

Supplemental Information

Methods

Plant material

Processed methylomes of MA lines MA3_WT (GSE153055), MA_cmt3 (GSE178684) and MA_suvh4/5/6 (GSE178684) were taken from [1]. H2AZ data (GSE128434) were taken from [2].

Metaprofile creation

The genomic data of the MA-lines were partitioned into windows according to the position of a CG-site within its gene; only gbM genes were included. A single window spans 5% of the gene length, with overlapping windows starting every 1%. Sites up to 2kB up- and downstream of the gene were also included. The metaprofile tool is available as an open-source application (<https://github.com/constantingoeldel/alphabeta-rs/>).

Epimutation rate estimation

The CpG divergences of Fig. 1b were obtained using AlphaBeta [35] on both the full methylome and a filtered methylome containing only CG-sites within gbM genes. 95% confidence intervals were calculated with a two-sided inverse Student's t-distribution from the linear regression standard errors. A hypothesis test to demonstrate a significant difference between the regression coefficients [3] resulted in $p < 0.001$.

The epimutation rates of the metaprofiles were estimated using the ABneutral model of AlphaBeta [35], which, for performance reasons, was rewritten in Rust. It is also available as an open-source application through alphabeta-rs

(<https://github.com/constantingoeldel/alphabeta-rs/>). 95% confidence intervals were obtained through bootstrapping. We tested for significant differences in estimated epimutation rates between genotypes (Fig. S1), using a modified Welch's t-test where the sample size of the i th genotype is given by $(n - 1) \frac{n_i}{2}$, with n being the number of samples in the pedigree.

H2AZ analysis

To estimate the correlation between a windows' epimutation rate and its H2AZ content, we assorted the H2AZ sites into the same sliding windows along the gene, followed by a linear regression.

CG-density analysis

To estimate the correlation between a windows' epimutation rate and its CG-Density, we calculated the CG-Density of each sliding window, followed by a linear regression with the epimutation rates of each window.

Replicability

Metaprofile creation, statistical analysis and figure creation code is available as an interactive notebook.

<https://github.com/constantingoeldel/epigenomics/blob/master/plotting.ipynb>). Raw data and intermediary results are available as a spreadsheet (<https://docs.google.com/spreadsheets/d/1eldupjcWapVGPNMNTeKbnaPIroEXP6taFhrDMn5z3D4/edit?usp=sharing>).

Sources:

1. Hazarika, R. R., Serra, M., Zhang, Z., Zhang, Y., Schmitz, R. J., & Johannes, F. (2022). Molecular properties of epimutation hotspots. *Nature plants*, 8(2), 146–156. <https://doi.org/10.1038/s41477-021-01086-7>
2. Lu, Z., Marand, A.P., Ricci, W.A. et al. The prevalence, evolution and chromatin signatures of plant regulatory elements. *Nat. Plants* 5, 1250–1259 (2019). <https://doi.org/10.1038/s41477-019-0548-z>
3. Clogg, C. C., Petkova, E., & Haritou, A. (1995). Statistical methods for comparing regression coefficients between models. *American Journal of Sociology*, 100(5), 1261–1293. <https://doi.org/10.1086/230638>