**Supplemental Data for “Uncovering methylation-dependent genetic effects on regulatory element function in diverse genomes”**

Rachel M. Petersen1,2, Christopher M. Vockely3, Amanda J. Lea1,2,4,\*

1 Department of Biological Sciences, Vanderbilt University, Nashville, TN

2 Evolutionary Studies Initiative, Vanderbilt University, Nashville, TN

3 The Broad Institute of MIT and Harvard, Cambridge, MA

4 Vanderbilt Genetics Institute, Vanderbilt University, Nashville, TN

1) **replicateinfo.txt** contains metadata for each mSTARR-seq replicate, including replicate number, pool number, sample type (DNA vs RNA) and methylation status

2) **rnadnacounts\_400bpwin.txt** contains a count matrix with the number of DNA and RNA reads falling within each 400 bp genomic window for each replicate. Columns are replicate names, rows are genomic windows.

3) **model\_results.zip** contains

* model1\_methonly\_results.txt: results from linear modeling to identify windows with regulatory function in the methylated condition
* model1\_unmethonly\_results.txt: results from linear modeling to identify windows with regulatory function in the methylated condition
* model2\_mashr\_results.txt: results from mashr analysis to identify windows with methylation-dependent regulatory function
* ASE\_meth\_results.txt: results from allele specific expression analysis to identify ASE in the methylated condition
* ASE\_unmeth\_results.txt: results from allele specific expression analysis to identify ASE in the unmethylated condition
* ASE\_mashr\_results.txt: results from mashr analysis to identify sites with methylation-dependent ASE

4) **Comparison\_datasets.zip** contains

* Johnston\_eLife\_mSTARR\_counts\_K562.txt: counts matrix from Johnston et al. 2024, adapted to use 200 bp windows. Original dataset can be found here: https://zenodo.org/records/7949036#.ZGZ5UnbMJq9
* Lea\_eLife\_mSTARR\_counts.txt: counts matrix from Lea et al. 2018

5) **Joint\_genotyping.vcf**contains results from joint genotyping analysis using DNA sequences generated in the current study from 25 individuals accessed through the 1000 Genomes Project.

6) **GWAS\_EWAS\_overlap\_files.zip** contains

* GWAShits\_siteformat.txt: GWAS associations accessed through the NHGRI-EBI catalog in March 2024, formatted for use in R
* EWAS\_Atlas\_associations.tsv: EWAS associations accessed through the EWAS Open Platform Data Hub in March 2024
* EWAS\_Atlas\_probe\_annotations.tsv: genomic locations of EWAS probes
* ASE\_mashr\_GWASOverlap.bed: methylation-dependent genetic effect sites that are located within 400 bp of a GWAS hit (results of bedtools intersect)
* ASE\_mashr\_EWASOverlap.bed: methylation-dependent genetic effect sites that are located within 400 bp of a EWAS hit (results of bedtools intersect)