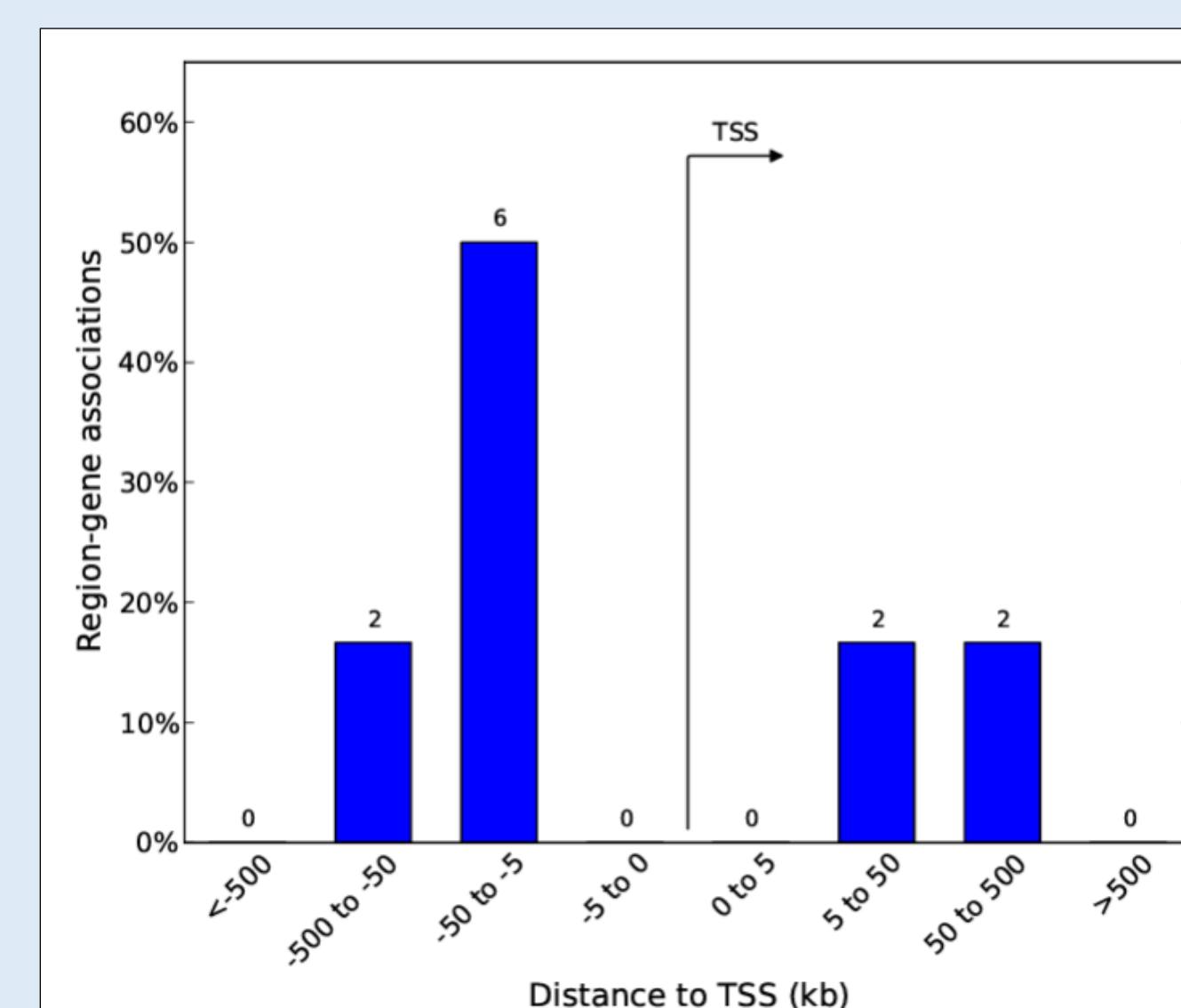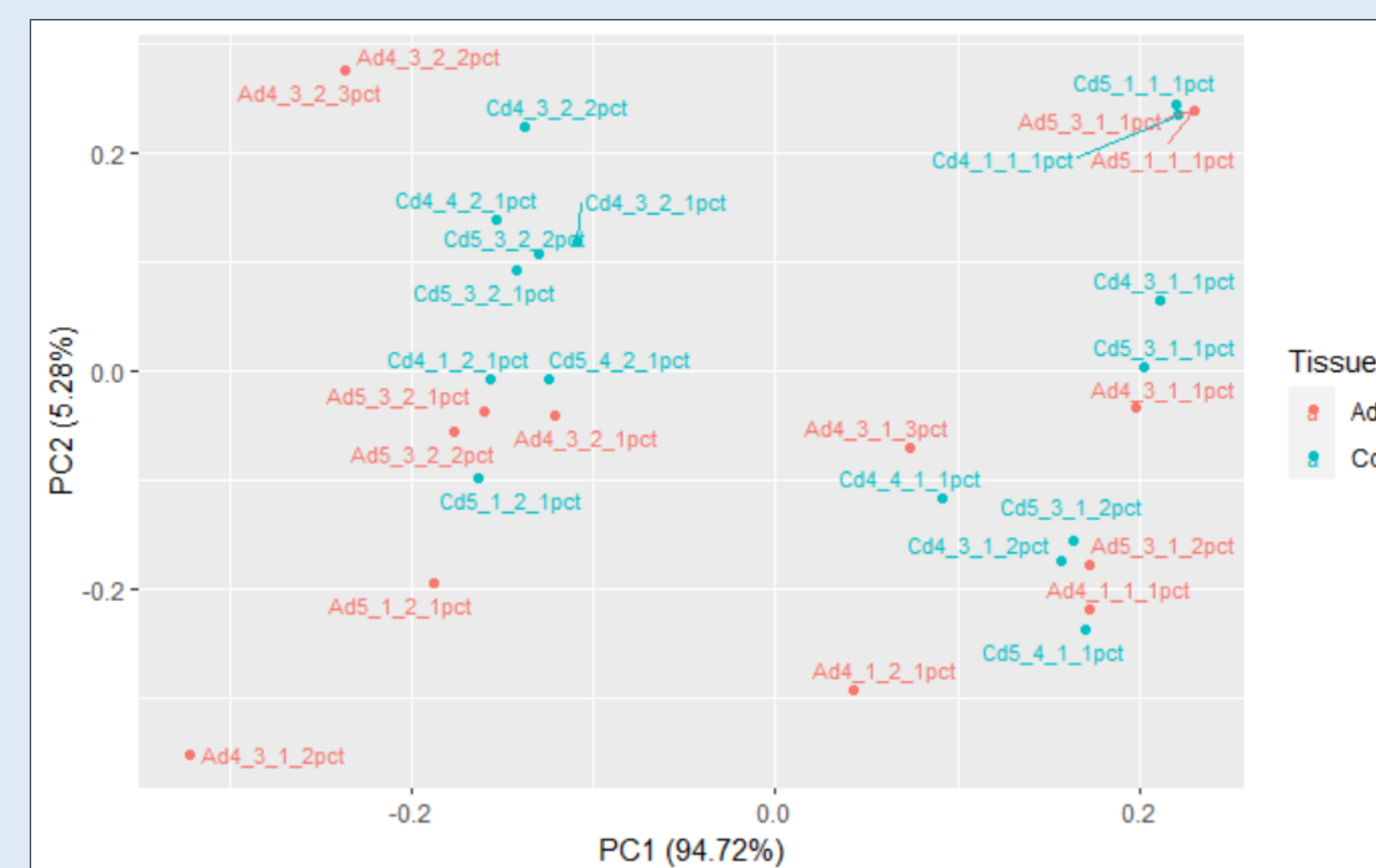Figure 5: Distance of each DMR to transcription start site in the genome

Figure 6: PCA analysis of RRBS data

Discussion & Next Steps

- Very **few DMRs** observed in both datasets
  - **WGBS**: Observed in a previous trial
  - **RRBS**: Observed in chromosomes 10, 11, 15, 2, 7, 9
- **BSBolt** was the more effective pipeline
- New roadblock: **Variation** in data due to **segmentation**
- Need to **fuse** files for each sample together and **rerun**
- Can look at **differentially methylated cytosines** (DMCs) instead of DMRs
- Can begin to investigate the effects of other harmful compounds, such as **nicotine**

References

1. Farrell, C., Thompson, M., Tosevska, A., Oyetunde, A., & Pellegrini, M. (2021). BiSulfite Bolt: A bisulfite sequencing analysis platform. *GigaScience*, *10*(5), giab033. https://doi.org/10.1093/gigascience/giab033
2. Ooi, S. K. T., Jiang, H., Kang, Y., & Allard, P. (n.d.). Examining the Developmental Trajectory of an in Vitro Model of Mouse Primordial Germ Cells following Exposure to Environmentally Relevant Bisphenol A Levels. *Environmental Health Perspectives*, *129*(9), 097013. https://doi.org/10.1289/EHP8196
3. Pascual, F. (n.d.). Off to a Rough Start: Environmental Exposures May Alter Germ Cell Development. *Environmental Health Perspectives*, *130*(1), 014001. https://doi.org/10.1289/EHP10550
4. Wöste, M., Leitão, E., Laurentino, S., Horsthemke, B., Rahmann, S., & Schröder, C. (2020). wg-blimp: An end-to-end analysis pipeline for whole genome bisulfite sequencing data. *BMC Bioinformatics*, *21*(1), 169. https://doi.org/10.1186/s12859-020-3470-5
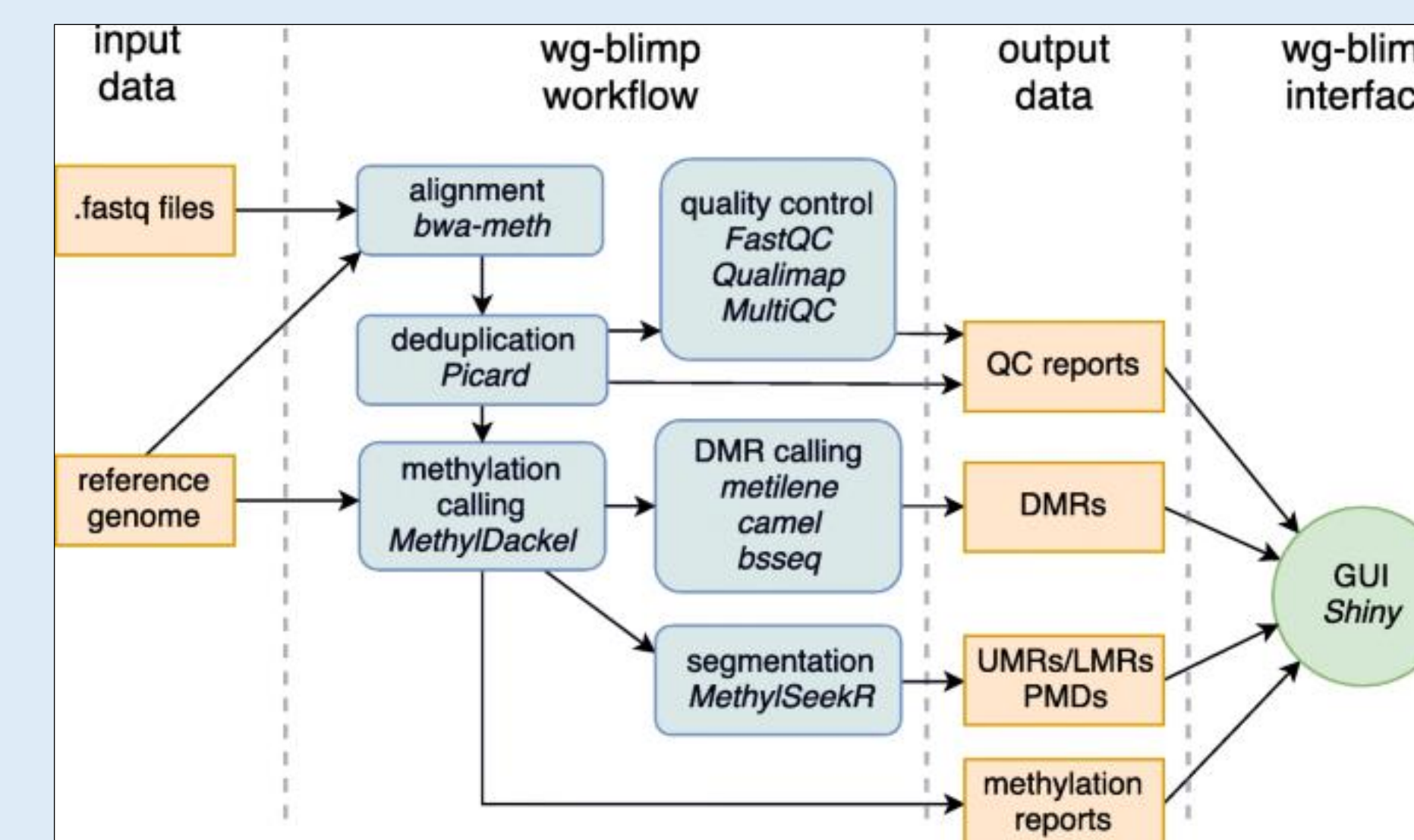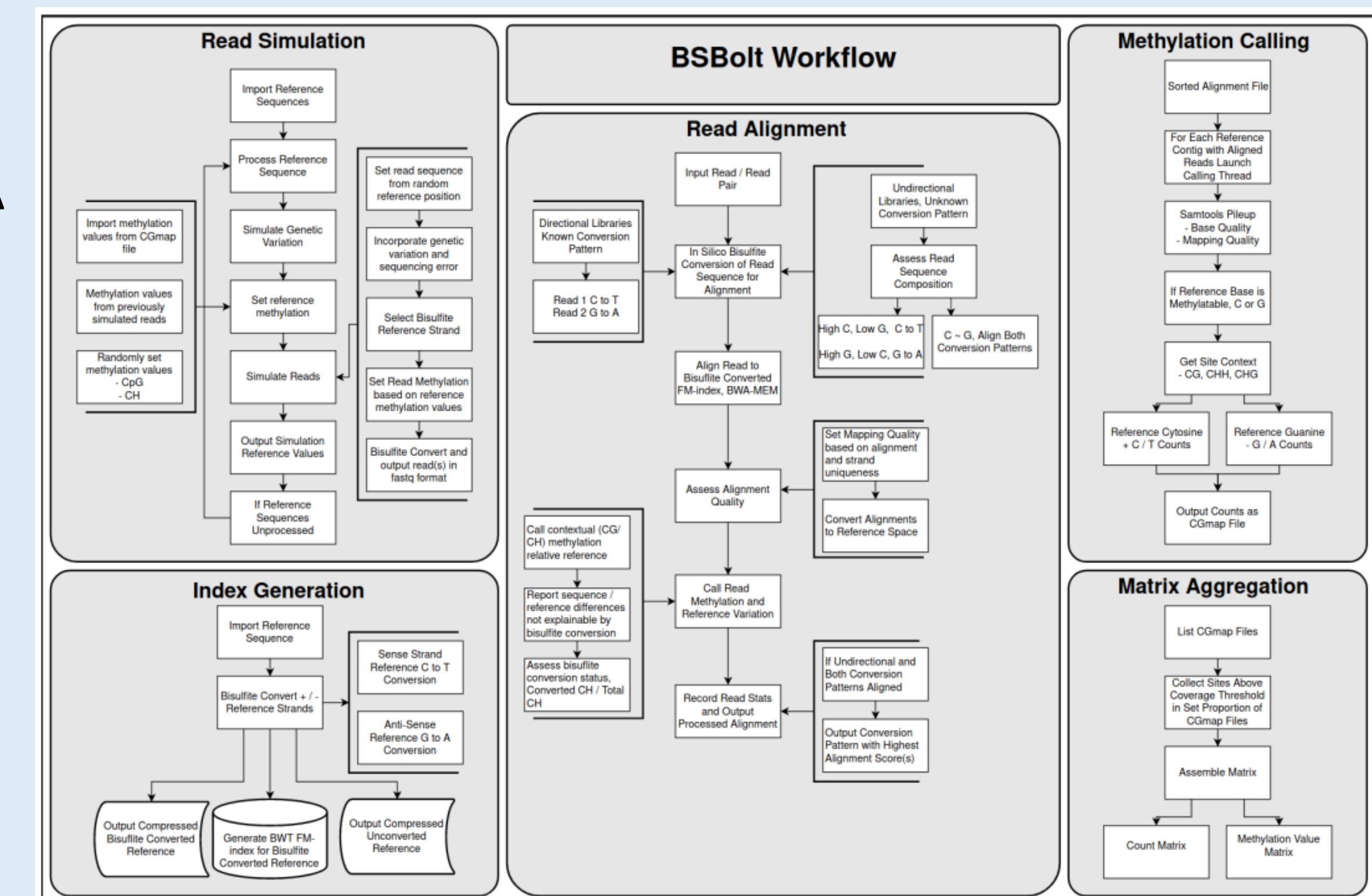
Acknowledgments