

Whole-Genome Bisulfite Sequencing

Lee Carlin

February 12, 2018

1 Literature

1.1 Methodology

- (Jaffe et al., 2012)
 - Bump hunting to identify differentially methylated regions in epigenetic epidemiology studies
- (Akalın et al., 2012)
 - Analysis of genome-wide DNA methylation profiles
- (Hansen, Langmead, & Irizarry, 2012)
 - Hansen presents an end-to-end WGBS analysis pipeline that includes alignment, quality control, inference, and DMRs production tools for low coverage data that can handles biological differences.
- (Michels et al., 2013)
 - Recommendations and suggestions for Epigenome-wide association study (EWAS)
- (Assenov et al., 2014)
 - A software tool for large-scale analysis and interpretation of DNA methylation data, providing a user-friendly analysis work-flow that yields detailed hypertext reports
- (Feng, Conneely, & Wu, 2014)
 - Although developed for a single nucleotide resolution, these authors suggests a Bayesian hierarchical model to detect differentially methylated loci from single nucleotide resolution sequencing data
- (Dolzhenko & Smith, 2014)
 - The use of beta-binomial regression as a general approach for modeling whole-genome bisulfite data to identify differentially methylated sites and genomic intervals.
- (Sun et al., 2014)
 - MOABS: model based analysis of bisulfite sequencing data
- (Saito, Tsuji, & Mituyama, 2014)
 - Accurate detection of methylated cytosines and differentially methylated regions
- (Park, Figueroa, Rozek, & Sartor, 2014)
 - MethylSig takes into account both read coverage and biological variation by utilizing a beta-binomial approach across biological samples for a CpG site or region, and identifies relevant differences in CpG methylation. It can also incorporate local information to improve group methylation level and/or variance estimation for experiments with small sample size.
- (Ziller, Hansen, Meissner, & Aryee, 2015)
 - Coverage recommendations for methylation analysis by whole genome bisulfite sequencing
- (Wu et al., 2015)

- Detection of differentially methylated regions from whole-genome bisulfite sequencing data without replicates
- (Lee & Morris, 2015)
 - Identification of differentially methylated loci using wavelet-based functional mixed models
- (Yu & Sun, 2016)
 - identifying differentially methylated regions using a hidden Markov model (no pdf)
- (Benjamini, Taylor, & Irizarry, 2016)
 - Selection Corrected Statistical Inference for Region Detection with High-throughput Assays
- (Wen, Chen, Zhang, Zhuang, & Li, 2016)
 - Detection of differentially methylated regions in whole genome bisulfite sequencing data using local Getis-Ord statistics
- (Park & Wu, 2016)
 - Differential methylation analysis for BS-seq data under general experimental design
- (Korthauer, Chakraborty, Benjamini, & Irizarry, 2017)
 - Fitting a generalized least squares (GLS) regression model with a nested autoregressive correlated error structure for the effect of interest on transformed methylation proportions and inference based on a pooled null distribution.
- (Chen, Pal, Visvader, & Smyth, 2017)
 - This package is to be used with RRBS, but it contains interesting techniques.
- (Shafi, Mitrea, Nguyen, & Draghici, 2017)
 - A survey of the approaches for identifying differential methylation using bisulfite sequencing data

1.2 Practice

- (Deaton & Bird, 2011)
 - CpG islands and the regulation of transcription
- (Hansen et al., 2011)
 - Increased methylation variation in epigenetic domains across cancer types
- (Khamis et al., 2016)
 - CpG traffic lights are markers of regulatory regions in humans

Begin by looking at s in leading journals that use BS-seq: Cell, Nature Bioinformatics. How was BS-seq used?

References

- Akalın, A., Kormaksson, M., Li, S., Garrett-Bakelman, F. E., Figueroa, M. E., Melnick, A., & Mason, C. E. (2012). methylkit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome biology*, 13(10), R87.
- Assenov, Y., Müller, F., Lutsik, P., Walter, J., Lengauer, T., & Bock, C. (2014). Comprehensive analysis of DNA methylation data with RnBeads. *Nature methods*, 11(11), 1138.
- Benjamini, Y., Taylor, J., & Irizarry, R. A. (2016). Selection corrected statistical inference for region detection with high-throughput assays. *bioRxiv*, 082321.
- Chen, Y., Pal, B., Visvader, J. E., & Smyth, G. K. (2017). bisulfite sequencing experiments using edgeR [version 1].
- Deaton, A. M., & Bird, A. (2011). CpG islands and the regulation of transcription. *Genes & development*, 25(10), 1010–1022.
- Dolzhenko, E., & Smith, A. D. (2014). Using beta-binomial regression for high-precision differential methylation analysis in multifactor whole-genome bisulfite sequencing experiments. *BMC bioinformatics*, 15(1), 215.
- Feng, H., Conneely, K. N., & Wu, H. (2014). A bayesian hierarchical model to detect differentially methylated loci from single nucleotide resolution sequencing data. *Nucleic acids research*, 42(8), e69–e69.
- Hansen, K. D., Langmead, B., & Irizarry, R. A. (2012). BSmooth: from whole genome bisulfite sequencing reads to differentially methylated regions. *Genome biology*, 13(10), R83.
- Hansen, K. D., Timp, W., Bravo, H. C., Sabuncian, S., Langmead, B., McDonald, O. G., ... others (2011). Increased methylation variation in epigenetic domains across cancer types. *Nature genetics*, 43(8), 768.
- Jaffe, A. E., Murakami, P., Lee, H., Leek, J. T., Fallin, M. D., Feinberg, A. P., & Irizarry, R. A. (2012). Bump hunting to identify differentially methylated regions in epigenetic epidemiology studies. *International journal of epidemiology*, 41(1), 200–209.
- Khamis, A. M., Lioznova, A. V., Artemov, A. V., Ramensky, V., Bajic, V. B., & Medvedeva, Y. A. (2016). CpG traffic lights are markers of regulatory regions in humans. *bioRxiv*, 095968.
- Korthauer, K., Chakraborty, S., Benjamini, Y., & Irizarry, R. A. (2017). Detection and accurate false discovery rate control of differentially methylated regions from whole genome bisulfite sequencing. *bioRxiv*, 183210.
- Lee, W., & Morris, J. S. (2015). Identification of differentially methylated loci using wavelet-based functional mixed models. *Bioinformatics*, 32(5), 664–672.
- Michels, K. B., Binder, A. M., Dedeurwaerder, S., Epstein, C. B., Grealis, J. M., Gut, I., ... others (2013). Recommendations for the design and analysis of epigenome-wide association studies. *Nature methods*, 10(10), 949.
- Park, Y., Figueroa, M. E., Rozek, L. S., & Sartor, M. A. (2014). MethylSig: a whole genome DNA methylation analysis pipeline. *Bioinformatics*, 30(17), 2414–2422.
- Park, Y., & Wu, H. (2016). Differential methylation analysis for BS-seq data under general experimental design. *Bioinformatics*, 32(10), 1446–1453.
- Saito, Y., Tsuji, J., & Mituyama, T. (2014). Bisulfighter: accurate detection of methylated cytosines and differentially methylated regions. *Nucleic acids research*, 42(6), e45–e45.
- Shafi, A., Mitrea, C., Nguyen, T., & Draghici, S. (2017). A survey of the approaches for identifying differential methylation using bisulfite sequencing data. *Briefings in bioinformatics*, bbx013.
- Sun, D., Xi, Y., Rodriguez, B., Park, H. J., Tong, P., Meong, M., ... Li, W. (2014). Moabs: model based analysis of bisulfite sequencing data. *Genome biology*, 15(2), R38.
- Wen, Y., Chen, F., Zhang, Q., Zhuang, Y., & Li, Z. (2016). Detection of differentially methylated regions in whole genome bisulfite sequencing data using local getis-ord statistics. *Bioinformatics*, 32(22), 3396–3404.
- Wu, H., Xu, T., Feng, H., Chen, L., Li, B., Yao, B., ... Conneely, K. N. (2015). Detection of differentially methylated regions from whole-genome bisulfite sequencing data without replicates. *Nucleic acids research*, 43(21), e141–e141.
- Yu, X., & Sun, S. (2016). HMM-DM: identifying differentially methylated regions using a hidden Markov model. *Statistical applications in genetics and molecular biology*, 15(1), 69–81.

Ziller, M. J., Hansen, K. D., Meissner, A., & Aryee, M. J. (2015). Coverage recommendations for methylation analysis by whole-genome bisulfite sequencing. *Nature methods*, 12(3), 230.