**IntroPainter**

This is a home made pipeline. The description of the method to calculate introgression regions is as follows:

The variant sites where the missing rate in each of the ancestor populations less than 20% were retained. The sites were further filtered out where the allele frequency between two ancestral populations was <0.8. This criterion allowed us to only retain the diagnostic SNPs and assign alleles of each parental population. After filtering, we calculated the ratio of every remaining site in CN\_Cs to which ancestors it belonged. If the site in the tested individual was missing, it was also removed. Then, we calculated the average ratios of the parental alleles with sliding windows of window size at 10,000 SNPs and step size at 5,000 SNPs. This method was applied to every individual and the whole population.

Please cite our paper if you use this pipeline:

Chien, C.C., Seiko, T., Muto, C., Ariga, H., Wang, Y.C., Chang, C.H., Sakai, H., Naito, K.\* and Lee, C.R.\* (2025) A single domestication origin of adzuki bean in Japan and the evolution of domestication gene. Science, 388, eads2871. DOI: 10.1126/science.ads2871

Dependencies: bcftools, conda environment for R, R packages: ggplot2, dplyr

Usage:

Usage: perl IntroPainter\_v2.pl -vcf PATH -list LIST\_FILE [-win INT] [-step INT] [-pre PREFIX] [-p1c COLOR] [-p2c COLOR] [-ci] [-rcal] [-rplot] [-conda NAME] [-sn SN] [-m FLOAT] [-d FLOAT] [-n] [-mem INT] [-local] [-ow] [-exc]

-list: trios list, population\_name follows by sample\_names. 3 lines: 2 parents at first two lines, and the testing population in the 3rd line.

-win: window size (SNP number). Default: 1000

-step: step size (SNP number). Default: 500

-p1c: represent color of ancestor 1 in the figure. If you use color code, please add ''. Default: blue

-p2c: represent color of ancestor 2 in the figure. If you use color code, please add ''. Default: red

-ci: shown 95% confidence interval or not. Default: False

-m: missing rate of the sites. Default: 0.8

-d: allele differential rate of the sites between two parents. Default: 0.8

-rcal: re-calcutate introgression ratios only.

-rplot: re-plot introgression figures only.

-conda: conda environment name. Default: R-4.1

-n: threads to use. Default: 1

-mem: memory to use. Default: 12Gb

-ow: over-write the outputs.

-sn: serial number for the run. Required if -rcal, -rplot, or -ow is used.

-local: run the pipeline in the local machine. (Warning: It might be very slow)

-exc: send jobs to execute.