Ecological Genomics: Homework #3

Brendan Case

https://github.com/brendandaisy/ecological-genomics

10 April 2020

1 Background

The copepod *Acartia tonsa* inhabits estuaries and coastal waters across the world, tolerating a number of different environmental stressors. Often appearing with high abundance, these resilient zooplankton are an important base to food chains and other ecosystem functions. In the turbulent times of the anthropocene, understanding the mechanisms facilitating rapid adaptation is paramount. While it is known that DNA methylation has a potential role in adaptation, the relative importance of epigenenic factors in stress response is still largely unexplored [2, 4]. Here we look for differences in CpG methylation levels between experimental *A. tonsa* populations under heat and/or CO_2 stress.

2 Bioinformatics Pipeline

Wild *A. tonsa* samples were raised in common garden conditions for 3 generations prior to being split into the 4 treatments $\times$ 4 replicates with $\sim 4,000$ individuals each. Treatments were control (AA), CO_2 stress (AH), heat stress (HA), and combined stress (HH). Individuals were then let to evolve in these conditions, and sequencing was performed at both the F0 and F25 generations. For this study, in the F0 generation we chose only to use the AA populations. In order to quantify methylation frequencies in each group, we used reduced representation bisulfite sequencing. Read quality before and after trimming was performed with `fastqc`. In addition, samples were spiked with unmethylated *E. coli* DNA to test bisulfite conversion efficiency. Alignment was done using `bismark` with a reference *A. tonsa* genome. Samples with expectionally low or high (read depth < 20, or $>$ 97.5th percentile) coverage were removed, then percent methylation was calculated for each site. Then, differential methylation was evaluated between the F0 and F25 groups, along with the AA_F25 and AH_F25 groups using hierarchical clustering and χ^2 tests using the `methylKit` package in R [1]. Finally, sites with significant methylation differences between the AA_F25 and AH_F25 groups were annotated using `bedtools`.

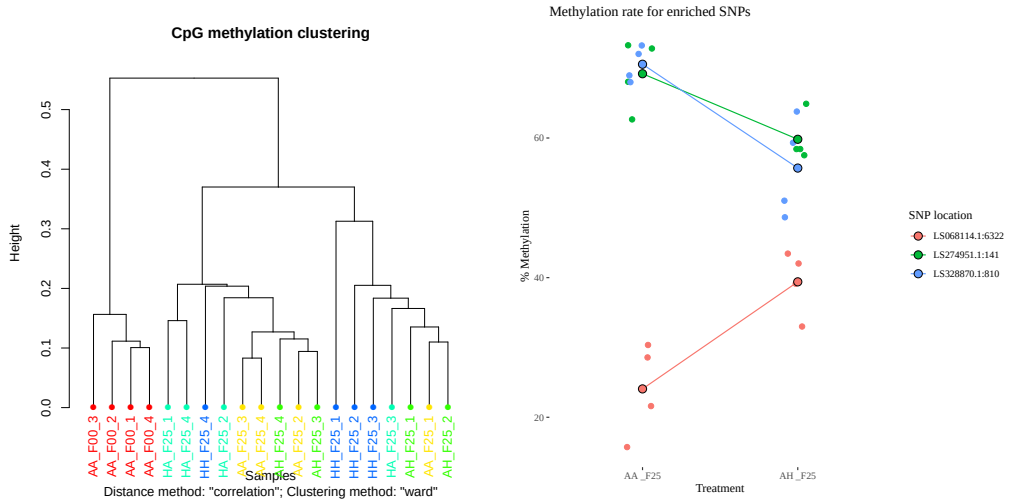

Figure 1: Left: hierarchical clustering of methylated percentage in all replicate groups. Right: percent methylation in the 3 sites which showed significant differences between the AA_F25 and AH_F25 groups, and were functionally enriched. Larger points connected by lines show means for each group, while lines show degree of hyper/hypo-methylation.

3 Results

After filtering, coverage was between 30 and 70 in all groups, with the highest coverage occurring in the F0 samples. Average percent methylation was highest in the AA_F0 group (%41.3), and lowest in the AH_F25 group (%38.0). Hierarchical clustering between replicates revealed large-scale differences in methylation percentages between the F0 generation and the F25 populations (Fig. 1, left). Further, 3 out of the 4 AA_F25 and HH_F25 groups were distinct from each other, falling into separate "second level" branches in the clustering tree. Between the AA_F25 and AH_F25 groups, the χ^2 test found 30 sites which were differently methylated (6 hyper-methylated in AH_F25, 24 hypo-methylated). Of these sites, 3 were functionally enriched, with matching GO terms for DNA integration, dUTP catabolic processes, and various aspects of endonuclease activity. The two of these SNPs associated with endonuclease activity were hypo-methylated with similar methylation levels, while the dUTP-associated SNP was hyper-regulated (Fig. 1, right).

4 Conclusion

Our results show CpG methylation is present in *A. tonsa* under a variety of environmental conditions. The largest differences in methylation was found between F0 and F25 individuals, with methylation decreasing in the F25 population. While this could suggest a decrease in methylation associated with adaptation to stress [3], it is also could be an artifact of our sequencing and/or mapping methods. We also found significantly different methylation levels in several locations between AA_F25 and AH_F25 individuals, which

shows an association between methylation and prolonged CO_2 exposure. Two of these sites were associated with endonuclease activity. While it was outside the scope of these study to investigate functional roles in these genes and their association to methylation in response to stress, this is an avenue for future work.

References

[1] A. Akalin, M. Kormaksson, S. Li, F. E. Garrett-bakelman, M. E. Figueroa, A. Melnick, and C. E. Mason. methylKit : a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biology*, 13(R87), 2012.

[2] A. Boyko, T. Blevins, Y. Yao, A. Golubov, A. Bilichak, Y. Ilnytskyy, F. Meins, and I. Kovalchuk. Transgenerational Adaptation of Arabidopsis to Stress Requires DNA Methylation and the Function of Dicer-Like Proteins. *PLoS ONE*, 5(3), 2010.

[3] D.-h. Cao, X. Gao, J. Liu, J. N. Kimatu, S.-j. Geng, J. Zhao, and D.-c. Shi. Methylation sensitive amplified polymorphism (MSAP) reveals that alkali stress triggers more DNA hypomethylation levels in cotton (Gossypium hirsutum L .) roots than salt stress. *African Journal of Biotechnology*, 10(82), 2011.

[4] K. B. Flores, F. Wolschin, and G. V. Amdam. Integrative and Comparative Biology The Role of Methylation of DNA in Environmental Adaptation. *Integrative and Comparative Biology*, 53(2):359–372, 2013.