

Integrating population information using DNA methylation to explore the spatiotemporal dynamics and developmental mechanisms of populations

Wei Liu¹, Chu Jiang¹, Zheng-Feng Wang¹, Jie Chen², Hong-Ling Cao¹, Yong-Biao Lin¹, Xue-Jun Oyang¹, Ru-Fang Deng¹, Yang Liu³, Wanhui Ye¹, and Ju-Yu Lian¹

¹South China Botanical Garden, Chinese Academy of Sciences

²Research Institute of Tropical Forestry Chinese Academy of Forestry

³University of the Chinese Academy of Sciences

April 28, 2020

Abstract

Serious disconnection among researches on individual physiological adaption, population genetic diversity and spatiotemporal demography has obstructed us in the knowledge of plant population ecology. Here we develop an approach to integrate those three research aspects by taking advantage of recent knowledge about DNA methylation and multivariate analysis. We show that by using various epigenetic parameters corresponding to individual physiological metabolic reprogramming potential, gene expression repression degree and physiological reaction characteristics, the contribution of various biotic and abiotic factors to an individual state and population structure can be quantified. Furthermore, population dynamics can be narrowly estimated by analysing DNA methylation of populations at different developmental stages. This study demonstrates an approach for the overall analysis of plant populations and exploring the spatiotemporal dynamics and developmental mechanisms of a population using *Castanopsis chinensis* as the model species.

Introduction

As the basic functional and assembly units of communities, populations play a critical role in ecological and evolutionary research. Because individual heterogeneity is ubiquitous in natural populations, to accurately assess population structure, status and destiny requires a basis in an integrative measure of differential responses of different genotypic individuals in different life cycle stages to environmental changes (Valpine *et al.*, 2014). By now, a robust body of literature has discussed the forming and maintenance mechanisms of plant population (Crone *et al.*, 2011; Fabritius, Singer, Pennanen, & Snäll, 2019). These papers focus on the genetic analysis of population structure, demographic analysis of population dynamics and physiological analysis of individual adaptive mechanism (Collins & Gardner, 2009; Wang *et al.*, 2014; Quintana-Ascencio *et al.*, 2018). However, different ecological definitions and scales among these three research aspects cut off their mutual connections and hinders us in exploring the spatiotemporal dynamics and developmental mechanisms of population, especially at the within-population level, although ecologists have tended to evaluate the predictive capability of models by integrating more and more population information into meta-matrix models (Crone *et al.*, 2011; Fabritius *et al.*, 2019; Plard *et al.*, 2019).

First, although population genetic structure has been extensively studied, neutral molecular markers (e.g., microsatellites or RFLP) used in genetic diversity research have failed to respond to environmental change (Andersen & Lubberstedt, 2003). Variation at marker loci is not an accurate predictor of genetic variation at loci contributing to phenotype, that is, adaptive variation. Because of this, analysis of population genetic structures focuses on seed and pollen dispersal and its effects on gene flow (Browne, Ottewell, Sork

& Karubian, 2018). Both seed and pollen dispersal are only preliminary steps for population expansion. The final population colonisation and expansion, as well as the formation of spatial distribution pattern, are determined by the interactions between individuals and habitat selection (HilleRisLambers, Adler, Harpole, Levine & Mayfield, 2012). Second, because of the lack of genetic diversity background of populations, experimental individuals used in physiological analysis can maybe not represent the whole population well enough. At the same time, physiological indicators, such as enzymes, products or substrates of biochemical reaction, among others, only represent one aspect of individual physiological state. Because of the randomness of individual and physiological parameter selection, population research results may be biased (Plard *et al.* , 2019). Third, in recent years, demographic population models have tried to address individual heterogeneity due to age, size and so on (Plard *et al.* , 2019). However, those indicators are not correlative with individual basic and instantaneous genetic physiological states. They are just the description of age-related cumulative growth (Browne *et al.* , 2018, Hamel *et al.* , 2017; Plard *et al.* , 2019).

Our research is motivated by those disconnections among research on populations. A key to eliminate the gap among different population research directions is to search for parameters that are definite, measurable and directly related with environment and individual genetic background. Responding to environmental variation, DNA methylation may trigger different gene expression patterns by blocking the promoters at which activating transcription factors could bind and then controlling individual developmental recombination to produce different phenotypes (Suzuki & Bird, 2008; Schubeler 2015). DNA methylation has attracted ecologists' attention in recent years, but relevant research has surrounded the relationship between population genetic structure and population DNA methylation-variable positions (MVPS) variation pattern and the qualitative relationship between MVPS variation and environmental factors (Heer, Mounger, Boquete, Richards & Opgenoorth, 2018; Moler *et al.* , 2018). The non-quantitative analysis and DNA methylation parameters' insufficiency makes it so DNA methylation in population research is still a largely unexplored area.

DNA methylation includes full-methylation and hemi-methylation. Full-methylation was proved to inhibit gene expression in a large body of evidence (Heard & Martienssen, 2014). However, hemi-methylation can either activate or repress gene expression, and activation seems to be more common than repression (Couldrey, Brauning, Henderson, & Mcewan, 2015; Fang *et al.* , 2016). Hemi-methylated CpG sites were proven to be enriched at core pluripotency loci, and DNA demethylation was enriched in these loci, which promoted somatic cell reprogramming (He *et al.* , 2019). Therefore, the DNA hemi-methylation rate is closely related to development reprogramming potential (Liu *et al.* , 2013; Heard & Martienssen, 2014; Couldrey *et al.* , 2015; Schubeler, 2015; Fang *et al.* , 2016; Harrison *et al.* , 2016; Heer *et al.* , 2018; Moler *et al.* , 2018; He *et al.* , 2019). Genome-wide MVPS variation at is a rounded description of individual physiological reaction (Couldrey *et al.* , 2015). By appropriately using various epigenetic parameters, DNA methylation can completely represent individual physiological information and then the population spatiotemporal dynamic (Baubec *et al.* , 2015; Schubeler, 2015).

Here, we use *Castanopsis chinensis* populations in Dinghu Mountain (DHS) as a model to demonstrate an approach in which we explore the spatiotemporal dynamics and developmental mechanisms of population by integrating various information on population with DNA methylation. *C. chinensis* is one of the most important constructive species widely distributed from South China to Vietnam. *C. chinensis* populations in DHS are in two different development stages. One population has been conserved since it was cut down more than 60 years ago (hereafter, the recovering population). The other population has never been disturbed as it is protected by the nearby Qinyun Temple for over 400 years (hereafter, the native population) (Chen, Rui, Zhou, Ye & Liu, 2016). These populations were studied for their population genetic structures (Wang *et al.* , 2014), demographic dynamics (Wang *et al.* , 2014; Chen *et al.* , 2016), habitat conditions (Chen *et al.* , 2016) and physiological adaptations. Now, integrating these various studies on *C. chinensis* population is much-needed for us to narrowly estimate population spatiotemporal dynamics.

Materials and Methods

Material, sample site and micro-environmental

A 20-hectare plot was selected in an evergreen broad-leaved forest located in DHS National Nature Reserve (112deg30' 39"-112° 33' 41" E, 23° 09' 21"-23° 11' 30" N). This forest is characterised by a south subtropical monsoon climate. This choice plot was divided into 500 20×20 m smaller plots and one *C. chinensis* individual in each small plot was selected randomly. We collected 308 individuals in recovering stands and 73 individuals in native stands (Fig. 1a-b). From each individual, 3–5 new completely spread leaves were collected and mixed to extract DNA using the CTAB method. For individual variables, we recorded diameter at breast height (DBH) and 14 environmental variables including soil water content (SWC), soil bulk density (BD), soil pH values, soil organic matter (SOM), available potassium (AK), available phosphorus (AP), available nitrogen (AN), total nitrogen (TN), total phosphorus (TP), total potassium (TK), altitude, aspect, slope, and convexity as described by Chen *et al.* (2016). *C. chinensis* is a wind-pollinated plant, and seeds are dispersed by animals (Peng & Xu, 2005). The general information about the soil and topography of the two stands is given (Fig. 1c).

MSAP and SSR

SSR data were extracted from published SSR experimental results of *C. chinensis* in DHS (Wang *et al.*, 2014). Methylation sensitive amplified polymorphism (MSAP) is widely used in studies of population epigenetic variation for its being simple and relatively inexpensive (Latiff, 2009). MSAP in this paper follows the way described by Schulz *et al.* (2014). A set gradient parameters method and orthogonal experiment method was used to optimise the optimal reaction system and adapter sequences (see supporting information: The optimal reaction system, Table S1). First, 128 primers were randomly combined based on the EcoR I and Hap II/Msp I universal primers. After polymorphic amplifying, the products were electrophoresed and screened by ABI 3730XL DNA. Based on the visualised peak figure, six primers that could amplify clear, complete and high polymorphism bands were selected (Table S2). GS500 was used as the DNA standard. To ensure that the data were credible, using GeneMarker version 2.2.0 (SoftGenetics), 25–30 well-defined sites were selected from ~100–500 bp fragments for each primer in later research. The fragment statistic also followed Schulz's way (Schulz *et al.*, 2014).

A controversial defect in MSAP is that scoring of MSAP data may lead to ambiguous interpretation of MSAP data because a cytosine in any sequence context can be methylated with decreased frequency for CG, CHG and CHH motifs (H = A or T or C) (Fulneček & Kovařík, 2014). In addition, full methylation of external cytosines (mCCGG) cannot be detected by MSAP (Xia, Zou, Zhang, Feng & Wang, 2014). However, a screening of literature revealed that, for quantification of genome-wide cytosine methylation, methylation sites assayed by MSAP was greater than those obtained by high-performance liquid chromatography, which is recommended because of its global assessment, accuracy and reproducibility (Alonso, Pérez, Bazaga, Medrano & Herrera, 2016). Furthermore, full-methylation and hemi-methylation show clear distinction in MSAP analysis.

Data analysis

General statistical analysis

Before elucidating the relationships among the methylation states, environmental factors and individual characteristics, we used the Kolmogorov-Smirnov test to detect whether the assumptions of normality were satisfied. If not, a non-parametric test was employed. All statistical tests were performed in R v. 3.1.0. The spatial analysis, multiple regressions on distance matrices analysis (MRM) and multivariate variation partitioning were performed as has been shown before (Chen *et al.*, 2016). The details are as follows.

Spatial analysis

Individual methylation rate and epigenetic variation may be associated with geographical distance. We used principle components of neighbourhood matrices (PCNM) implemented in the R package “vegan” (Oksanen, 2015) to model those relationships. Moran's eigenvector mapping (MEM) was used to calculate spatial variables with R. Before calculating the semivariance, polynomial trend-surface regressions on each variable were performed using the method described in Chen (Chen *et al.*, 2016). Residuals from the trend-surface

regressions were used to generate the semivariograms and the confidence envelopes in spatial autocorrelation analyses. In short, in each lag class, sample values were randomly selected and a semivariogram was calculated based on the selected data in 10,000 randomisations. A 95% confidence envelope was then obtained by calculating the 2.5 and 97.5 percentiles of the randomly generated semivariance values in each lag class.

The relative contribution of edaphic, topography, DBH and spatial to methylation rate

Variation-partitioning is a method using tables of various variables. By adjusting the coefficient of determination, partitioning provides an unbiased estimate in both multiple regression and canonical redundancy analysis on the variation of methylation level data table that is explained by the various tables of explanatory variables. In here, the multivariate variation-partitioning approach was conducted by using the R package “varpart” to quantify the relative contributions of microhabitat, individual life stage and spatial attributes, as well as their potential intersections, to the variation of methylation rate. Before the variation partitioning analysis, the variables most significantly influential to methylation level were chosen by permutational forward model selection by using the method developed previously (Chen *et al.*., 2016) with the R package “packfor”.

Statistical estimation the effect of edaphic, topography, DBH, genetic variation and spatial on MVPS

The relative influence of microhabitat, individual life stage, genetic variation and spatial attributes on MVPS were determined by using MRM in R 3.1.0 with the “ecodist” packages (Goslee & Urban, 2007). MRM, an extension of partial Mantel analysis, can employ non-parametric or non-linear multiple regression. In this study, independent variables increasing one by one from single edaphic variables to multiple variables, such as edaphic, topography, spatial, individual life stage and genetic variation were used to construct five MRM models. We first used Spearman correlations to assess the relationships between environmental distance matrices and the MVPS matrix; then the significant explanatory variables were identified using a forward-selection procedure; finally, the significance of MRM models was tested with 1,000 permutations on the objects of the response distance matrix. The r^2 from MRM analyses was used to assess MVPS variation explained by these significant environmental variables, and the significantly correlated variables were excluded within each group of variables (edaphic, topography, individual life stage, genetic variation and spatial variables). The relationship between DBH and methylation level was also analysed with MRM.

Results

The environment, demographic and epigenetic difference between two stands

There are significant difference between two populations in habitat condition and demography (Fig. 1a–e). Except for available phosphorus, other measured edaphic elements are higher in native stand than in recovering stand (Fig. 1c). The proportion of large diameter *C. chinensis* individuals (DBH>40 cm) changes from less than 15% in recovering forests to more than 75% in native forests (Fig. 1d–e). The mean methylation ratios in recovering and native *C. chinensis* populations were 19.94% and 20.06%, and there was no significant difference between them ($p>0.05$). However, the h- and f-type methylations in recovering populations were 1.135 and 0.91 times those in native populations (Fig. 1f).

Effects of environment and DBH on individual methylation level

Spatial variable, edaphic variables, topographic variables and DBH can explain over 40% of various methylation rate variations in recovering stands but less than 30% in native stands (Fig. 2a–g). Their contribution models were different in two populations. Contribution of soil factors (including alone and jointly with spatial factors and topography) to total methylation rate variation in two stands declined slowly, by 28.4% and 23.2% in recovering and native stands, respectively (Fig. 2d, 2g). However, the contribution of spatial factors changed greatly in different stands. Pure contribution and joint contribution of spatial factors to total methylation rate variation were 21.8% and 20.1% in recovering stands, but declined to 6.2% and 15.3%, respectively, in native stands. Contribution of topography to methylation increased from 3.2% in recovering stands to 13% in native stands.

For special edaphic factors, in recovering stands, most edaphic factors provided various degrees of contribution to full-methylation level variation (Fig. 2b), but only soil bulk density (BD), total nitrogen (TN), total phosphorus (TP), available phosphorus (AP) and elevation showed significant correlation with the variation of hemi-methylation levels (Fig. 2c). In native stands, the edaphic contribution to full-methylation level variation was mainly provided by total potassium (TK), TP, BD and elevation (Fig. 2e), and only TK and elevation had significant correlation with the level of hemi-methylation variation (Fig. 2f).

DBH did not have independent contribution to methylation rate in either recovering or native stands (Fig. 2b–g). However, in native stands, together with soil, topography and spatial factors affected 8.3% of the full-methylation rate variation, which was over one-third of all full-methylation variation interpreted by all four variables (Fig. 2e). Multiple regressions on distance matrices analysis (MRM) showed that DBH had a negative correlation with the hemi-methylation rate and positive correlation with the full-methylation rate (Fig. 3a). DBH also showed a significant inverse relationship with individual relative growth rates (Fig. 3b).

Effects of biotic and abiotic factors on the MVPS variation

The contributions of edaphic, topography and spatial factors to the variation of MVPS are similar to their contributions to the variation of methylation rates. In the recovering forest, soil, topography and spatial factors showed significant contribution to the variation of MVPS via BD, P, mean elevation, convex and spatial distance, respectively, but DBH and individual genetic background did not show significant effects on MVPS (Table 1). In native forests, only the aspect of topographic factors showed significant effects on MVPS (Table 1). Therefore, an individual reactive mechanism in recovering forests has a significant relationship with soil and spatial factors, but in native forest, this relationship disappears. MVPS variation in the recovering stand is negatively correlated with TP and AP and was significantly positively correlated with BD.

Epigenetic spatial autocorrelation analysis showed that only among DBH < 40 cm individuals in recovering stands is there a significant positive spatial autocorrelation in the 140 m range and a significant negative correlation in the range of 220–440 m (Fig. 4a). For other subpopulations, there was no significant spatial autocorrelation (Fig. 4b, c, d). *C. chinensis* populations did not show significant genetic spatial autocorrelation in all four individual sets (Fig. 4e–h).

Discussion

The analysis on *C. chinensis* shows that the contributions of various biotic and abiotic factors to individual physiological states and population structure can be quantified by integrating various information on population with DNA methylation analysis, and then the population dynamics can be narrowly estimated.

During population development from expansion or early stage maturation (in the recovering stand) to the top stage of maturation (in the native stand), accompanied by the change of habitat environment and individual DBH range, relative contribution of soil and spatial factors to methylation decreased, and MVPS variation spatial autocorrelation of low DBH subpopulations disappeared (Fig. 4a–d, Table 1). All of that indicates that, following the development of *C. chinensis* population, the effects of soil and spatial factors on individual development decline and population structure tends to disorder. The spatial epigenetic and genetic autocorrelation represents the aggregation distribution of functional group and kin individuals, respectively (Wang *et al.*., 2012; Huang *et al.*., 2015). The disappearance of epigenetic spatial autocorrelation, which accompanied the decline in individual amount but increase in mean DBH (Fig. 1d–e), indicates a functional self-thinning process. This functional self-thinning may result from the competitive exclusion between same functional individuals, then a functional diversity distribution around high DBH individuals forms.

At an individual level, the positive and negative correlations of DBH with the full-methylation rate and hemi-methylation rate are consistent with the result that the relative growth rates decline as the DBH increases in *C. chinensis* (Fig. 3a,b), because full-methylation and hemi-methylations determine the repressed degree of individual vitality and the potential of development reprogramming. Furthermore, following the increase of edaphic N concentration and decrease of edaphic AP concentration from the recovering stand to the native

stand, soil AP constraints aggravate (Fig. 1c) (Hou *et al.* 2014; Chen *et al.* 2016; Turner, Brenes-Arguedas & Condit, 2018). However, AP lost its significant effect on DNA methylation variation in the native stand, in which over 75% individuals' DBH is over 40 cm and the DBH effect on total full-methylation rate rises to 8.3% through interactions with environmental factors (Fig. 2b–g, Table 1). DBH as an important plant phenotypic indicator may affect individual physiological characteristics (Bustos-Segura *et al.* 2017). The transformation of plant nutritional needs provides a new way for us to redefine nutrient limitations during primary succession of soil formation in subtropical forests (Fig. 2b–g, Table 1) (Turner *et al.* 2018). The temporal change on both from individual active to inactive and the transformation of plant nutritional needs are favourable for population maintenance.

The presented approach has several advantages. First, the statistical data of genome-wide, methylation rates that determine individual physiological state and the MVPS that determines the reaction model represent individual characteristics well enough (Suzuki & Bird, 2008). In addition, because the variation of methylation rate and MVPS are direct-acting results of biotic and abiotic factors on individuals and are the direct reason for individual physiological adjustment and population structure change (Schubeler, 2015), all analyses do not create uncertainty because of indirect relationships between variables. Thus, the parsing results and their variation trends of the contribution of various biotic and abiotic impact factors on individual DNA methylation are the closest presentation for a logical and a time relationship of individual development fitness to microhabitat. Similarly, epigenetic spatial autocorrelation analysis results are the intuitive description of population functional group structure.

Second, both the MRM analysis and multivariate variation-partitioning analysis can be employed in non-parametric or non-linear analysis, and their analytical precision can be improved by expanding the number of explanatory matrices, allowing more environmental variables to be represented by their own distance matrices (Goslee & Urban, 2007; Oksanen, 2015). For example, as a stable influencing factor, topography abidingly affects surface runoff and the distribution of soil elements (Latiff, 2009) and also relates to light, humidity and temperature in microhabitats. How topographic factors affect plant physiological characteristics can be illuminated by adding those microenvironmental and microclimate factor matrices into analysis after permutational forward model selection.

Third, DNA methylation contains both instantaneous and cumulative information (Suzuki & Bird, 2008; Heard & Martienssen, 2014; Schubeler, 2015; Harrison *et al.* 2016; Heer *et al.* 2018; Moler *et al.* 2018). DNA methylation parameters in this research were obtained from leaf organisation, which renewed their response mechanism to both the instantaneous environmental condition and age-related DBH indicator. The implications of this bilayer extend the range of methylation research to individuals and populations on time scales.

Many studies on the relationship between genetic structure and epigenetic structure have reached different conclusions (Herrera, Medrano, Monica & Bazaga, 2016; Heer *et al.* 2018; Moler *et al.* 2018). For different spatial scales and different species, the contribution of genetic background to DNA methylation may be different. Therefore, that contribution should always be noted in research. On the other hand, *C. chinensis* has many common characteristics of constructive species of a subtropical, evergreen, broad-leaved forest, such as being perennial, outcrossing and occupying the top space of the community. All those characteristics may reduce genetic diversity within the population, and then the relationship between genetic structure and epigenetic structure (Herrera *et al.* 2016; Moler *et al.* 2018). Thus, the functional self-thinning and its effect on population development and community construction, which appears in *C. chinensis* populations, may provide enlightenment in within-population research of other subtropical constructive species.

The presented approach still has some room for improvement. First, DNA methylation analysis can join with a next-generation sequencing platform (Xia *et al.* 2014). The new method can eliminate constraints associated with the methylation-resolving power of the gel and explain adaptive variation at gene level. Secondly, besides the three most distinctive DNA methylation indicators used in this study, other DNA methylation indicators, such as epigenetic distance between individuals, also represent some detailed information on population function diversity (Herrera *et al.* 2016). How to integrate more DNA methylation indicators into

this presented approach is an interesting topic.

ACKNOWLEDGEMENTS

We are grateful to L. Wu for assistance in the field. We thank Proof-Reading-Service.com Ltd in Letchworth Garden city for the helpful feedback on a previous draft of the manuscript. This study supported by Strategic Priority Research Program of the Chinese Academy of Sciences (XDB31030000), the National Key R&D Program of China (grand No. 2017YFC0505802), National Natural Science Foundation of China (No. 31370437), National Natural Science Foundation of China (No. 41371078) and Chinese Forest Biodiversity Monitoring Network.

References

- Alonso, C., Perez, R., Bazaga, P., Medrano, M., & Herrera, C. M. (2016). MSAP markers and global cytosine methylation in plants: A literature survey and comparative analysis for a wild-growing species. *Molecular Ecology Resources*, 16 (1), 80-90.
- Andersen, J. R., & Lubberstedt, T. (2003). Functional markers in plants. *Trends Plant Sci.*, 8 (11), 554–560.
- Baubec, T., Colombo, D. F., Wirbelauer, C., Schmidt, J., Burger, L., Krebs, A. R., ... Schubeler, D. (2015). Genomic profiling of dna methyltransferases reveals a role for dnmt3b in genic methylation. *Nature*, 520, 243-247.
- Browne, L., Ottewell, K., Sork, V. L., & Karubian, J. (2018). The relative contributions of seed and pollen dispersal to gene flow and genetic diversity in seedlings of a tropical palm. *Molecular Ecology*, 27 (15), 3159–3173.
- Bustos-Segura, C., Poelman, E. H., Reichelt, M., Gershenzon, J., Gols, R., & Scherber, C. (2017). Intraspecific chemical diversity among neighbouring plants correlates positively with plant size and herbivore load but negatively with herbivore damage. *Ecology Letters*, 20 (1), 87-97.
- Chen, J., Rui, Y., Zhou, X., Ye, W., & Liu, W. (2016). Determinants of the biodiversity patterns of ammonia-oxidizing archaea community in two contrasting forest stands. *Journal of Soils and Sediments*, 16 (3), 878-888.
- Collins, S., & Gardner, A. (2009). Integrating physiological, ecological and evolutionary change: a Price equation approach. *Ecol. Lett.*, 12 (8), 744– 757.
- Couldrey, C., Brauning, R., Henderson, H. V., & McEwan, J. C. (2015). Genome-wide dna methylation analysis: no evidence for stable hemimethylation in the sheep muscle genome. *Animal Genetics*, 46 (2), 185-189.
- Crone, E. E., Menges, E. S., Ellis, M. M., Bell, T., Bierzychudek, P., Ehrlén, J., ... Williams, J. L. (2011). How do plant ecologists use matrix population models?. *Ecology Letters*, 14 (1), 1-8.
- Fabritius, H., Singer, A., Pennanen, J., & Snäll, T. (2019). Estimation of metapopulation colonization rates from disturbance history and occurrence-pattern data. *Ecology*, 100 (10), 10.1002/ecy.2814.
- Fang, J., Cheng, J., Wang, J., Zhang, Q., Liu, M., Gong, R., ... Xu Y. (2016). Hemi-methylated dna opens a closed conformation of uhrf1 to facilitate its histone recognition. *Nature Communications*, 7, 11197.
- Fulneček, J., & Kovařík, A. (2014). How to interpret Methylation Sensitive Amplified Polymorphism (MSAP) profiles?. *BMC Genet* 15, 2. doi:10.1186/1471-2156-15-2.
- Goslee, S. C., & Urban, D. L. (2007). The ecodist Package for Dissimilarity-based Analysis of Ecological Data. *Journal of Statistical Software*, 22 (7), 1-19.
- Hamel, S., Gaillard, J. M., Douhard, M., Festa-Bianchet, M., Pelletier, F., & Yoccoz, N. G. (2017). Quantifying individual heterogeneity and its influence on life-history trajectories: different methods for different questions and contexts. *Oikos*, 127 (5), 687– 704.

- Harrison, J. S., Cornett, E. M., Goldfarb, D., DaRosa, P. A., Li, Z. M., Yan, F., ... Rothbart, S. B. (2016). Hemi-methylated DNA regulates DNA methylation inheritance through allosteric activation of H3 ubiquitylation by UHRF1. *eLife Sciences*, 6 (10152) 10.7554/eLife.17101.
- He, S., Wang, F., Zhang, Y., Chen, J., Liang, L., Li, Y., ... Zheng, H. (2019). Hemi-methylated cpg sites connect dnmt1-knockdown-induced and tet1-induced dna demethylation during somatic cell reprogramming. *Cell Discovery*, 5 (11), 10.1038/s41421-018-0074-6.
- Heard, E., & Martienssen, R. A. (2014). Transgenerational epigenetic inheritance: myths and mechanisms. *Cell*, 157 (1), 95–109.
- Heer, K., Mounger, J., Boquete, M. T., Richards, C. L., & Opgenoorth, L. (2018). The diversifying field of plant epigenetics. *New Phytologist*, 217 , 988– 992.
- Herrera, C. M., Medrano, Mónica, & Bazaga, P. (2016). Comparative spatial genetics and epigenetics of plant populations: heuristic value and a proof of concept. *Molecular Ecology* , 25 (8), 1653-1664.
- HilleRisLambers, J., Adler, P. B., Harpole, W. S., Levine, J. M., & Mayfield M. M. (2012). Rethinking community assembly through the lens of coexistence theory. *Annu Rev Ecol Evol Syst* 43 , 227–248.
- Hou, E. Q., Chen, C. R., Wen, D. Z., & Liu, X. (2014). Relationships of phosphorous fractions to organic carbon content in surface soils in mature subtropical forests in Dinghushan, China. *Soil Research* 52, 55-63.
- Huang, C. L., Chen, J. H., Tsang, M. H., Chung, J. D., Chang, C. T., & Hwang, S. Y. (2015). Influences of environmental and spatial factors on genetic and epigenetic variations in rhododendron oldhamii (ericaceae). *Tree Genetics & Genomes*, 11 (1), 823. DOI: 10.1007/s11295-014-0823-0
- Latiff, A. (2009). Effect of topography and soil on the distribution of under canopy trees of garcinia (guttiferae) in lowland forest of peninsular malaysia. *International Journal of Botany*, 5 (4), 287-294.
- Liu, X., Gao, Q., Li, P., Zhao, Q., Zhang, J., Li, J., ... Wong, J. (2013). Uhrf1 targets dnmt1 for dna methylation through cooperative binding of hemi-methylated dna and methylated h3k9. *Nature Communications* 4, 1563.
- Moler, E., Abakir, A., Eleftheriou, M., Johnson, J. S., Krutovsky, K. V., & Lewis, L. C. (2018). Population Epigenomics: Advancing Understanding of Phenotypic Plasticity, Acclimation, Adaptation and Diseases. In: Rajora O. (eds) Population Genomics. Population Genomics. Springer, Cham
- Oksanen, J. (2015). *The vegan package. Community ecology package* . (R package version 2.3-2, 2015).
- Peng S. J., & Xu, G. L. (2005). Seed traits of castanopsis chinensis and its effects on seed predation patterns in dinghushan biosphere reserve. *Ecology and Environment*, 14 (4), 493-497.
- Plard, F. Fay, R., Kery, Marc., Cohas, A., & Schaub, M. (2019). Integrated population models: powerful methods to embed individual processes in population dynamics models. *Ecology*, 100 (6), 10.1002/ecy.2715.
- Quintana-Ascencio, P. F., Koontz, S. M., Smith, S. A., Sclater, V. L., David, A. S., & Menges, E. S. (2018). Predicting landscape-level distribution and abundance: integrating demography, fire, elevation, and landscape habitat configuration. *Journal of Ecology*, 106 (6), 2395– 2408.
- Schubeler, D. (2015). Function and information content of DNA methylation. *Nature*, 517 (7534), 321–326.
- Schulz, B., Eckstein, R. L., & Durka, W. (2014). Epigenetic variation reflects dynamic habitat conditions in a rare floodplain herb. *Molecular Ecology*, 23 (14), 3523– 3537.
- Suzuki, M. M., & Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet.*, 9 (6), 465-476.
- Turner, B.L., Brenes-Arguedas, T., & Condit R. (2018). Pervasive phosphorus limitation of tree species but not communities in tropical forests. *Nature*, 555 (7696), 367-370.

Valpine, P., Scranton, K., Knape, J., Ram, K., Mills, N.J., & Liebhold, A. (2014). The importance of individual developmental variation in stage-structured population models. *Ecology Letters*, 17 (8), 1026–1038.

Wang, Z. F., Lian, J. Y., Huang, G. M., Ye, W. H., Cao, H. L., & Wang, Z. M. (2012). Genetic groups in the common plant species *castanopsis chinensis* and their associations with topographic habitats. *Oikos*, 121 (12), 2044–2051.

Wang, Z. F., Lian, J. Y., Huang, G. M., Ye, W. H., Cao, H. L., & Wang, Z. M. (2014). The spatial genetic pattern of *Castanopsis chinensis* in a large forest plot with complex topography. *Ecology Manage*, 318 , 318–325.

Xia, Z., Zou, M., Zhang, S., Feng, B., & Wang, W. (2014). AFSM sequencing approach: a simple and rapid method for genome-wide SNP and methylation site discovery and genetic mapping. *Sci Rep.*, 4, 7300 doi:10.1038/srep07300.

Data Accessibility Statement

The data supporting the results and the R scripts used to generate the analyses are archived in the Figshare data repository: <https://doi.org/10.6084/m9.figshare.11953764>. Data are available from the corresponding author upon reasonable request.

Author Contributions

WL, HC, and WY, conceived and designed the study. WL, JL, CJ, XO, RD, YBL and YL, carried out the experiments and collected data. WL, JC, and ZW, performed statistical analyses. WL, JL, and CJ, wrote the first draft of the manuscript, and all authors contributed to revisions. All authors reviewed and approved the final paper.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table 1. Different multiple regression models predicting the responses of MVPS variation to the variables. Independent variables of models increased one by one from the single edaphic variables to include topography, spatial, individual life stage, genetic variation and edaphic variables.

Stand		Variables	Coefficient [*]	p	r ²	p
Recovering	edaphic	BD	84.116	0.002 ⁺	0.043	0.00
		BD	82.723	0.002 ⁺	0.054	0.00
		AP	-2.359	0.043 ⁺		
		convex	-0.641	0.01 ⁺		
	edaphic + topographic + geographical distance	BD	37.945	0.05	0.089	0.00
		TP	-118.108	0.026 ⁺		
		AP	-2.781	0.032 ⁺		
		convex	-0.535	0.024 ⁺		
		mean elevation	-0.145	0.002 ⁺		
		Geographical Distance	0.079	0.001 ⁺		
Native	topographic	aspect	0.090	0.017 ⁺	0.070	0.2

^{*} Coefficients of Spearman correlations, r² (ranged 0–1), 10,000 permutations; ⁺ p<0.05

BD soil bulk density, TP total phosphorus, AP available phosphorus,

Figure Legends

Fig. 1. The difference between two populations in distribution, environment, demography and DNA methylation rates. **(a)**, **(b)**: individual distribution in recovering stand and native stand. **(c)**: environment of two populations. **(d)**, **(e)**: Diameter at breast height (DBH) distribution of population in recovering stand and native stand. **(f)**: Methylation rates of two populations; data show averages (+ SD), and different superscript letters indicate significant differences among different stands, $P < 0.05$.

Fig. 2. Effects of biotic and abiotic factors on individual methylation rate. **(a)–(g)**: Variation partitioning of *C. chinensis* methylation rates variation in **(b)–(d)**, recover, and **(e)–(g)**, native stand. The number is the significant interpretation rate; the residual is the unexplained part. **(a)**: visualization outline **(b)**, **(e)**: full methylation rate **(c)**, **(f)**: hemi-methylation rate **(d)**, **(g)**: total methylation rate. The significant edaphic and topographic variables are presented in the table of each subgraph with the contributed variation.

Fig. 3. The relationships between DBH and methylation rates and relative growth rate. **(a)**: MRM correlation coefficients for the relationship between individual DBH and methylation rates. * Significant at the 0.1 level. An edaphic, topography, spatial, DBH and genetic variable model was used in this MRM analysis. **(b)**: DBH distribution of relative growth rates (RGR) of all individuals of *C. chinensis* in DHS.

Fig. 4. Spatial autocorrelation of populations. Epigenetic **(a)–(d)** and genetic **(e)–(h)** autocorrelation analysis of different diameter groups. **(a)**, **(e)**. DBH < 40 cm individuals in recovering forests, **(b)**, **(f)**: DBH > 40 cm individuals in recovering forests, **(c)**, **(g)**: DBH < 40 cm individuals in native forests and **(d)**, **(h)**: DBH > 40 cm individuals in native forests, respectively. The solid line represents the autocorrelation coefficient R (with positive and negative deviations), and the dashed line represents the 95% confidence interval.

Hosted file

Fig1.docx available at <https://authorea.com/users/314967/articles/445316-integrating-population-information-using-dna-methylation-to-explore-the-spatiotemporal-dynamics-and-developmental-mechanisms-of-populations>

Hosted file

Fig2.docx available at <https://authorea.com/users/314967/articles/445316-integrating-population-information-using-dna-methylation-to-explore-the-spatiotemporal-dynamics-and-developmental-mechanisms-of-populations>

Hosted file

Fig3.docx available at <https://authorea.com/users/314967/articles/445316-integrating-population-information-using-dna-methylation-to-explore-the-spatiotemporal-dynamics-and-developmental-mechanisms-of-populations>

Hosted file

Fig4.docx available at <https://authorea.com/users/314967/articles/445316-integrating-population-information-using-dna-methylation-to-explore-the-spatiotemporal-dynamics-and-developmental-mechanisms-of-populations>