Natural patterns of plant epigenetic variation:

Understanding the link between DNA methylation and environmental conditions

Dissertation

„kumulativ“

zur Erlangung des Grades eines

Doktor der Naturwissenschaften

(Dr. rer.nat.)

des Fachbereichs Biologie der Philipps-Universität Marburg

Vorgelegt von

Bárbara Díez Rodríguez

Aus Burgos, Spanien

Marburg, Januar 2023

Die vorliegende Dissertation wurde von 04/2018 bis 01/ 2023 am Fachbereich Biologie, Pflanzenökologie und Geobotanik unter Leitung von Prof. Dr. Lars Opgenoorth angefertigt.

Vom Fachbereich Biologie der Philipps-Universität Marburg
(Hochschulkennziffer 1180) als Dissertation angenommen am _____

Erstgutachter(in): Prof. Dr. Lars Opgenoorth
Zweitgutachter(in): Prof. Dr. Katrin Heer

Tag der Disputation: 28 März 2023

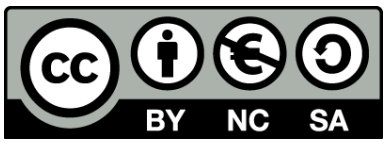
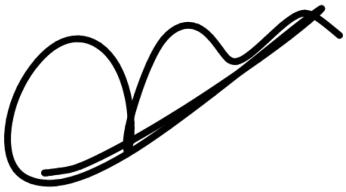


Table of contents

Declaration

I hereby declare that this PhD thesis represents my own work and that I have not received unauthorized help. I declare that this dissertation has not yet been submitted to any other university in its present or similar form and has not served any other examination purposes.

Burgos, 16.01.2023 Bárbara Díez Rodríguez



Acknowledgements

I want to start by thanking my supervisors, Lars Opgenoorth and Katrin Heer, for giving me the opportunity of being part of the EpiDiverse project. Without them and their invaluable help and support, this research would not have been possible. I also want to thank my external advisor, Koen Verhoeven, for his support over the years and all the advice he has given us.

A big thank you goes to the EpiDiverse family, for all the great times we spent together. The summer (winter) schools wouldn't be the same without them, and I hope we can keep meeting and being the last ones to go home. However, some people deserve a special mention. First, the Poplar group, in particular Paloma and Cristian, because the time we spent discussing results and their help with data analysis contributed more than I can express to this research. But also, Emanuele, who was not my supervisor officially, but took the same role nonetheless, and whose input was always very valuable. Second, the Natural Patterns people, Iris and Dario. This thesis would not be the same without them, and I will be always grateful for our discussions about cross-species patterns and their patience with my questions. And third, Adam, because I can honestly say he saved all of us with his help. I could not ask for better friends and co-authors.

I also want to thank Sarina, Belén and Jill. Marburg would have been something else without our dinner and caipi nights. Sarina, thank you for letting me complain as much as I did, and for ignoring my inability to navigate the German culture so many times. I haven't found a better 9am coffee mate yet. Belén, thank you for being an amazing friend during all these years, it was great to have a piece of Spain there. And Jill, thank you for being a sound wall of chillness, you helped me avoid a very deep well of craziness. I'd also like to extend my gratitude to all the other people I met in our working groups: Birgit, Tina, Sascha, Galina, Cristian, Silke, Mona and Eric. Thank you for all your help with everything.

Furthermore, I would like to thank Benedicte Albrectsen and Karen Kloth, for introducing me to the world of plant research, and supporting me still, after all these years.

Finally, I want to thank all the members of my family, for supporting me all the time, with everything I have wanted to do. I would not be here today without the support of my parents and my brother, who have been with me through all my academic life without complaining (too much). Special mention to my uncle Eloy, who has financed my coffee needs since I first became addicted to it. I am who I am today because of them.

Declaration of author contributions

This thesis, “Natural patterns of epigenetic variation: understanding the link between DNA methylation and environmental conditions”, is based on the work I did during my PhD at the University of Marburg, supervised by Prof. Dr. Lars Opgenoorth and Prof. Dr. Katrin Heer. **Chapters II–VI** of this thesis include five independent scientific manuscripts, each containing co-authorship, and all are or will be published. The contribution of the authors for each chapter is stated as following:

Chapter II – Using common gardens to study natural epigenetic variation

Publication: *“An uncommon garden experiment: microenvironment has stronger influence on phenotypic variation than epigenetic memory in the clonal Lombardy poplar”*

Bárbara Díez Rodríguez, Cristian Peña Pontón, Paloma Pérez Bello, Julius Bette, Lena Lerbs, Tabea Mackenbach, Sven Wulle, Emanuele De Paoli, Koen J.F. Verhoeven, Katrin Heer, Lars Opgenoorth

Experimental design, data collection, data analysis and interpretation, manuscript writing. Total contribution: 90%

Status in publication process: Submitted to *Annals of Botany*

Chapter III – Effects of environmental conditions on DNA methylation

Publication: *“High-resolution methylome analysis in the clonal Populus nigra cv. 'Italica' reveals environmentally sensitive hotspots and drought-responsive TE superfamilies”*

Cristian Peña Pontón, **Bárbara Díez Rodríguez**, Paloma Pérez Bello, Claude Becker, Lauren M. McIntyre, Wim van der Putten, Emanuele De Paoli, Katrin Heer, Lars Opgenoorth; Koen J.F. Verhoeven

Data interpretation, manuscript editing. Total contribution: 10%

Status in publication process: in preparation

Chapter IV – Natural patterns of DNA methylation

Publication: *“Epigenetic variation in the Lombardy poplar across climatic gradients is independent of genetic structure and persists across clonal reproduction”*

Bárbara Díez Rodríguez, Dario Galanti, Adam Nunn, Cristian Peña-Ponton, Paloma Pérez-Bello, Iris Sammarco, Katharina Jandrasits, Claude Becker, Emanuele de Paoli, Koen J.F Verhoeven, Lars Opgenoorth, Katrin Heer

Experimental design, data collection, data analysis and interpretation, manuscript writing. Total contribution: 90%

Status in publication process: In preparation

Chapter V – Epigenetic transgenerational inheritance

Publication: *"DNA methylation in the wild: epigenetic transgenerational inheritance can mediate adaptation in clones of the wild strawberry (Fragaria vesca)"*

Iris Sammarco, Zuzana Münzbergová, Claude Becker, Oliver Bossdorf, **Bárbara Díez Rodríguez**, Dario Galanti, Adam Nunn, Vít Latzel

Data analysis support, manuscript editing. Total contribution: 10%

Status in publication process: In preparation.

Chapter VI – Tailored tool for ecological epigenetics

Publication: *"EpiDiverse Toolkit: a pipeline suite for the analysis of bisulfite sequencing data in ecological plant epigenetics"*

Adam Nunn, Sultan Nilay Can, Christian Otto, Mario Fasold, **Bárbara Díez Rodríguez**, Noé Fernández-Pozo, Stefan A Rensing, Peter F Stadler, David Langenberger

Data analysis (testing of existing code components), empirical data, manuscript editing. Total contribution: 10%

Published in *NAR Genomics and Bioinformatics*, Volume 3, Issue 4 (December 2021)

Summary

In the last couple of decades, extreme weather events have been increasing, exceeding plants' tolerance thresholds, and driving mass mortalities in many tree species. Furthermore, many studies suggest that due to their longevity, trees are not able to adapt rapidly enough to keep pace with global climate change. Understanding how trees respond to such weather events and other environmental conditions (such as biotic stress) has thus become crucial for conservation policies and forest management programs. To cope with unpredictable environmental conditions, plants have evolved the ability to alter their physiology, morphology, or development or, in other words, the ability to produce different phenotypes from one genotype. This ability is called phenotypic plasticity, and it plays a major role in plant adaptation. Three factors have been suggested to increase phenotypic variability and thus potentially the resilience of tree populations: intraspecific genetic variability, (micro-)environmental variation, and epigenetic variation. Indeed, a growing body of literature suggests that epigenetic variation might contribute to local adaptation of natural plant populations. Epigenetic mechanisms, such as DNA methylation, can in fact quickly alter phenotypes in response to environmental changes. Variation in DNA methylation can be under genetic control, arise stochastically, or be induced by environmental conditions. Furthermore, phenotypic changes induced by epigenetic variation can be inherited across several generations (especially across clonal generations), suggesting that variation in DNA methylation might contribute to heritable phenotypic variation and, eventually, to adaptation. Although considerable progress has been made in recent years, the link between epigenetic variation and phenotypic variation remains poorly understood.

With this thesis, I aimed to further investigate this link, addressing the following questions: i) To what extent does epigenetic variation contribute to local adaptation? More specifically, ii) does DNA methylation vary in response to different environmental cues? iii) Can DNA methylation variation be independent of the underlying genetic structure? And, if so, iv) are DNA methylation patterns transmitted to the clonal offspring? **Chapter I** is a general introduction to the field of ecological epigenetics, including how epigenetic variation can affect phenotypic plasticity, epigenetic mechanisms, and the effects of epigenetic variation in some ecological processes. In **Chapter II**, I address the suitability of a common garden experimental design for studies on epigenetic variation. I used the *Populus nigra* "Italica" cultivar to assess whether phenotypic differences in functional traits could be observed under common garden conditions. I used linear mixed models and climatic data to study if these phenotypic differences can be traced back to

the environmental origin of the poplar clones. **Chapter III** uses a subset of the poplar individuals collected across Europe for the common garden to determine if DNA methylation is affected by cues associated with biotic (rust, herbivory and salicylic acid) and abiotic (heat, cold and drought) stress conditions. In **Chapter IV**, I used the EpiDiverse toolkit (described in **Chapter VI**) and other statistical methods to analyze epigenomic data and study natural methylation patterns associated with climatic conditions at a landscape level, and assess if methylation profiles can be transmitted to the clonal offspring. Finally, **Chapter V**, uses epigenomic data to study the contribution of climatic conditions to variation in DNA methylation in *Fragaria vesca* wild populations and the potential effects of this variation in gene expression. My work shows that i) phenotypic differences associated to geographic origin can be observed under common garden conditions despite the effects of microenvironment on phenotypic plasticity, thus proving that common garden experiments are suitable to study natural epigenetic variation (**Chapter II**). In poplar, abiotic stress conditions, such as heat and drought, appear to elicit a stronger response than biotic stresses in the methylome **(Chapter III)**. Furthermore, methylome stress responses can be global or stress-specific and might persist as epialleles (epigenetic alleles) in natural conditions. I also found that ii) variation in the methylome can be independent of the underlying genetic variation and iii) is associated with the historical climatic conditions of the geographic origin of the poplar clones **(Chapter IV).** I showed that iv) methylation patterns can be partially transmitted to the clonal offspring. Finally, the study on *Fragaria vesca* showed that v) natural patterns of epigenetic variation are species dependent and that environmentally induced variation in the methylome can potentially affect gene expression, thus leading to functional phenotypic variation (**Chapter V**). Altogether, these findings provide further evidence that the methylome can react to environmental cues and that changes in methylation patterns ca be partially transmitted across clonal generations. Thus, if methylation patterns persist over time, they can potentially play a role in plant adaptation processes. My research further advances our knowledge in the field of ecological epigenetics.

Zusammenfassung

Toleranzschwelle von Pflanzen überschreiten und bei vielen Baumarten zu einem Massensterben führen. Außerdem deuten viele Studien darauf hin, dass Bäume aufgrund ihrer Langlebigkeit nicht in der Lage sind, sich schnell genug anzupassen, um mit dem globalen Klimawandel Schritt zu halten. Das Verständnis dafür, wie Bäume auf solche Wetterereignisse und andere Umweltbedingungen (wie biotischen Stress) reagieren, ist daher für Naturschutzmaßnahmen und Waldbewirtschaftungsprogramme von entscheidender Bedeutung. Um mit unvorhersehbaren Umweltbedingungen zurechtzukommen, haben Pflanzen die Fähigkeit entwickelt, ihre Physiologie, Morphologie oder Entwicklung zu verändern, oder anders ausgedrückt, die Fähigkeit, aus einem Genotyp verschiedene Phänotypen hervorzubringen. Diese Fähigkeit wird als phänotypische Plastizität bezeichnet und spielt eine wichtige Rolle bei der Anpassung von Pflanzen. Es wird angenommen, dass drei Faktoren die phänotypische Variabilität und damit potenziell die Widerstandsfähigkeit von Baumpopulationen erhöhen: intraspezifische genetische Variabilität, (Mikro-)Umweltvariation und epigenetische Variation. In der Tat deutet ein wachsender Teil der Literatur darauf hin, dass epigenetische Variation zur lokalen Anpassung natürlicher Pflanzenpopulationen beitragen könnte. Epigenetische Mechanismen, wie z. B. die DNA-Methylierung, können in der Tat Phänotypen als Reaktion auf Umweltveränderungen schnell verändern. Variationen in der DNA-Methylierung können unter genetischer Kontrolle stehen, stochastisch entstehen oder durch Umweltbedingungen induziert werden. Darüber hinaus können durch epigenetische Variationen hervorgerufene phänotypische Veränderungen über mehrere Generationen hinweg vererbt werden (insbesondere über klonale Generationen hinweg), was darauf schließen lässt, dass Variationen in der DNA-Methylierung zu vererbbaren phänotypischen Variationen und schließlich zur Anpassung beitragen könnten. Obwohl in den letzten Jahren beträchtliche Fortschritte erzielt wurden, ist der Zusammenhang zwischen epigenetischer Variation und phänotypischer Variation nach wie vor kaum verstanden.

Mit dieser Arbeit wollte ich diesen Zusammenhang weiter untersuchen und mich mit folgenden Fragen befassen: i) Inwieweit trägt epigenetische Variation zur lokalen Anpassung bei? Genauer gesagt, ii) variiert die DNA-Methylierung als Reaktion auf verschiedene Umwelteinflüsse? iii) Kann die DNA-Methylierungsvariation unabhängig von der zugrunde liegenden genetischen Struktur sein? Und, falls ja, iv) werden DNA-Methylierungsmuster an die klonalen Nachkommen weitergegeben? Kapitel I ist eine allgemeine Einführung in das Gebiet der ökologischen Epigenetik, einschließlich der Frage, wie epigenetische Variation die phänotypische Plastizität beeinflussen kann,

epigenetischer Mechanismen und der Auswirkungen epigenetischer Variation auf einige ökologische Prozesse. In **Kapitel II** befasse ich mich mit der Eignung eines gewöhnlichen Gartenversuchsplans für Studien zur epigenetischen Variation. Ich habe die Sorte *Populus nigra* "Italica" verwendet, um festzustellen, ob phänotypische Unterschiede in funktionellen Merkmalen unter normalen Gartenbedingungen beobachtet werden können. Ich verwendete lineare gemischte Modelle und Klimadaten, um zu untersuchen, ob diese phänotypischen Unterschiede auf den ökologischen Ursprung der Pappelklone zurückgeführt werden können. In **Kapitel III** wird anhand einer Teilmenge der in ganz Europa für den Gemeinschaftsgarten gesammelten Pappelindividuen untersucht, ob die DNA-Methylierung durch biotische (Rost, Herbivorie und Salicylsäure) und abiotische (Hitze, Kälte und Trockenheit) Stressfaktoren beeinflusst wird. In **Kapitel IV** habe ich das EpiDiverse-Toolkit (beschrieben in **Kapitel VI**) und andere statistische Methoden verwendet, um epigenomische Daten zu analysieren und natürliche Methylierungsmuster in Verbindung mit klimatischen Bedingungen auf Landschaftsebene zu untersuchen und zu bewerten, ob Methylierungsprofile auf die klonalen Nachkommen übertragen werden können. In **Kapitel V** schließlich werden epigenomische Daten verwendet, um den Beitrag der klimatischen Bedingungen zur Variation der DNA-Methylierung in Wildpopulationen von *Fragaria vesca* und die möglichen Auswirkungen dieser Variation auf die Genexpression zu untersuchen. Meine Arbeit zeigt, dass i) phänotypische Unterschiede, die mit der geografischen Herkunft zusammenhängen, unter gewöhnlichen Gartenbedingungen trotz der Auswirkungen der Mikroumgebung auf die phänotypische Plastizität beobachtet werden können, was beweist, dass gewöhnliche Gartenexperimente zur Untersuchung der natürlichen epigenetischen Variation geeignet sind (**Kapitel II**). Bei der Pappel scheinen abiotische Stressbedingungen wie Hitze und Trockenheit eine stärkere Reaktion im Methylom hervorzurufen als biotische Stressfaktoren (**Kapitel III**). Darüber hinaus können Methylom-Stressreaktionen global oder stressspezifisch sein und als Epiallele (epigenetische Allele) unter natürlichen Bedingungen fortbestehen. Ich fand auch heraus, dass ii) die Variation im Methylom unabhängig von der zugrunde liegenden genetischen Variation sein kann und iii) mit den historischen klimatischen Bedingungen des geografischen Ursprungs der Pappelklone zusammenhängt (**Kapitel IV**). Ich habe gezeigt, dass iv) Methylierungsmuster teilweise an die klonale Nachkommenschaft weitergegeben werden können. Schließlich hat die Studie an *Fragaria vesca* gezeigt, dass v) natürliche Muster epigenetischer Variation artenabhängig sind und dass umweltbedingte Variationen im Methylom potenziell die Genexpression beeinflussen können, was zu funktionellen phänotypischen Variationen führt (**Kapitel V**). Insgesamt liefern diese Ergebnisse weitere Belege dafür, dass das Methylom auf Umwelteinflüsse

reagieren kann und dass Veränderungen in den Methylierungsmustern teilweise über klonale Generationen hinweg weitergegeben werden können. Wenn Methylierungsmuster also im Laufe der Zeit bestehen bleiben, können sie möglicherweise eine Rolle bei Anpassungsprozessen von Pflanzen spielen. Meine Forschung bringt unser Wissen auf dem Gebiet der ökologischen Epigenetik weiter voran.

CHAPTER I

General introduction

Bárbara Díez Rodríguez

Plant adaptation in a changing world

Plants have colonized practically every habitat on Earth, even some of the most extreme ones, and are constantly exposed to changing environmental conditions. Understanding how plants can adapt to so many different environments has driven endless research across all fields of science. However, over the last decades extreme weather events have been increasing in frequency (IPCC: 'Climate Change 2022: Impacts, Adaptation and Vulnerability). For example, in 2018 Central Europe experienced one of the most severe and long-lasting summer drought and heat wave ever recorded (Schuldt *et al.,* 2020). This heat wave affected ecologically (and economically) important tree species and resulted in unprecedented tree mortality. Furthermore, these extreme events have a major influence on plant populations and communities, because environmental conditions change faster than plants' ability to adapt to them (Cox *et al.,* 2000; Stenseth *et al.,* 2002; Björklund *et al.,* 2009; Anderson *et al.,* 2012). Understanding how plants respond to such weather events and other environmental conditions has thus become crucial for conservation policies and sustainable development. To cope with unpredictable environmental conditions, plants have evolved the ability to alter their physiology, morphology, or development or, in other words, the ability of one genotype to produce different phenotypes (Callahan *et al.,* 1997). This ability is called phenotypic plasticity, and it plays a major role in plant adaptation to their habitat.

Epigenetic variation: a source of phenotypic plasticity.

Over the years, considerable effort has been made to identify and understand the sources of phenotypic variation. We know that the total phenotypic variation observed in a particular trait is the result of genetic variation, environmental variation, and the interaction between genotypes and environments (Scheiner & Goodnight, 1984). Although at first this may seem straightforward, the study of phenotypic variation is incredibly complex, because each layer of variation is in turn composed of several factors, all of them interacting among each other in ways we are only beginning to grasp. Often considered the missing link of heritable variation (Maher, 2008; Danchin, 2013; Banta & Richards, 2018), epigenetic variation is one of these interacting factors.

The term "epigenetics" refers to the study of alterations in gene expression which are not caused by changes in the underlying genomic sequence, and the molecular mechanisms and processes that cause them (Riggs & Porter, 1996; Richards, 2006; Boquete *et al.,* 2021). The first example of epigenetically induced phenotypic variation was originally

described more than 250 years ago by Linnaeus (Linnaeus & Rudberg, 1744). In his work, he described a naturally occurring mutant of *Linaria vulgaris* (common toadflax) in which the symmetry of the flower is changed from bilateral to radial. Centuries later, Cubas and colleagues (Cubas *et al.,* 1999) showed that the occurrence of this mutant was correlated with the DNA methylation levels of the *Lcyc* gene, identifying one of the most famous epialleles. In the mutant, this gene is extensively methylated and thus rendered silent. Furthermore, this epigenetic modification is heritable and co-segregates with the mutant phenotype. Their results offered concrete evidence that epigenetic modifications can affect phenotypic variation, opening a whole new field of research. Since then, several studies have further confirmed that epigenetic variation plays a bigger role in plant adaptation than previously thought. Some other notable examples are the hypermethylation of the Colorless non-ripening (*Cnr*) gene promoter that inhibits fruit ripening in Tomato (Manning *et al.,* 2006), the hypomethylation of a retrotransposon responsible for the mantled abnormality in African oil palm (Ong-Abdullah *et al.,* 2015), and the epialleles (epigenetic alleles) involved in cold tolerance in crofton weed (Xie *et al.,* 2015, but see also Martin *et al.,* 2009; Miura *et al.,* 2009 and Quadrana *et al.,* 2014). Beyond these few clear examples, there is barely any knowledge about the heritable impact of spontaneous or environmentally induced epigenetic variation in plants. A growing body of literature has attempted to describe a causal relationship between epigenetic variation and phenotypic variation, but the current research trend seems to indicate that the few examples previously cited are also notable exceptions. Nonetheless, enormous progress has been made in studying the effects of epigenetic variation on gene expression there is substantial evidence that epigenetic variation can indeed regulate gene expression (Zilberman *et al.,* 2006; Vaughn *et al.,* 2007; Johannes *et al.,* 2009; Schmitz *et al.,* 2013; also reviewed in Bräutigam *et al.,* 2013; Thiebaut *et al.,* 2019; and Liu & Chang, 2021). Not only that, but extensive efforts have tried to understand the role of epigenetic variation in adaptive responses of plant species (Bräutigam *et al.,* 2013; Amaral *et al.,* 2020; Lloyd & Lister, 2022). In this context, if epigenetically induced phenotypic plasticity can persist across generations and be under selection, it can therefore be adaptive.

Epigenetic mechanisms: DNA methylation in context

Epigenetic modifications are the result of several mechanisms that are part of a complex and dynamic network. These mechanisms are controlled by distinct molecular machineries, and work together to regulate genome function and stability (reviewed in

Maeji & Nishimura, 2018; Liu & Chang, 2021). Epigenetic mechanisms include histone post-translational modifications, processes associated with non-coding RNAs, and DNA methylation. Histone modifications are chemical modifications of histone N-terminal tails, such as acetylation, methylation, phosphorylation, and ubiquitination, and generally affect chromatin structure and gene expression (Tessarz & Kouzarides, 2014). RNAs also play an important role in epigenetic regulation and include small RNAs (sRNAs) and long non-coding RNAs (lncRNAs). These RNAs are usually involved in plant development, maintenance of genome integrity and plant responses to biotic and abiotic stresses (Carthew & Sontheimer, 2009; Deng *et al.,* 2016). The most studied and well characterized epigenetic mechanism, however, is DNA methylation (Zemach *et al.,* 2013; Matzke & Mosher, 2014; Heer *et al.,* 2018).

DNA methylation is a base modification in which a methyl (CH_3) group is covalently added to the 5th carbon of a cytosine (resulting in a 5-methylcytosin, or 5mC), one of the four bases forming the DNA molecule (Finnegan *et al.,* 1998; Moore *et al.,* 2012). In plants, methylation of cytosines occurs at three different sequence contexts: CG, CHG and CHH sites (hereafter mCG, mCHG and m CHH), where H = A, T or C but never G. The specific functions of each methylation context in plants are not clear, but we know they are established and maintained by different enzymes and pathways (reviewed in Law & Jacobsen, 2010; He *et al.,* 2013; Zhang *et al.,* 2018). CG and CHG sites are considered symmetrical because there is a mirroring cytosine in the opposite strand (Gruenbaum, 1981), while CHH sites are asymmetrical (Meyer *et al.,* 1994). The symmetrical nature of the CG and CHG contexts is crucial to how methylation in these sites is maintained. In plants, maintenance of methylation in the CG context is done by an enzyme called METHYLTRANSFERASE 1 (MET1). Due to the semiconservative replication of DNA, newly replicated DNA molecules are hemi-methylated. Another enzyme called VARIANT IN METHYLATION 1 (VIM1-5) recognizes these mCG sites and recruits MET1 to methylate the symmetrical site on the opposing strand. This way, mCG is maintained across all cell divisions. Methylation in CHG sites is maintained in a similar fashion by CHROMOMETHYLASE 3 (CMT3). On the other hand, since mCHH is asymmetrical, there is no mirrored cytosine to act as template. Thus, methylation in this context is established *de novo* each time by one of two mechanisms. The first involves another member of the CMT family (CMT2) and the second involves a pathway that is common to all three contexts, RNA-directed DNA methylation (RdDM) (Matzke & mosher, 2014). To make the matters more complex, there is a certain degree of crosstalk between all these mechanisms (He *et al.,* 2011; Niederhuth & Schmitz, 2017; Zhang *et al.,* 2018). As a result of the different mechanisms involved in DNA methylation

maintenance, different sequence contexts differ in their degrees of *mitotic stability*, which are mainly dictated by their symmetry. In the symmetrical contexts, methylation is guided, and thus stably inherited across mitotic divisions (Niederhuth & Schmitz 2014). On the other hand, methylation in the asymmetrical context is maintained mainly by *de novo* establishment and thus less stable across cell divisions (Peter Meyer & Lohuis, 1994)

Furthermore, methylation levels across the genome can vary depending on the feature that is methylated. Although it is increasingly clear that variation in DNA methylation is highly species-dependent (reviewed in Niederhuth *et al.,* 2016) and the molecular function of DNA methylation in some genomic regions remains a mystery, there are some widely recognized patterns. For example, TEs have high levels of DNA methylation in all three sequence contexts, which is involved in silencing TE expression and preventing TE mobilization (Slotkin & Martienssen 2007a; Fultz, Choudury, & Slotkin 2015). CG methylation in promoters is associated with repression of gene expression, and gene body methylation is associated with constitutively expressed genes (reviewed in: Bewick & Schmitz, 2017). The function of mCHH, which is usually associated with actively expressed genes when found in promoters, is not clear, but recent evidence suggests that it might play a regulatory role (Gent *et al.,* 2013; Rajkumar *et al.,* 2020).

In contrast, the effects of DNA methylation in plant physiological processes are better characterized. For example, DNA methylation plays an important role in plant development (e.g. Gehring *et al.,* 2009; Ibarra *et al.,* 2012; reviewed in Bräutigam & Cronk, 2018), vegetative growth (e.g. Candaele *et al.,* 2014; Wang *et al.,* 2016; reviewed in Kumar & Mohapatra, 2021), fruit development (e.g. Zhong *et al.,* 2013) and in several responses to environmental stimuli (Sahu *et al.,* 2013; Zhang *et al.,* 2018; Ashapkin *et al.,* 2020; Parker, Wilkinson and Ton, 2021). Although most studies on epigenetic effects involve short-lived plant species, methylation differences can also affect morphological and physiological processes in tree species (reviewed in Bräutigam *et al.,* 2013). For example, a study by Raj *et al.,* (2011) suggested that there might be a possible epigenetic basis for differences in transcriptomic profiles between poplar trees growing in distinct environments. In addition, white mangrove trees can exhibit striking morphological differences that might be related to a higher epigenetic diversity (Lira-Medeiros *et al.,* 2010). Despite the sometimes-overwhelming amount of literature reviews that summarize past and current knowledge on DNA methylation and its role on plant responses (for a recent one, see Lloyd & Lister, 2022), these processes remain significantly understudied.

Natural patterns of epigenetic variation

In the last couple of decades, epigenetic variation has become of relevance for ecologists. In this context, research is more interested in a) the origins and drivers of this variation, b) patterns of natural epigenetic variation and c) its ecological consequences (Bossdorf *et al.,* 2008; Richards *et al.,* 2017). Variation in DNA methylation can arise stochastically, be under genetic control, or be induced by environmental conditions. Stochastic methylation variants, also called 'spontaneous epimutations', arise spontaneously due to errors in the maintenance of DNA methylation during DNA replication. (Becker *et al.,* 2011; Schmitz *et al.,* 2011; Van der Graaf *et al.,* 2015; Johannes and Schmitz, 2018). Genetically induced methylation variants are caused by genetic variants (Richards, 2006). These genetic variants can induce local (*cis*-acting) or genome-wide (*trans*-acting) epigenetic changes (Schmitz *et al.,* 2013; Dubin *et al.,* 2015; Hagmann *et al.,* 2015; Seymour & Becker, 2017, Galanti *et al.,* 2022). Finally, environmentally induced methylation variants are driven by environmental cues, such as temperature, soil composition or water availability (Lira-Madeiros *et al.,* 2010; Raj *et al.,* 2011; Kawakatsu *et al.,* 2016; Wibowo *et al.,* 2016; Sammarco *et al.,* 2022), or in response to biotic interactions (Dowen *et al.,* 2012; Yu *et al.,* 2013; Martínez *et al.,* 2014; reviewed in Parker *et al.,* 2021)

Plant epigenomes are highly dynamic and vary considerably at different levels, such as within-individuals, among individuals, and among populations (reviewed in Niederhuth *et al.,* 2016; Lloyd & Lister, 2022). Epigenomic variation between cell types and tissues has been documented in a few species, such as the model species *Arabidopsis thaliana* (Gutzat *et al.,* 2020), soybean (Song *et al.,* 2013), rice (Higo *et al.,* 2020) and *Prunus dulcis (*D'Amico-Willman *et al.,* 2022*)*. In these studies, an increase in non-CG methylation levels of the shoot apical meristem in Arabidopsis and rice was observed after the transition from vegetative to reproductive growth, and in Arabidopsis, the columella was found to have a different genome-wide methylome when compared to other root cells. As the cost of sequencing continues to decease, recent research has revealed substantial variation in DNA methylation profiles among different plant species. For example, in a comparative study using 34 angiosperms, Niederhuth and colleagues showed that global methylation mCG levels ranged from ~30.5% *A. thaliana* to 92.5% in *Beta vulgaris* (reviewed in Niederhuth *et al.,* 2016). Methylation levels in CHG varied even more among species, ranging from ~9.3 in *Eutrema salsugineum* to ~81.2% in *B. vulgaris*. mCHH levels, comparatively, were the lowest in all species, but also the most variable, with a range between less than ~2% and ~20%.

However, very little is known about the spatial patterns of epigenetic diversity in natural populations. The study of any kind of variation (phenotypic, genetic or epigenetic), specially at the population or landscape levels, has always been challenged by economical and/or experimental limitations. From a traditional perspective, even studies on genetic diversity at large-scale levels (i.e., over large areas and for many species) are still demanding, given the need for field sampling and the still more or less high costs of sequencing (Taberlet *et al.,* 2012). As a result, empirical work on natural patterns of epigenetic variation has been slowly catching up (Richards *et al.,* 2017). Nevertheless, population epigenetic studies have shown that epigenetic variation can be spatially structured among and within populations (Lira-Medeiros *et al.,* 2010; Richards *et al.,* 2012; Medrano *et al.,* 2014), and that this structure is often associated with environmental and phenotypic variation (Avramidou *et al.,* 2015, Dubin *et al.,* 2015; Nicotra *et al.,* 2015; Gugger *et al.,* 2016; Kawkatsu *et al.,* 2016; Herrera *et al.,* 2017; de Kort *et al.,* 2020; Sammarco *et al.,* 2022; Galanti *et al.,* 2022)

Ecological consequences of epigenetic variation

When including epigenetic diversity into large-scale studies, several layers of complexity are added to the mix. DNA methylation variation can result from genetic control, environmental induction and stochastic epimutations, and can in principle be shaped further by drift and natural selection. In contrast to genetic or genomic patterns, the strength, the effects of epigenetic diversity at a landscape level, and its evolutionary implications are poorly understood. A persistent problem we face in understanding the importance of epigenetics in ecology is that there is a complex relationship between genetic and epigenetic effects. Some advances have been made in this regard by using clonally propagated species and common garden experiments (Bossdorf *et al.,* 2008; Richards *et al.,* 2017, Díez Rodríguez *et al.,* 2022). Despite the limitations, many recent landscape-level studies have investigated the role of epigenetics in intraspecific trait variation and adaptation (Medrano *et al.,* 2014; Dubin *et al.,* 2015; Preite *et al.,* 2015; Foust *et al.,* 2016; Gugger *et al.,* 2016; Herrera *et al.,* 2016; Keller *et al.,* 2016; Alakärpa *et al.,* 2018; Gáspár, Bossdorf & Durka, 2018). These studies focus on the relationship between genetic and epigenetic variation at the landscape level, correlations between environmental variables and epigenetic marks, and correlations between epigenetic marks and plant phenotypic traits (Whipple & Holeski, 2016, Richards *et al.,* 2017).

Although epimutations may arise spontaneously, a significant fraction of all epigenetic variation found within a population has a genetic and environmental basis (Niederhuth *et al.,* 2016; Lloyd & Lister, 2022). It is thus reasonable to assume that epigenetic variation can also influence populations and communities, and processes at the ecosystem or landscape levels. The most commonly documented ecological effects of epigenetic diversity on populations involve the productivity or fitness of the population studied, and they can occur through different mechanisms. In 2010, Bossdorf and colleagues found that experimental alteration of DNA methylation strongly affected growth, fitness, and phenology in *A. thaliana* individuals. In another study, Latzel *et al.,* (2013) suggested that epigenetic variability might affect how different epiRILs ("Epigenetic Recombinant Inbred Lines" Johannes *et al.,* 2009) respond to treatment with salicylic acid and jasmonate, which are the main hormones involved in plant defense responses. In a series of experiments, Fieldes and colleagues showed that demethylation agents affected the fitness and phenological traits of *Linum usitatissimum* (Fieldes 1994; Fieldes & Amyot 1999; Fieldes *et al.,* 2005). Moreover, evidence is growing that shortfalls in genetic diversity can be balanced by epigenetic diversity and facilitate plant population success (Latzel & Klimešová, 2010; Rollins *et al.,* 2013; Verhoeven & Preite, 2014; Dodd & Douhovnikoff, 2016; Mounger *et al.,* 2021). Although several studies have explored the importance of genetic diversity at the community level (Downing *et al.,* 2002; Helm *et al.,* 2009; Taberlet *et al.,* 2012, Lamy *et al.,* 2016), very few studies have tried to assess the contribution of epigenetic variation to ecosystem dynamics in plant communities (Balao, Paun & Alonso, 2017; Herrera, Medrano & Bazaga, 2017; Mounger *et al.,* 2020).

The Lombardy poplar as a model for ecological epigenetics

As mentioned before, one of the challenges of studying the role of epigenetics in plant adaptation and other ecosystem processes, is the tight link between genetic and epigenetic variation. Already in 2006, Richards explained the problem: some epigenetic effects are entirely determined by genotype, other effects may be “facilitated” by specific genotypes, and some effects may be completely independent from genotype. Disentangling these effects is complicated both in experimental design and practical terms, not only because we know little about genetic-epigenetic interactions, but also because the genetic basis of most complex traits is still not well understood. A way to overcome these difficulties is to use asexually reproducing species. Clonal propagation often results in low genetic diversity (Balloux *et al.,* 2003; Rasmussen & Kollmann, 2007;

Meloni *et al.,* 2013; Fu, 2015; Ingvarsson & Dahlberg, 2018), so species that present this method of reproduction can offer a unique system (Verhoeven & Preite, 2014; Vanden Broeck *et al.,* 2018; Shi *et al.,* 2019; Díez Rodríguez *et al.,* 2022). So far, in-depth documentation of intraspecific epigenetic variation has been restricted to model plant species such as *Arabidopsis thaliana*, *Oryza sativa*, and *Zea mays* for which extensive genomic and epigenomic resources exist (Becker *et al.,* 2011; Schmitz *et al.,* 2013; Van Der Graaf *et al.,* 2015; Kawakatsu *et al.,* 2016). However, with high throughput sequencing techniques becoming more affordable and accurate every day, new plant species are being added to the pool for which genomic resources are readily available, increasing research on natural genetic and epigenetic variation in non-model species (Richards *et al.,* 2017).

Poplars (*Populus* sp.), member of the *Salicaceae*, have a wide distribution in the world and can easily be propagated vegetatively. The Lombardy poplar *(Populus nigra* cv. 'Italica*'* Duroi) is probably the most distributed tree clone. The cultivar likely originated in between 1700 and 1720 from one single male mutant tree of *P. nigra* located in central Asia (Elwes & Henry, 1913), from where it was distributed worldwide (Wood, 1994). In the mid-eighteenth century, the Lombardy poplar was spread by cuttings worldwide from Italy, reaching France in 1749, England in 1758, and North America in 1784 (Wood, 1994). It has been widely introduced for use as windbreaks, screens, avenue trees, and landscape plantings all over the temperate regions of the world (in Europe, North and South America, South Africa, Australia, New Zealand, and China) even in subtropical environments where it appears to perform poorly (CABI, 2022). It is assumed that most Lombardy poplars originate from artificial propagation performed by humans, with plant material that has been grown locally for centuries. It is therefore expected, that the large-scale geographic, but artificial expansion of this cultivar may have resulted in the accumulation of lineage-specific, selectively neutral spontaneous epimutations, and in environmental-directed epigenetic effects that are potentially heritable and may have generated different local phenotypes. Furthermore, since the *P. trichocarpa* genome was first assembled in 2006, the amount of genetic, genomic, and biochemical resources available have increased considerably, and *Populus* species have become a model tree species for studies on plant adaptation (Taylor, 2002; Tuskan *et al.,* 2006; Jansson & Douglas, 2007).

Aims of my thesis

In this thesis, I investigated natural patterns of DNA methylation in *Populus nigra* cv 'Italica' (Lombardy poplar) and its effects on phenotypic variation. In **Chapter II**, I investigated the suitability of the cultivar for studies on natural epigenetic variation, focusing on phenotypic differences among individuals sampled along a wide geographical and climatic gradient. I concluded that phenotypic differences can indeed be observed under common garden conditions and are associated with the geographic origin of the individuals. In **Chapter III**, the same cultivar was used to assess the effects of biotic and abiotic stress conditions on methylation patterns. This study provided evidence that environmental cues can induce variation in the poplar methylome. In **Chapter IV,** I studied methylation patterns at a landscape level and the association between methylation and climatic variables. This is the first study that assesses DNA methylation patterns on a clonal tree species along a European-scale climatic gradient. With this system, I aimed to investigate if epigenetic variation that is independent of the underlying genetic structure can contribute to local acclimation and, potentially, local adaptation. **Chapter V** dealt with the effects of transgenerational epigenetic variation on gene expression, using natural populations of a different clonally propagated species (*Fragaria vesca*). And finally, **Chapter VI** describes a pipeline suit of bioinformatic tools for the analysis of bisulfite sequencing data in non-model plant species. The pipeline aims to help researchers with a basic understanding of bioinformatic tools in analyzing epigenomic data.

References

Alakärppä E, Salo HM, Valledor L, Cañal MJ, Häggman H, Vuosku J. **2018**. Natural variation of DNA methylation and gene expression may determine local adaptations of Scots pine populations. *Journal of Experimental Botany* **69**: 5293–5305.

Amaral J, Ribeyre Z, Vigneaud J, Sow MD, Fichot R, Messier C, Pinto G, Nolet P, Maury S. **2020**. Advances and Promises of Epigenetics for Forest Trees. *Forests 2020, Vol. 11, Page 976* **11**: 976.

Anderson JT, Panetta AM, Mitchell-Olds T. **2012**. Evolutionary and Ecological Responses to Anthropogenic Climate ChangeUpdate on Anthropogenic Climate Change. *Plant Physiology* **160**: 1728–1740.

Ashapkin V v., Kutueva LI, Aleksandrushkina NI, Vanyushin BF. **2020**. Epigenetic Mechanisms of Plant Adaptation to Biotic and Abiotic Stresses. *International Journal of Molecular Sciences 2020, Vol. 21, Page 7457* **21**: 7457.

Avramidou E v., Ganopoulos I v., Doulis AG, Tsaftaris AS, Aravanopoulos FA. **2015**. Beyond population genetics: natural epigenetic variation in wild cherry (Prunus avium). *Tree Genetics and Genomes* **11**: 1–9.

Balao F, Paun O, Alonso C. **2018**. Uncovering the contribution of epigenetics to plant phenotypic variation in Mediterranean ecosystems (JD Thompson, Ed.). *Plant Biology* **20**: 38–49.

Balloux F, Lehmann L, de Meeûs T. **2003**. The Population Genetics of Clonal and Partially Clonal Diploids. *Genetics* **164**: 1635–1644.

Banta JA, Richards CL. **2018**. Quantitative epigenetics and evolution. *Heredity 2018 121:3* **121**: 210–224.

Becker C, Hagmann J, Müller J, Koenig D, Stegle O, Borgwardt K, Weigel D. **2011**. Spontaneous epigenetic variation in the Arabidopsis thaliana methylome. *Nature 2011 480:7376* **480**: 245–249.

Bewick AJ, Schmitz RJ. **2017**. Gene body DNA methylation in plants. *Current Opinion in Plant Biology* **36**: 103–110.

Björklund M, Ranta E, Kaitala V, Bach LA, Lundberg P, Stenseth NC. **2009**. Quantitative Trait Evolution and Environmental Change. *PLOS ONE* **4**: e4521.

Booth RE, Grime JP. **2003**. Effects of genetic impoverishment on plant community diversity. *Journal of Ecology* **91**: 721–730.

Boquete MT, Muyle A, Alonso C. **2021**. Plant epigenetics: phenotypic and functional diversity beyond the DNA sequence. *American Journal of Botany* **108**: 553–558.

Bossdorf O, Arcuri D, Richards CL, Pigliucci M. **2010**. Experimental alteration of DNA methylation affects the phenotypic plasticity of ecologically relevant traits in Arabidopsis thaliana. *Evolutionary Ecology* **24**: 541–553.

Bossdorf O, Richards CL, Pigliucci M. **2008**. Epigenetics for ecologists. *Ecology Letters* **11**: 106–115.

Bräutigam K, Cronk Q. **2018**. DNA methylation and the evolution of developmental complexity in plants. *Frontiers in Plant Science* **9**: 1447.

Bräutigam K, Vining KJ, Lafon-Placette C, Fossdal CG, Mirouze M, Marcos JG, Fluch S, Fraga MF, Guevara MÁ, Abarca D, *et al.,* 2013. Epigenetic regulation of adaptive responses of forest tree species to the environment. *Ecology and Evolution* **3**: 399–415.

vanden Broeck A, Cox K, Brys R, Castiglione S, Cicatelli A, Guarino F, Heinze B, Steenackers M, vander Mijnsbrugge K. **2018**. Variability in DNA Methylation and Generational Plasticity in the Lombardy Poplar, a Single Genotype Worldwide Distributed Since the Eighteenth Century. *Frontiers in Plant Science* **9**: 1635.

Callahan HS, Pigliucci M, Schlichting CD. **1997**. Developmental phenotypic plasticity: Where ecology and evolution meet molecular biology. *BioEssays* **19**: 519–525.

Candaele J, Demuynck K, Mosoti D, Beemster GTS, Inzé D, Nelissen H. **2014**. Differential Methylation during Maize Leaf Growth Targets Developmentally Regulated Genes. *Plant Physiology* **164**: 1350–1364.

Carthew RW, Sontheimer EJ. **2009**. Origins and Mechanisms of miRNAs and siRNAs. *Cell* **136**: 642–655.

Climate Change 2022: Impacts, Adaptation and Vulnerability | Climate Change 2022: Impacts, Adaptation and Vulnerability.

Cox PM, Betts RA, Jones CD, Spall SA, Totterdell IJ. **2000**. Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature 2000 408:6809* **408**: 184–187.

Cubas P, Vincent C, Coen E. **1999**. An epigenetic mutation responsible for natural variation in floral symmetry. *Nature* **401**: 157–161.

D'Amico-Willman KM, Ouma WZ, Meulia T, Sideli GM, Gradziel TM, Fresnedo-Ramírez J. **2022**. Whole-genome sequence and methylome profiling of the almond [Prunus dulcis (Mill.) D.A. Webb] cultivar 'Nonpareil'. *G3 Genes|Genomes|Genetics* **12**.

Danchin É. **2013**. Avatars of information: towards an inclusive evolutionary synthesis. *Trends in Ecology & Evolution* **28**: 351–358.

Deng X, Song X, Wei L, Liu C, Cao X. **2016**. Epigenetic regulation and epigenomic landscape in rice. *National Science Review* **3**: 309–327.

Díez Rodríguez B, Peña C, Pérez-Bello P, Bette J, Lerbs L, Mackenbach T, Wulle S, de Paoli E, Verhoeven KJF, Heer K, *et al.,* 2022. An Uncommon Garden Experiment: Microenvironment Has Stronger Influence on Phenotypic Variation than Epigenetic Memory in the Clonal Lombardy Poplar. *SSRN Electronic Journal*.

Dodd RS, Douhovnikoff V. **2016**. Adjusting to Global change through clonal growth and epigenetic variation. *Frontiers in Ecology and Evolution* **4**: 86.

Dowen RH, Pelizzola M, Schmitz RJ, Lister R, Dowen JM, Nery JR, Dixon JE, Ecker JR. **2012**. Widespread dynamic DNA methylation in response to biotic stress. *Proceedings of the National Academy of Sciences of the United States of America* **109**: E2183–E2191.

Downing AL, Leibold MA. **2002**. Ecosystem consequences of species richness and composition in pond food webs. *Nature* **416**: 837–841.

Dubin MJ, Zhang P, Meng D, Remigereau MS, Osborne EJ, Casale FP, Drewe P, Kahles A, Jean G, Vilhjálmsson B, *et al.,* 2015. DNA methylation in Arabidopsis has a genetic basis and shows evidence of local adaptation. *eLife* **4**.

Elwes HJ, Henry A. **1913**. The trees of Great Britain & Ireland. R. & R. Clark, Limited, Edinburgh, 2022 pp.

Eriksson ME, Hoffman D, Kaduk M, Mauriat M, Moritz T. **2015**. Transgenic hybrid aspen trees with increased gibberellin (GA) concentrations suggest that GA acts in parallel with FLOWERING LOCUS T2 to control shoot elongation. *New Phytologist* **205**: 1288–1295.

Escalante-Pérez M, Jaborsky M, Lautner S, Fromm J, Müller T, Dittrich M, Kunert M, Boland W, Hedrich R, Ache P. **2012**. Poplar extrafloral nectaries: two types, two strategies of indirect defenses against herbivores. *Plant physiology* **159**: 1176–91.

Fieldes MA. **1994**. Heritable effects of 5-azacytidine treatments on the growth and development of flax (*Linum usitatissimum*) genotrophs and genotypes. *Genome* **37**: 1–11.

Fieldes MA, Amyot L. **1999**. Epigenetic control of early flowering in flax lines induced by 5-azacytidine applied to germinating seed. *Journal of Heredity* **90**: 199–206.

Fieldes MA, Schaeffer SM, Krech MJ, Brown JCL. **2005**. DNA hypomethylation in 5-azacytidine-induced early-flowering lines of flax. *Theoretical and Applied Genetics* **111**: 136–149.

Fisher RA. **1958**. *The genetical theory of natural selection*. Рипол Классик.

Fu YB. **2015**. Understanding crop genetic diversity under modern plant breeding. *Theoretical and Applied Genetics* **128**: 2131–2142.

Fultz D, Choudury SG, Slotkin RK. **2015**. Silencing of active transposable elements in plants. *Current Opinion in Plant Biology* **27**: 67–76.

Galantiid D, Ramos-Cruzid D, Nunnid A, Rodríguez-Aré Valoid I, Scheepensid JF, Beckerid C, Bossdorfid O. **2022**. Genetic and environmental drivers of large-scale epigenetic variation in Thlaspi arvense (NM Springer, Ed.). *PLOS Genetics* **18**: e1010452.

Gáspár B, Bossdorf O, Durka W. **2019**. Structure, stability and ecological significance of natural epigenetic variation: a large-scale survey in Plantago lanceolata. *New Phytologist* **221**: 1585–1596.

Gehring M, Bubb KL, Henikoff S. **2009**. Extensive demethylation of repetitive elements during seed development underlies gene imprinting. *Science* **324**: 1447–1451.

Gent JI, Ellis NA, Guo L, Harkess AE, Yao Y, Zhang X, Dawe RK. **2013**. CHH islands: de novo DNA methylation in near-gene chromatin regulation in maize. *Genome Research* **23**: 628–637.

van der Graaf A, Wardenaara R, Neumann DA, Taudt A, Shaw RG, Jansen RC, Schmitz RJ, Colomé-Tatché M, Johannes F. **2015**. Rate, spectrum, and evolutionary dynamics of spontaneous epimutations. *Proceedings of the National Academy of Sciences of the United States of America* **112**: 6676–6681.

Gruenbaum Y, Naveh-many T, Cedar H, Razin A. **1981**. Sequence specificity of methylation in higher plant DNA. *Nature 1981 292:5826* **292**: 860–862.

Gugger PF, Fitz-Gibbon S, Pellegrini M, Sork VL. **2016**. Species-wide patterns of DNA methylation variation in Quercus lobata and their association with climate gradients. *Molecular Ecology* **25**: 1665–1680.

Gutzat R, Rembart K, Nussbaumer T, Hofmann F, Pisupati R, Bradamante G, Daubel N, Gaidora A, Lettner N, Donà M, *et al.,* 2020. Arabidopsis shoot stem cells display dynamic transcription and DNA methylation patterns. *The EMBO Journal* **39**: e103667.

Hagmann J, Becker C, Müller J, Stegle O, Meyer RC, Wang G, Schneeberger K, Fitz J, Altmann T, Bergelson J, *et al.,* 2015. Century-scale Methylome Stability in a Recently Diverged Arabidopsis thaliana Lineage. *PLOS Genetics* **11**: e1004920.

Hannan Parker A, Wilkinson SW, Ton J. **2022**. Epigenetics: a catalyst of plant immunity against pathogens. *New Phytologist* **233**: 66–83.

He XJ, Chen T, Zhu JK. **2011**. Regulation and function of DNA methylation in plants and animals. *Cell Research 2011 21:3* **21**: 442–465.

Heer K, Mounger J, Boquete MT, Richards CL, Opgenoorth L. **2018**. The diversifying field of plant epigenetics. *New Phytologist* **217**: 988–992.

Helm A, Oja T, Saar L, Takkis K, Talve T, Pärtel M. **2009**. Human influence lowers plant genetic diversity in communities with extinction debt. *Journal of Ecology* **97**: 1329–1336.

Herrera CM, Medrano M, Bazaga P. **2016**. Comparative spatial genetics and epigenetics of plant populations: heuristic value and a proof of concept. *Molecular Ecology* **25**: 1653–1664.

Herrera CM, Medrano M, Bazaga P. **2017**. Comparative epigenetic and genetic spatial structure of the perennial herb *Helleborus foetidus* : Isolation by environment, isolation by distance, and functional trait divergence. *American Journal of Botany* **104**: 1195–1204.

Higo A, Saihara N, Miura F, Higashi Y, Yamada M, Tamaki S, Ito T, Tarutani Y, Sakamoto T, Fujiwara M, *et al.,* 2020. DNA methylation is reconfigured at the onset of reproduction in rice shoot apical meristem. *Nature Communications 2020 11:1* **11**: 1–12.

Hughes AR, Inouye BD, Johnson MTJ, Underwood N, Vellend M. **2008**. Ecological consequences of genetic diversity. *Ecology Letters* **11**: 609–623.

Hughes AR, Stachowicz JJ. **2004**. Genetic diversity enhances the resistance of a seagrass ecosystem to disturbance. *Proceedings of the National Academy of Sciences* **101**: 8998–9002.

Ibarra CA, Feng X, Schoft VK, Hsieh TF, Uzawa R, Rodrigues JA, Zemach A, Chumak N, Machlicova A, Nishimura T, *et al.,* 2012. Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. *Science* **337**: 1360–1364.

Ingvarsson PK, Dahlberg H. **2018**. The effects of clonal forestry on genetic diversity in wild and domesticated stands of forest trees. *https://doi.org/10.1080/02827581.2018.1469665* **34**: 370–379.

Jansson S, Douglas CJ. **2007**. *Populus*: A Model System for Plant Biology. *Annual Review of Plant Biology* **58**: 435–458.

Johannes F, Porcher E, Teixeira FK, Saliba-Colombani V, Simon M, Agier N, Bulski A, Albuisson J, Heredia F, Audigier P, *et al.,* 2009. Assessing the Impact of Transgenerational Epigenetic Variation on Complex Traits. *PLOS Genetics* **5**: e1000530.

Johannes F, Schmitz RJ. **2018**. Spontaneous epimutations in plants. *New Phytologist*.

Johnson MTJ, Lajeunesse MJ, Agrawal AA. **2006**. Additive and interactive effects of plant genotypic diversity on arthropod communities and plant fitness. *Ecology Letters* **0**: 051012084514001.

Kagiya S, Yasugi M, Kudoh H, Nagano AJ, Utsumi S. **2018**. Does genomic variation in a foundation species predict arthropod community structure in a riparian forest? *Molecular Ecology* **27**: 1284–1295.

Kawakatsu T, Huang S shan C, Jupe F, Sasaki E, Schmitz RJJ, Urich MAA, Castanon R, Nery JRR, Barragan C, He Y, *et al.,* 2016. Epigenomic Diversity in a Global Collection of Arabidopsis thaliana Accessions. *Cell* **166**: 492–505.

Keller TE, Lasky JR, Yi S v. 2016. The multivariate association between genomewide DNA methylation and climate across the range of *Arabidopsis thaliana*. *Molecular Ecology* **25**: 1823–1837.

de Kort H, Panis B, Deforce D, van Nieuwerburgh F, Honnay O. **2020**. Ecological divergence of wild strawberry DNA methylation patterns at distinct spatial scales. *Molecular Ecology* **29**: 4871–4881.

Kumar S, Mohapatra T. **2021**. Dynamics of DNA Methylation and Its Functions in Plant Growth and Development. *Frontiers in Plant Science* **12**: 858.

Lamy T, Laroche F, David P, Massol F, Jarne P. **2017**. The contribution of species-genetic diversity correlations to the understanding of community assembly rules. *Oikos* **126**: 759–771.

Latzel V, Allan E, Bortolini Silveira A, Colot V, Fischer M, Bossdorf O. **2013**. Epigenetic diversity increases the productivity and stability of plant populations. *Nature Communications* **4**: 2875.

Latzel V, Klimešová J. **2010**. Transgenerational plasticity in clonal plants. *Evolutionary Ecology* **24**: 1537–1543.

Linnaeus C, Rudberg D. **1744**. Dissertatio botanica de Peloria. *Amoenitates academicae* **1**: 280–297.

Lira-Medeiros CF, Parisod C, Fernandes RA, Mata CS, Cardoso MA, Ferreira PCG. **2010**. Epigenetic Variation in Mangrove Plants Occurring in Contrasting Natural Environment. *PLOS ONE* **5**: e10326.

Liu J, Chang C. **2021**. Concerto on Chromatin: Interplays of Different Epigenetic Mechanisms in Plant Development and Environmental Adaptation. *Plants 2021, Vol. 10, Page 2766* **10**: 2766.

Lloyd JPB, Lister R. **2021**. Epigenome plasticity in plants. *Nature Reviews Genetics 2021 23:1* **23**: 55–68.

Madritch M, Donaldson JR, Lindroth RL. **2006**. Genetic identity of *Populus* tremuloides litter influences decomposition and nutrient release in a mixed forest stand. *Ecosystems* **9**: 528–537.

Maeji H, Nishimura T. **2018**. Epigenetic Mechanisms in Plants. *Advances in Botanical Research* **88**: 21–47.

Maher B. **2008**. Personal genomes: The case of the missing heritability. *Nature* **456**: 18–21.

Manning K, Tör M, Poole M, Hong Y, Thompson AJ, King GJ, Giovannoni JJ, Seymour GB. **2006**. A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nature Genetics 2006 38:8* **38**: 948–952.

Martin A, Troadec C, Boualem A, Rajab M, Fernandez R, Morin H, Pitrat M, Dogimont C, Bendahmane A. **2009**. A transposon-induced epigenetic change leads to sex determination in melon. *Nature 2009 461:7267* **461**: 1135–1138.

Martinez G, Castellano M, Tortosa M, Pallas V, Gomez G. **2014**. A pathogenic non-coding RNA induces changes in dynamic DNA methylation of ribosomal RNA genes in host plants. *Nucleic Acids Research* **42**: 1553–1562.

Matzke MA, Mosher RA. **2014**. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nature Reviews Genetics 2014 15:6* **15**: 394–408.

Medrano M, Herrera CM, Bazaga P. **2014**. Epigenetic variation predicts regional and local intraspecific functional diversity in a perennial herb. *Molecular Ecology* **23**: 4926–4938.

Meloni M, Reid A, Caujapé-Castells J, Marrero Á, Fernández-Palacios JM, Mesa-Coelo RA, Conti E. **2013**. Effects of clonality on the genetic variability of rare, insular species: the case of Ruta microcarpa from the Canary Islands. *Ecology and Evolution* **3**: 1569–1579.

Miura K, Agetsuma M, Kitano H, Yoshimura A, Matsuoka M, Jacobsen SE, Ashikari M. **2009**. A metastable DWARF1 epigenetic mutant affecting plant stature in rice. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 11218–11223.

Moore LD, Le T, Fan G. **2012**. DNA Methylation and Its Basic Function. *Neuropsychopharmacology 2013 38:1* **38**: 23–38.

Mounger J, Ainouche M, Bossdorf O, Cavé-Radet A, Li B, Parepa M, Salmon A, Yang J, Richards C. **2020**. Epigenetics and the success of invasive plants.

Nicotra AB, Segal DL, Hoyle GL, Schrey AW, Verhoeven KJF, Richards CL. **2015**. Adaptive plasticity and epigenetic variation in response to warming in an Alpine plant. *Ecology and Evolution* **5**: 634–647.

Niederhuth CE, Bewick AJ, Ji L, Alabady MS, Kim K do, Li Q, Rohr NA, Rambani A, Burke JM, Udall JA, *et al.,* 2016. Widespread natural variation of DNA methylation within angiosperms. *Genome Biology* **17**.

Niederhuth CE, Schmitz RJ. **2014**. Covering Your Bases: Inheritance of DNA Methylation in Plant Genomes. *Molecular Plant* **7**: 472–480.

Niederhuth CE, Schmitz RJ. **2017**. Putting DNA methylation in context: from genomes to gene expression in plants. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1860**: 149–156.

Ong-Abdullah M, Ordway JM, Jiang N, Ooi SE, Kok SY, Sarpan N, Azimi N, Hashim AT, Ishak Z, Rosli SK, *et al.,* 2015. Loss of Karma transposon methylation underlies the mantled somaclonal variant of oil palm. *Nature 2015 525:7570* **525**: 533–537.

Peter Meyer IN, Lohuis M Ten. **1994**. Evidence for cytosine methylation of non-symmetrical sequences in transgenic Petunia hybrida. *The EMBO journal* **13**: 2084–2088.

Preite V, Snoek LB, Oplaat C, Biere A, van der Putten WH, Verhoeven KJF. **2015**. The epigenetic footprint of poleward range-expanding plants in apomictic dandelions. *Molecular Ecology* **24**: 4406–4418.

Quadrana L, Almeida J, Asís R, Duffy T, Dominguez PG, Bermúdez L, Conti G, Corrêa Da Silva J v., Peralta IE, Colot V, *et al.,* 2014. Natural occurring epialleles determine vitamin E accumulation in tomato fruits. *Nature Communications 2014 5:1* **5**: 1–11.

Raj S, Bräutigam K, Hamanishi ET, Wilkins O, Thomas BR, Schroeder W, Mansfield SD, Plant AL, Campbell MM. **2011**. Clone history shapes *Populus* drought responses. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 12521–12526.

Rajkumar MS, Gupta K, Khemka NK, Garg R, Jain M. **2020**. DNA methylation reprogramming during seed development and its functional relevance in seed size/weight determination in chickpea. *Communications Biology 2020 3:1* **3**: 1–13.

Rasmussen KK, Kollmann J. **2007**. Low genetic diversity in small peripheral populations of a rare European tree (Sorbus torminalis) dominated by clonal reproduction. *Conservation Genetics 2007 9:6* **9**: 1533–1539.

Reusch TBH, Ehlers A, Hämmerli A, Worm B. **2005**. Ecosystem recovery after climatic extremes enhanced by genotypic diversity. *Proceedings of the National Academy of Sciences* **102**: 2826–2831.

Richards EJ. **2006**. Inherited epigenetic variation — revisiting soft inheritance. *Nature Reviews Genetics 2006 7:5* **7**: 395–401.

Richards CL, Alonso C, Becker C, Bossdorf O, Bucher E, Colomé-Tatché M, Durka W, Engelhardt J, Gaspar B, Gogol-Döring A, *et al.,* 2017. Ecological plant epigenetics: Evidence from model and non-model species, and the way forward. *Ecology Letters* **20**: 1576–1590.

Richards CL, Schrey AW, Pigliucci M. **2012**. Invasion of diverse habitats by few Japanese knotweed genotypes is correlated with epigenetic differentiation (M Vellend, Ed.). *Ecology Letters* **15**: 1016–1025.

Riggs AD, Porter TN. **1996**. Overview of epigenetic mechanisms. *Cold Spring Harbor monograph series* **32**: 29–46.

Sahu PP, Pandey G, Sharma N, Puranik S, Muthamilarasan M, Prasad M. **2013**. Epigenetic mechanisms of plant stress responses and adaptation. *Plant Cell Reports* **32**: 1151–1159.

Sammarco I, Münzbergová Z, Latzel V. **2022**. DNA Methylation Can Mediate Local Adaptation and Response to Climate Change in the Clonal Plant *Fragaria vesca*: Evidence From a European-Scale Reciprocal Transplant Experiment. *Frontiers in Plant Science* **13**: 435.

Scheiner SM, Goodnight CJ. **1984**. The Comparison of Phenotypic Plasticity and Genetic Variation in Populations of the Grass Danthonia spicata. *Evolution* **38**: 845.

Schmitz RJ, Schultz MD, Lewsey MG, O'Malley RC, Urich MA, Libiger O, Schork NJ, Ecker JR. **2011**. Transgenerational epigenetic instability is a source of novel methylation variants. *Science* **334**: 369–373.

Schmitz RJ, Schultz MD, Urich MA, Nery JR, Pelizzola M, Libiger O, Alix A, McCosh RB, Chen H, Schork NJ, *et al.,* 2013. Patterns of population epigenomic diversity. *Nature 2013 495:7440* **495**: 193–198.

Schweitzer JA, Bailey JK, Hart SC, Whitham TG. **2005**. Nonadditive effects of mixing cottonwood genotypes on litter decomposition and nutrient dynamics. *Ecology* **86**: 2834–2840.

Seymour DK, Becker C. **2017**. The causes and consequences of DNA methylome variation in plants. *Current Opinion in Plant Biology* **36**: 56–63.

Shi W, Chen X, Gao L, Xu CY, Ou X, Bossdorf O, Yang J, Geng Y. **2019**. Transient stability of epigenetic population differentiation in a clonal invader. *Frontiers in Plant Science* **9**: 1851.

Slotkin RK. **2016**. Plant epigenetics: from genotype to phenotype and back again. *Genome Biology* **17**: 57.

Slotkin RK, Martienssen R. **2007**. Transposable elements and the epigenetic regulation of the genome. *Nature Reviews Genetics 2007 8:4* **8**: 272–285.

Song QX, Xiang L, Li QT, Chen H, Hu XY, Ma B, Zhang WK, Chen SY, Zhang JS. **2013**. Genome-Wide Analysis of DNA Methylation in Soybean. *Molecular Plant* **6**: 1961–1974.

Stenseth NC, Mysterud A, Ottersen G, Hurrell JW, Chan KS, Lima M. **2002**. Ecological effects of climate fluctuations. *Science* **297**: 1292–1296.

Taberlet P, Zimmermann NE, Englisch T, Tribsch A, Holderegger R, Alvarez N, Niklfeld H, Coldea G, Mirek Z, Moilanen A, *et al.,* 2012. Genetic diversity in widespread species is not congruent with species richness in alpine plant communities. *Ecology Letters* **15**: 1439–1448.

TAYLOR G. **2002**. *Populus*: Arabidopsis for Forestry. Do We Need a Model Tree? *Annals of Botany* **90**: 681–689.

Tessarz P, Kouzarides T. **2014**. Histone core modifications regulating nucleosome structure and dynamics. *Nature Reviews Molecular Cell Biology 2014 15:11* **15**: 703–708.

Thiebaut F, Hemerly AS, Ferreira PCG. **2019**. A role for epigenetic regulation in the adaptation and stress responses of non-model plants. *Frontiers in Plant Science* **10**: 246.

Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam M, Ralph S, Rombauts S, Salamov A, *et al.,* 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**: 1596–1604.

Vaughn MW, Tanurdžić M, Lippman Z, Jiang H, Carrasquillo R, Rabinowicz PD, Dedhia N, McCombie WR, Agier N, Bulski A, *et al.,* 2007. Epigenetic Natural Variation in Arabidopsis thaliana (JC Carrington, Ed.). *PLoS Biology* **5**: e174.

Verhoeven KJF, Preite V. **2014**. Epigenetic variation in asexxually reproducing organisms. *Evolution* **68**: 644–655.

Wang Y, Xue X, Zhu JK, Dong J. **2016**. Demethylation of ERECTA receptor genes by IBM1 histone demethylase affects stomatal development. *Development (Cambridge)* **143**: 4452–4461.

Whipple A v., Holeski LM. **2016**. Epigenetic inheritance across the landscape. *Frontiers in Genetics* **7**: 189.

Wibowo A, Becker C, Marconi G, Durr J, Price J, Hagmann J, Papareddy R, Putra H, Kageyama J, Becker J, *et al.,* 2016. Hyperosmotic stress memory in arabidopsis is mediated by distinct epigenetically labile sites in the genome and is restricted in the male germline by dna glycosylase activity. *eLife* **5**.

Xie HJ, Li H, Liu D, Dai WM, He JY, Lin S, Duan H, Liu LL, Chen SG, Song XL, *et al.,* 2015. ICE1 demethylation drives the range expansion of a plant invader through cold tolerance divergence. *Molecular Ecology* **24**: 835–850.

Yu A, Lepère G, Jay F, Wang J, Bapaume L, Wang Y, Abraham AL, Penterman J, Fischer RL, Voinnet O, *et al.,* 2013. Dynamics and biological relevance of DNA demethylation in Arabidopsis antibacterial defense. *Proceedings of the National Academy of Sciences of the United States of America* **110**: 2389–2394.

Zemach A, McDaniel IE, Silva P, Zilberman D. **2010**. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* **328**: 916–919.

Zhang H, Lang Z, Zhu JK. **2018**. Dynamics and function of DNA methylation in plants. *Nature Reviews Molecular Cell Biology 2018 19:8* **19**: 489–506.

Zhong S, Fei Z, Chen YR, Zheng Y, Huang M, Vrebalov J, McQuinn R, Gapper N, Liu B, Xiang J, *et al.,* 2013. Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. *Nature Biotechnology 2013 31:2* **31**: 154–159.

Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S. **2006**. Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. *Nature Genetics 2006 39:1* **39**: 61–69.

CHAPTER II

Using common gardens to study natural epigenetic variation

Bárbara Díez Rodríguez, Cristian Peña Pontón, Paloma Pérez Bello, Julius Bette, Lena Lerbs, Tabea Mackenbach, Sven Wulle, Emanuele De Paoli, Koen J.F. Verhoeven, Katrin Heer and Lars Opgenoorth

This chapter is in preparation for submission to *Scientific Reports*

Abstract

Environmental changes can trigger phenotypic variation in plants through epigenetic mechanisms, but strong genetic influences on epigenetic variation and phenotypes make it difficult to isolate and study these effects. In this study, we aimed to investigate the effect of previous environments on the phenotypic variation of the next clonal generation of Lombardy poplar ramets growing under common garden conditions. We studied phenotypic plasticity using the Lombardy poplar *(Populus nigra* cv. 'Italica' Duroi), a globally distributed clonal tree. We surveyed 14 functional traits related to tree growth, ecophysiological and phenological processes in poplar ramets collected along a wide geographical range in Europe and planted under common garden conditions. We investigated whether phenotypic variation was related to geography and historical bioclimatic data of the ramets' sites of origin using linear mixed effect models. We found significant differences among ramets from different geographic origins in tree height, number of stems per ramet, and duration of bud flush. However, microenvironmental variation in the common garden, captured via block effects, had an even bigger impact on phenotypic variation than the environmental conditions at the sites of origin. Our results show that phenotypic variation in the ramets might be associated to the climate origin from different climates, suggesting possible epigenetic memory. However, such legacy effects might be quickly outweighed by new environmental conditions.

Key words: *Populus nigra* cv. 'Italica' Duroi, Phenotypic variation, parental effects, epigenetic memory, microenvironment, common garden

Introduction

Recent climatic extremes have shown that climate change already has severe impacts on temperate tree populations (Vitasse *et al.,* 2019; Schuldt *et al.,* 2020) and many studies suggest that due to their longevity, trees are not able to adapt rapidly enough to keep pace with global climate change (Aitken *et al.,* 2008; Bisbing *et al.,* 2021). However, three factors have been suggested that might increase phenotypic variability and thus potentially the resilience of tree populations: intraspecific genetic variability (Benito Garzón *et al.,* 2011; Pfenninger *et al.,* 2021), (micro-)environmental variation (Slavov *et al.,* 2010; Sork *et al.,* 2013; Scotti *et al.,* 2016), and epigenetic acclimation (Guarino *et al.,* 2015; Richards *et al.,* 2017a; Sow *et al.,* 2020). For example, several studies on various *Populus sp*. genotypes have shown that patterns of phenotypic variation observed under common garden conditions usually follow latitudinal clines (Howe *et al.,* 2000; Luquez *et al.,* 2008; Ma *et al.,* 2010; Rohde *et al.,* 2011a; McKown *et al.,* 2014). Phenotypic variation associated to clinal gradients has been observed in multiple functional traits, such as stomatal density(Gornall & Guy, 2007) , specific leaf area(de Frenne *et al.,* 2013), herbivory damage and herbivore abundance (Schemske *et al.,* 2009; Robinson *et al.,* 2012), stem height, relative growth rate, and total phenolic content (Luquez *et al.,* 2008). Furthermore, many of these traits tend to show high heritability values in *Populus* species, suggesting that phenotypic differences have a major genetic component (Howe *et al.,* 2000; McKown *et al.,* 2014). Several factors that act as selective forces can contribute to this clinal variation, including temperature, precipitation, soil nutrient availability, and biotic agents(de Frenne *et al.,* 2013). However, all these studies dealt with genetically diverse (source) populations and since this genetic variability interacts with or might overshadow the impact of the microenvironmental and epigenetic marks on the phenotype, it is difficult to disentangle them in long lived organisms and thus quantify their relevance in structuring phenotypic variability. Here, we collected clonal Lombardy Poplar individuals across Europe and transferred them to a common garden environment. The genetic uniformity of this clone allows pinpointing epigenetic effects on phenotypic variation which can be induced by microenvironmental and large-scale environmental variation.

Epigenetic modifications are chemical modifications in the DNA (for example, DNA methylation) that influence gene expression or function without altering the underlying DNA sequence (Bossdorf *et al.,* 2008). These modifications can arise spontaneously or

can be triggered by environmental conditions and might be transmitted to the offspring (Latzel & Klimešová, 2010a; Verhoeven & Preite, 2014a; Münzbergová & Hadincová, 2016). If environment induces epigenetic marks, epigenetic variation can lead to "epigenetic memory" (Latzel *et al.,* 2016a; Dodd & Douhovnikoff, 2016a). Clones thus offer a unique system to study epigenetically mediated plasticity(Latzel & Klimešová, 2010a; Richards *et al.,* 2017a; Heer *et al.,* 2018). First, because clones are characterized by low to zero genetic diversity, effects of epigenetic variation on trait variation will not be confounded by effects of genetic variation. Second, since clonal reproduction circumvents the epigenetic resetting associated with meiosis, epigenetic marks might be stable between clonal generations(Verhoeven & Preite, 2014a).
Within single poplar genotypes, in absence of genetic effects on trait variation, phenotypic variation among plants can also arise based on transmission of parental environmental effects, potentially mediated by epigenetic mechanisms(Raj *et al.,* 2011a). In our study, we worked with the so-called Lombardy poplar *(Populus nigra* cv. 'Italica*'* Duroi) which is probably the widest distributed tree clone globally (CABI, 2022). The cultivar likely originated in the 18th century from one single male mutant tree of *P. nigra* located in central Asia (Elwes & Henry, 1913) and its cuttings were introduced to Italy, from where its cuttings were distributed worldwide for ornamental purposes and as a source of timber (Wood, 1994). It is assumed that almost all Lombardy poplars are the result of artificial propagation performed by humans. This unique origin and the wide geographical distribution of the clone makes the cultivar a perfect study system to investigate epigenetically induced phenotypic variation in a long-lived plant species and its potential role in plant adaptation (vanden Broeck *et al.,* 2018). Despite the obvious advantages of this system, to date only one other study has used this cultivar as a model species. In 2018, Vanden Broeck and colleagues reported that a significant fraction of epigenetic variation in Lombardy poplar clones was distributed among the countries of origin of the clones and suggested that, in the Lombardy poplar, epigenetic marks might contribute to phenotypic differences in the timing of bud set and can be transferred to asexually reproducing offspring.

In this study, we aimed to investigate the effect of previous environments on the phenotypic variation of the next clonal generation of Lombardy poplar ramets growing under common garden conditions. We surveyed 14 functional traits related to tree growth, ecophysiological and phenological processes under common garden conditions in one to two growing seasons. We genotyped all ramets established in the garden to determine clonal identity. Using historical bioclimatic data from each region where the ramets were collected, we related phenotypic variation to geographic and climatic

gradients. We hypothesized that (1) phenotypic variation in functional traits would correlate with macro-climatic gradients from the ramets sites of origin and (2) that microenvironmental differences in the common garden would not have any effect on the clone phenotypes.

Material and methods

Plant material and common garden design

In early spring 2018, cuttings from *Populus nigra cv* 'Italica' clones were collected in Europe across geographical gradients that spanned from 41° to 60° N and -5° to 25° E approximately (Figure 1A). Twelve sampling sites were selected that covered seven different Köppen-Geiger climate subtypes (Peel *et al.,* 2007). At each site, cuttings of approximately 30 cm in length were sampled from 50 to 56 different individuals within a 25 km radius, except for the sites in Spain, Poland, and Lithuania, where only 24, 27, and 12 individual trees were found within the sampling radius, respectively. Source trees (ortets hereafter) were tagged and georeferenced. During the first week of May 2018, the cuttings (ramets hereafter) were planted on a lawn in the Marburg Botanical Garden (Germany). The common garden is located at 50º 48' 02.7" N, 8º 48' 24.8" E, at an elevation of 328 m within the Cfb (temperate oceanic climate) of the Köppen-Geiger climate classification. The trial area was covered with water-permeable plastic foil (Agrolys BL100 #25/12.5, Beaulieu Technical Textiles) with openings of about 10 cm x 10 cm placed over the ramets. Weeds that grew through the openings in the foil were removed manually. The common garden area was not shaded in any way, allowing the ramets to grow under direct sunlight. No herbicides, pesticides, or fertilizers were used. The area was fenced to exclude herbivory or other disturbances by deer and wild boars. The ramets were planted in a random block design (Figure S1) composed of 12 blocks in a 3 by 4 array with 40-45 cuttings per block. Two cuttings from 1-5 individual ortets from each sampling site were planted per block. After one of the ramets had successfully established, the other was removed from the garden. The ramets were planted with 1 m between trees and were watered frequently until the end of summer. In total, 549 poplar ramets were planted, of which 433 survived the summer of 2018.

Genotyping of ramets

To determine whether the ramets really belonged to a single clone line, leaf samples were collected, and genomic DNA was extracted between July and August 2018. DNA

samples were isolated with the PeqGOLD Plant DNA mini kit (PEQLAB Biotechnologie GmbH, Erlangen, Germany). The clones were genotyped at 5000 genomic loci, equally distributed across the 19 *P. nigra* chromosomes, and selected among a larger set of polymorphic sites identified in Scaglione *et al.,* (2019). Sequencing was carried out using the Allegro Targeted Genotyping protocol from NuGEN Technologies (TECAN) by IGATech (Udine, Italy). Three adult *P. nigra* cv 'Italica' clones from the botanical garden in Marburg and two *P. nigra* genotypes (described in Faivre-Rampant *et al.,* 2016) were included in the sequencing design as controls.

Trait measurement in the common garden

To study phenotypic variation across geographic and climatic gradients, 14 phenotypic traits were measured under common garden conditions. These were divided into five categories: tree growth, ecophysiology, biotic stress damage, leaf chemistry, and phenology. A short description of all traits can be found in Supplementary table 2.

1. Tree growth traits

Ramet diameters were measured with an electronic calliper before the ramets were planted in 2018. Tree height was measured at the ground level at the end of the growing season in 2018, 2019, and 2020, after all ramets had completed bud set. The number of stems was counted at the end of the 2018 growing period and corresponds to the number of initial sprouts that appeared in the ramets in the summer. The active growth rate was calculated as the ratio between the height gain (2018 to 2019) and the number of days of growth. The growth period was calculated as the number of days between the last stage of bud flush (Stage 5 in Azad 2012) and the first stage of bud set (Stage 2.5 in Rohde *et al.,* 2011a, see detailed description below).

2. Ecophysiological traits

Leaf traits

In July 2020, leaves from a subset of 163 ramets randomly chosen were sampled for assessing specific leaf area (SLA) and leaf mass per area (LMA). Three branches from each ramet were chosen randomly. In poplar, new leaves are produced from the apical shoot, so to measure a fully mature leaf, we collected the eighth leaf counting from the

first fully unfolded leaf in the apical shoot of the branch. Only healthy leaves with no signs of biotic or abiotic stress were sampled. If the eighth leaf did not meet the criteria to be considered undamaged, the next healthy leaf was sampled. To assess leaf area, leaves were scanned on a flat white background at 300 dpi using an Epson perfection V370 photo scanner. Based on these scans, leaf area was calculated using the WinFOLIA leaf area analysis software (Regent Instruments Inc.). Leave dry weight was determined by drying leaf samples in an oven at 60° C for 2 days and weighted with a balance precision of 0.001 g. SLA was calculated as the ratio between leaf area and leaf dry weight and LMA as 1/SLA.

Stomatal density

Immediately after the leaves were scanned, two of the three leaves were randomly chosen to determine stomatal density (SD). Epidermal impressions of the abaxial side of the leaves were obtained by creating an imprint of the leaf surface with clear fingernail polish (Maybelline Superstay 7 Days Gel Nail Colour). The polish imprints were mounted in permanent microscope slides with glass covers and photographed under a Zeiss Axio Lab.A1 microscope at 10x magnification in a 450 x 350 µm field. The number of stomata was counted using the StomataCounter software (Fetter *et al.,* 2019). The software automatically annotates visible stomata. In addition, we visually double checked the automated annotation and added undetected stomata manually.

3. Biotic Stress traits

Herbivory damage was assessed in July 2019. A single randomly chosen branch in each ramet was selected and fifty leaves in the branch were scored for the presence or absence of herbivore damage. The number of leaves was counted from the bottom of the branch, and the seven leaves around the shoot area were excluded from the scoring system. When the ramet was too small to have fifty leaves in a single branch, a lower number of leaves was selected, except for two ramets that were completely excluded from the surveys due to their small size. The percentage of damage was calculated as the ratio of the number of damaged leaves to the number of undamaged leaves.

4. Leaf chemistry traits

In June 2020 leaf anthocyanin content (Anth), chlorophyll content (Chl), flavonol content (Flav), and Nitrogen Balanced Index (NBI, a plant status indicator directly correlated with nitrogen content, measured as the ration between chlorophyll and flavonol content) were estimated using a Dualex Scientific meter (FORCE-A, Orsay, France) The meter measures the light transmittance ratio at two different wavelengths in the near-infrared and far-red range to calculate the chlorophyll content. Flavonol and anthocyanin content is calculated based on the amount of light absorbed by polyphenols and the amount that reaches the chlorophyll in the mesophyll. The same subset of clones used for the SLA analysis was used for leaf chemistry analysis. Ten mature healthy leaves per clone were randomly chosen. The Dualex meter measurements were taken on the centre of the adaxial side of the leaves next to the midrib.

5. Phenological traits

Bud phenology, specifically bud set and bud flush, was monitored by focusing on the main apical bud. Bud set was scored in autumn 2018 and 2019 according to the scoring scale designed by Rohde *et al.,* (2011a), and bud flush was scored in spring 2019 and 2020 using the scale suggested by Azad (2012). The bud set scale spans seven stages, from stage 3 (apical shoot fully growing) to stage 0 (bud set), while the bud flush scale consists of six stages, from stage 0 (dormant bud) to stage 5 (leaves fully unfolded). Bud stages were recorded every two to three days at the beginning of the monitoring period and then daily until the apical buds of all clones had reached the final stage of their respective scales. Because bud phenology in *P. nigra* depends greatly on day length and temperature and all ramets were exposed to the same cues, no variation in the day when bud set or bud flush started was expected. The duration of bud formation, however, has been shown to differ in identical genotypes growing under different conditions (Rohde *et al.,* 2011b). Therefore, phenological traits were defined as duration of bud set or duration of bud flush, which equalled the number of days it took each ramet to reach from the first stage to the last stage, respectively.

Climatic variables and climatic gradients

Climatic data for each of the locations of the ortets were obtained from the CHELSA time-series data set (Karger *et al.,* 2017). The CHELSA data set covers the period between 1979 and 2013 and provides gridded data at a resolution of 30 arcsec (~ 1km). A Principal Component Analysis (PCA) was performed using all bioclimatic variables (bioclims, BIO 1-19). Individual coordinates for PC1 and PC2 were obtained and included as variables that represented climatic gradients. The bioclims that contributed the most to PC1 were all related to temperature variables (except for BIO 19, precipitation of the coldest quarter), while the most contributing bioclims in PC2 were related to precipitation variables (except for BIO 2, mean diurnal range). All bioclims and their contributions to each PC are described in Table S1.

Statistical analysis

All statistical analyses were performed in R (version 4.0.3; R core team, 2020). Basic descriptive statistics were calculated for all traits (Table 1). To assess if variation in cutting diameter and developmental processes such as the number of stems produced might have an effect in other phenotypic traits, all traits were correlated with each other using Pearson's Product-moment Correlation with the *cor* function from the base R *stats* package. The effects of geographic origin and microenvironmental conditions on phenotypic variation were tested using linear mixed-effects models (LMMs). Phenotypic traits were the response variable in all models. Two models were fitted for each trait, one with bioclimatic variables (PC1 and PC2) and one with sampling site as fixed effects. To disentangle the possible effects of microenvironmental conditions in the common garden from the effects the ortet provenances had on phenotypic variation, the garden block in which the ramets were planted was included as a random effect in all the LMMs (Supplementary Table 2). Since many of the traits analysed were correlated with tree height and cutting diameter, this source of variation was accounted for by also including these variables in the models as fixed effects. Several models were fitted including tree height in 2018 or 2019, cutting diameter or both variables as fixed effects. Based on the best marginal R^2 values, which we calculated with the *rsquared* function (R package *piecewiseSEM,* version 2.1.2), we decided which variable was included in the model. The LMMs were fit using the *lmer* function from the *lme4* package (version 1.1-23; Bates *et al.,* 2015). Differences between groups were tested using the *emmeans* R package

(version 1.6.3) and p-values were corrected for multiple pairwise comparisons using the Bonferroni correction.

Table 1. Phenotypic trait data description under common garden conditions. Data shown includes the year when the trait was measured (year), the number of clones included in the analysis for each trait (N), minimum value, maximum value, median, mean and standard deviation (SD) for all phenotypic traits used in the statistical analysis. Flavonol and anthocyanin content are given in relative absorbance units (RAU). Nitrogen Balance Index is a unit-less trait

Phenotypic trait	Year	N	Min	Max	Median	Mean ± SD
Tree growth						
Cutting diameter (cm)	2018	373	0.48	4.35	2.27	2.43 ± 0.70
Growth rate (cm day^{-1})	2019	342	0.08	1.76	0.83	0.84 ± 0.18
Height Autumn (cm)	2018	374	13.00	229.50	101.75	100.28 ± 39.16
Height Autumn (cm)	2019	372	47.00	398.00	207.50	203.16 ± 55.45
Height Autumn (cm)	2020	371	70	420	269	263.12 ± 64.11
Stems (#)	2018	374	1	11	2	2.95 ± 1.86
Ecophysiology						
LMA (mg mm^{-2})	2020	163	0.04	0.10	0.06	0.06 ± 0.01
SLA (mm^2 mg^{-1})	2020	163	10.13	26.80	16.34	16.35 ± 2.06
Stomatal Density (mm^{-2})	2020	126	142.86	269.84	206.35	206.15 ± 28.71
Biotic stress						
Herbivory damage (%)	2019	367	2.00	52.00	28.00	27.49 ± 9.64

Table 1. Continuation

Phenotypic trait	Year	N	Min	Max	Median	Mean ± SD
Leaf Chemistry						
Anthocyanins (RAU)	2020	155	0.12	0.18	0.14	0.14 ± 0.01
Chlorophyll (µg cm^{-2})	2020	164	24.79	35.38	30.28	30.18 ± 2.11
Flavonols (RAU)	2020	164	1.88	2.16	2.07	2.07 ± 0.05
Nitrogen Balanced Index	2020	164	11.83	17.70	14.67	14.62 ± 1.12
Phenology						
Days to bud flush (d)	2019	368	14	28	21	21.33 ± 2.26
Days to bud flush (d)	2020	368	8	34	22	22.14 ± 5.88
Days to bud set (d)	2018	374	12	68	22	22.63 ± 6.60
Days to bud set (d)	2019	344	19	39	25	27.28 ± 4.42

Results

Genotypic variation along geographic gradients

The results of the genotyping analysis indicated the presence of three genetic clusters (Figure 1B). The major cluster grouped together with the Italica controls, confirming that most of the individuals sampled and planted in the common garden belonged to the same genetic cluster, which we considered to comprise true Italica clones. The other two clusters included a few individuals from the two sampling sites in Italy (Italian cluster), and Norway (Norwegian cluster). The sequencing yielded information for 8,218 SNP positions. Among these, 3,313 positions (40.3%) were monomorphic across all trees, bringing the number of informative SNPs to 4,905. Of the 433 ramets established in the common garden, 374 belonged to the Italica cluster, with a mean number of pairwise differences among individual

ramets (including the adult clones of the botanical garden) equal to 96 ± 40 (s.d.) SNPs, corresponding to 1.2% of all SNP positions analysed (Figure 1C).

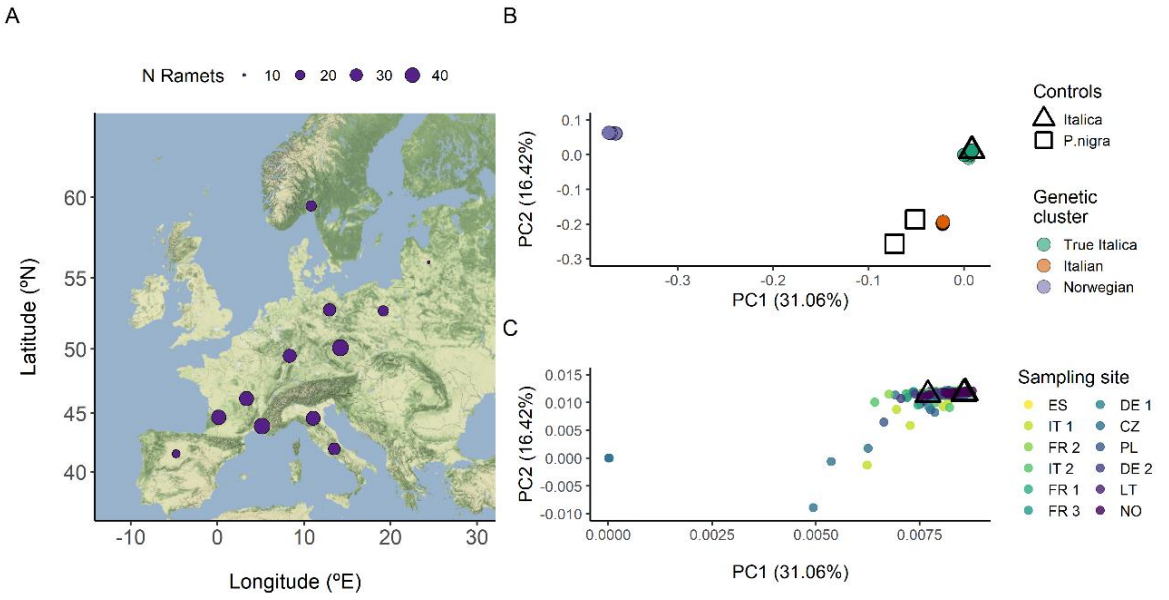


Figure 1. A. Distribution of sample sites of Lombardy poplars and number of ramets collected in each location. B. Genetic structure of all common garden individuals based on the genotyping results. The clusters are coloured according to genetic similarity: True Italica (dark orange), Italian (dark green) or Norwegian (purple). Black triangles indicate the Italica clones from the Marburg Botanical Garden included as controls. Black squares indicate the *P. nigra* individuals included as reference. C. Genetic structure of the True Italica cluster. Individual ramets are coloured according to latitude. Black triangles represent the same three adult *P. nigra cv "Italica*" clones from the Botanical Garden in Marburg that were included in the sequencing design as controls.

Phenotypic variation between geographic regions and along climatic gradients

Phenotypic differences along climatic gradients and among ramets originating from ortets of different geographical origins were tested using linear mixed models (Supplementary Table 3). Four phenotypic traits showed significant correlation with climatic gradients. The number of stems and herbivory damage ($p = 0.034$ and $p = 0.003$, respectively) were correlated with the temperature gradient (PC1), and active growth rate and duration of bud flush 2019 ($p = 0.036$ and $p = 0.044$, respectively) were correlated with the precipitation gradient (PC2).

Also, significant differences among trees from different site of origin (Figure 1) were found for the number of stems ($p = 0.025$) and the duration of bud flush in 2019 ($p = 0.019$). Autumn height in 2019, LMA, SLA, and leaf flavonol content showed a significant correlation with site of origin in the LMMs ($p < 0.05$ in all four traits), but no significant differences between groups after p-value correction. Supplementary figures 2 and 3 show boxplots for all traits where no significant differences were found after correcting the p-values for multiple testing.

Effects of common garden microenvironmental conditions on phenotypic variation

The linear mixed effects models also informed us about the respective importance of the ramets' site of origin vs. microenvironmental conditions in the common garden on phenotypic variation (Supplementary Table 2). In most traits, site of origin and climatic conditions of the ortets accounted for less than 10% of the total phenotypic variation, with the highest fraction of the variation in tree height 2019 and 2020 (marginal $R^2 = 0.102$ and 0.113, respectively) and flavonol content ($R^2 = 0.199$). For all traits, conditional R^2 values were considerably higher than marginal R^2 values (Supplementary Table 2), indicating that unknown microenvironmental variability in the common garden explained a larger fraction of the phenotypic variation among the ramets.

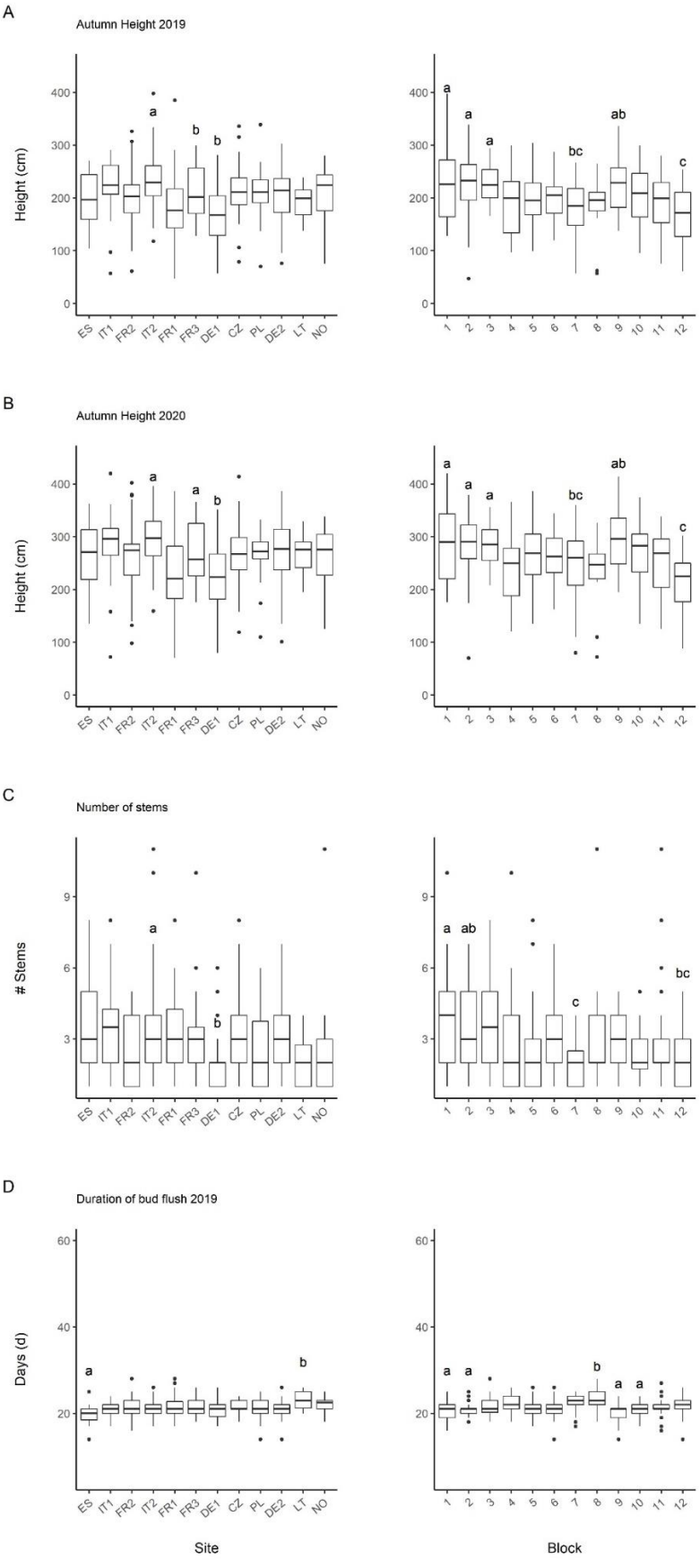

A
Autumn Height 2019
Height (cm)
B
Autumn Height 2020
Height (cm)
C
Number of stems
Stems
D
Duration of bud flush 2019
Days (d)
ES IT1 FR2 IT2 FR1 FR3 DE1 CZ PL DE2 LT NO
1 2 3 4 5 6 7 8 9 10 11 12
Site
Block

Figure 2. Phenotypic variation in traits for which significant differences among geographic origins were found. Boxplots are ordered by site (left column) or by common garden block (right column). Sites are ordered from south to north according to their geographic coordinates and labelled by the sample site code (ISO 3166 standard country code): ES: Spain; IT1: Italy 1; FR2: France 2; IT2: Italy 2; FR1: France 1; FR3: France 3; DE1: Germany 1; CZ: Czech Republic; PL: Poland; DE2: Germany 2; LT: Lithuania; NO: Norway. Sites or blocks that are significantly different ($p < 0.05$) are labelled with different letters.

Correlation among phenotypic traits

To assess if variation in cutting diameter and developmental processes such as number of stems produced might influence phenotypic traits, the relationships among phenotypic traits were assessed using Pearson's Product Moment Correlation (Figure S3). Unsurprisingly, most growth traits were intercorrelated. Tree height in 2018, 2019, and 2020 were positively correlated with active growth rate and negatively correlated with cutting diameter. Tree height was also correlated with ecophysiological traits (LMA and SLA), phenological traits (bud set 2018 and bud flush 2019), and traits related to leaf chemical compounds (flavonol content and NBI). Phenology and leaf chemistry traits also showed intercorrelation. The number of days that trees needed from the start of the bud flush period until the leaves were fully unfolded in 2019 (bud flush 2019 in the tables) were significantly correlated with the number of days required for a complete bud set in 2018 (bud set 2018). There was, however, no correlation between the duration of bud flush in 2020 and the duration of bud set in 2019. A few traits were found to be correlated with climatic variables from the ortet growing sites (PC1 and PC2). Cutting diameter and herbivory damage were positively and negatively correlated, respectively, with the temperature gradient (PC1). Growth rate and bud flush 2019 showed a weak negative correlation with the precipitation gradient (PC2). Bud flush 2020 and bud set 2018, on the other hand, were negatively correlated with PC2.

Discussion

In this study, we investigated the effect of previous environments on the clonal offspring of Lombardy poplar ramets growing under common garden conditions. We found that phenotypic variation of the ramets correlates with geographical and climatic gradients, but that uneven new microenvironmental conditions in the common garden might outweigh parental effects.

Our genotyping results indicated that the poplar clones established in our common garden were characterized by very low genetic diversity (Figure 1B-C). We found that the ramets that belonged to the Italica genetic cluster had a mean number of pairwise differences among individual ramets of around 96 SNPs out of the 4.906 investigated remaining positions. The Italica cultivar likely originated from a single male clone in Central Asia, from where it spread to Europe. It is widely accepted that this clone was further artificially propagated from an individual or group of individuals found in Lombardy, Italy (Elwes and Henry, 1913). Our results suggest that a major fraction of the clones found across Europe do indeed share a common line. The restricted number of SNPs available from the targeted genotyping could also explain why no population structure was observed in our study. We targeted 8,000 loci equally distributed across the 19 *P. nigra* chromosomes selected from a larger set identified in Scaglione et. al (2019), which should allow for accurate and effective genotyping of population groups.

Despite the low genetic variation, four of the phenotypic traits (growth rate, number of stems produced at sprouting, herbivory damage and duration of bud flush) showed a significant correlation with climatic gradients, and six traits (tree height, number of stems, LMA, SLA, flavonol content and duration of bud flush) showed significant differences among ramets from different geographical regions. Both bud phenology and tree growth can directly affect tree performance and fitness (Cooke *et al.,* 2012), and the differences observed in our common garden could have been influenced by the differences in the developmental stages of ortets and ramets. Hoewever, both traits have been shown to be epigenetically regulated to a certain point (Bräutigam *et al.,* 2013; Ríos *et al.,* 2014; Lu *et al.,* 2020). The presence of an epigenetic memory would be a major advantage for new saplings if the relevant environmental conditions remain relatively constant over time. The environmental requirements that are needed for dormancy break and bud flush are still not fully understood and are species-dependent, but it is widely accepted that temperate tree species rely on temperature and photoperiodic cues to trigger bud flush (Rohde & Bhalerao, 2007; Ibáñez

et al., 2010; Malyshev *et al.,* 2018). In poplar, after the chilling requirements and a certain day length threshold are reached, growing temperatures will trigger the processes related to bud flush (Singh *et al.,* 2017). Consequently, under the same temperature and light conditions, no differences in the number of days that ramets from the same genotype needed for bud flush would be expected. The fact that we found significant differences in tree growth traits and bud flush, suggests that environmentally induced epigenetic memory might play a role. Although differences in tree height persisted over two growing seasons, the differences in the duration of bud flush seem to have disappeared in 2020, suggesting that the parental effects might not be very stable (Shi *et al.,* 2019a). This could potentially also be an advantage, if the survival of the clones depends on the existence of a mechanism of rapid acclimation to unpredicted conditions. For example, multiple studies have suggested that high phenotypic plasticity levels associated with epigenetic diversity might contribute to the successful establishment of clonal (and often invasive) plant species (Davidson *et al.,* 2011; Richards *et al.,* 2012; Mounger *et al.,* 2020).

The phenotypic variation among ramets from different geographic origins observed in our common garden was generally low for the remaining traits (Supplementary figures 2 and 3). As evidenced by the considerably low marginal R^2 values of the linear models, in several traits this variation was not explained by the geographical origins of the ramets or by environmental gradients. Conditional R^2 values were, in comparison, larger than marginal R^2 values for all traits. In the models, the only variable considered as random effect was the block of the common garden where each ramet was planted. Our results indicate that microenvironmental conditions in the common garden explained a major fraction of the phenotypic variation found between ramets. The exact causes of environmental differences between blocks are unknown. However, we suspect that fine scale soil composition heterogeneity, combined with the strong summer droughts 2018 and 2019 (Schuldt *et al.,* 2020) and daily watering have provided sufficient stress and heterogeneity in growing conditions in the common garden field that epigenetic effects from the ortet origins are overshadowed. This fits to other results, where small-scale biotic and abiotic conditions experienced by individuals have been shown to dramatically influence phenotypic plasticity, genetic variation, and population persistence (Wu, 1996; Crutsinger, 2015; Denney *et al.,* 2020). Though we do have to note that other traits like cutting diameter might act as confounding factors, our results suggest that general plastic responses to extreme environmental conditions can outweigh parental effects and exacerbate even microenvironmental differences, masking any potentially inherited epigenetic variation.

Conclusions

Although intraspecific phenotypic variation in *Populus sp.* has been shown to have a large genetic component, our results indicate that the phenotypic differences found between genetically identical ramets under common garden conditions can partially be attributed to shared environmental conditions and might be transmitted as part of the epigenetic memory. However, we also found that uneven microenvironmental conditions in the common garden had a significant effect on the observed phenotypic variation, possibly overwriting parental effects and thus allowing for short term acclimation to new environmental conditions. In recent years, epigenetic variation has been shown to play a bigger role on phenotypic plasticity than previously thought. Further experimental research, in particular large-scale studies that combine phenotypic and epigenomic data, is necessary to understand the effects of natural epigenetic variation on phenotypic variation.

Acknowledgments

The authors want to thank Philipp Kurth and David Löning for their help with sampling, and Phillip Kaldeway, Leonie Braasch and Tabea Giese for their contributions to data collection. We thank Benjamin Dauphin for his help with bioclimatic data. We also thank all the members of the EpiDiverse Consortium for their support. We are grateful to An Vanden Broeck for her advice.

Funding

This work was supported by the European Training Network “EpiDiverse” and received funding from the EU Horizon 2020 program under Marie SkłodowskaCurie grant agreement No 764965.

Data availability

The data that support the findings of this study are openly available in Zenodo at

https://doi.org/10.5281/zenodo.5995424

Literature cited

Aitken SN, Yeaman S, Holliday JA, Wang T, Curtis-McLane S. **2008**. Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications* **1**: 95–111.

Akalin A, Kormaksson M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, Mason CE. **2012**. MethylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biology* **13**: 1–9.

Ashe A, Colot V, Oldroyd BP. **2021**. How does epigenetics influence the course of evolution? *Philosophical Transactions of the Royal Society B* **376**.

Azad MdS. **2012**. Observations on bud burst phenology in a field trial established with Poplar (*Populus spp*.). *Forestry Studies in China* **14**: 251–259.

Baduel P, Leduque B, Ignace A, Gy I, Gil J, Loudet O, Colot V, Quadrana L. **2021**. Genetic and environmental modulation of transposition shapes the evolutionary potential of Arabidopsis thaliana. *Genome Biology* **22**: 1–26.

Balloux F, Lehmann L, de Meeûs T. **2003**. The Population Genetics of Clonal and Partially Clonal Diploids. *Genetics* **164**: 1635–1644.

Banta JA, Richards CL. **2018**. Quantitative epigenetics and evolution. *Heredity 2018 121:3* **121**: 210–224.

Bates D, Mächler M, Bolker B, Walker S. **2015**. Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software* **67**.

Benito Garzón M, Alía R, Robson TM, Zavala MA. **2011**. Intra-specific variability and plasticity influence potential tree species distributions under climate change. *Global Ecology and Biogeography* **20**: 766–778.

Bisbing SM, Urza AK, Buma BJ, Cooper DJ, Matocq M, Angert AL. **2021**. Can long-lived species keep pace with climate change? Evidence of local persistence potential in a widespread conifer. *Diversity and Distributions* **27**: 296–312.

Boquete MT, Muyle A, Alonso C. **2021**. Plant epigenetics: phenotypic and functional diversity beyond the DNA sequence. *American Journal of Botany* **108**: 553–558.

Bossdorf O, Richards CL, Pigliucci M. **2007**. Epigenetics for ecologists. *Ecology Letters* **0**: 071117033013002-???

Bossdorf O, Richards CL, Pigliucci M. **2008**. Epigenetics for ecologists. *Ecology Letters* **11**: 106–115.

Bräutigam K, Vining KJ, Lafon-Placette C, Fossdal CG, Mirouze M, Marcos JG, Fluch S, Fraga MF, Guevara MÁ, Abarca D, *et al.,* 2013. Epigenetic regulation of adaptive responses of forest tree species to the environment. *Ecology and Evolution* **3**: 399–415.

vanden Broeck A, Cox K, Brys R, Castiglione S, Cicatelli A, Guarino F, Heinze B, Steenackers M, vander Mijnsbrugge K. **2018**. Variability in DNA Methylation and Generational Plasticity in the Lombardy Poplar, a Single Genotype Worldwide Distributed Since the Eighteenth Century. *Frontiers in Plant Science* **9**: 1635.

Callahan HS, Pigliucci M, Schlichting CD. **1997**. Developmental phenotypic plasticity: Where ecology and evolution meet molecular biology. *BioEssays* **19**: 519–525.

Climate Change 2022: Impacts, Adaptation and Vulnerability | Climate Change 2022: Impacts, Adaptation and Vulnerability.

Cooke JEK, Eriksson ME, Junttila O. **2012**. The dynamic nature of bud dormancy in trees: Environmental control and molecular mechanisms. *Plant, Cell and Environment* **35**: 1707–1728.

Cornes RC, van der Schrier G, van den Besselaar EJM, Jones PD. **2018**. An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets. *Journal of Geophysical Research: Atmospheres* **123**: 9391–9409.

Crutsinger GM. **2015**. A community genetics perspective: opportunities for the coming decade. *New Phytologist* **210**: 65–70.

Cubas P, Vincent C, Coen E. **1999**. An epigenetic mutation responsible for natural variation in floral symmetry. *Nature* **401**: 157–161.

Danchin É. **2013**. Avatars of information: towards an inclusive evolutionary synthesis. *Trends in Ecology & Evolution* **28**: 351–358.

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, *et al.,* 2021. Twelve years of SAMtools and BCFtools. *GigaScience* **10**: 1–4.

Darrow GM. **1966**. The strawberry. History, breeding and physiology. *The strawberry. History, breeding and physiology.*

Davidson AM, Jennions M, Nicotra AB. **2011**. Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta-analysis. *Ecology Letters* **14**: 419–431.

Denney DA, Jameel MI, Bemmels JB, Rochford ME, Anderson JT. **2020**. Small spaces, big impacts: contributions of micro-environmental variation to population persistence under climate change. *AoB PLANTS* **12**.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. **2013**. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)* **29**: 15–21.

Dodd RS, Douhovnikoff V. **2016a**. Adjusting to Global Change through Clonal Growth and Epigenetic Variation. *Frontiers in Ecology and Evolution* **4**: 86.

Dodd RS, Douhovnikoff V. **2016b**. Adjusting to Global Change through Clonal Growth and Epigenetic Variation. *Frontiers in Ecology and Evolution* **0**: 86.

Dubin MJ, Zhang P, Meng D, Remigereau MS, Osborne EJ, Casale FP, Drewe P, Kahles A, Jean G, Vilhjálmsson B, *et al.,* 2015. DNA methylation in Arabidopsis has a genetic basis and shows evidence of local adaptation. *eLife* **4**.

Edger PP, VanBuren R, Colle M, Poorten TJ, Wai CM, Niederhuth CE, Alger EI, Ou S, Acharya CB, Wang J, *et al.,* 2018. Single-molecule sequencing and optical mapping yields an improved genome of woodland strawberry (*Fragaria vesca*) with chromosome-scale contiguity. *GigaScience* **7**.

Elwes HJ, Henry A. **1913**. The trees of Great Britain & Ireland. R. & R. Clark, Limited, Edinburgh, 2022 pp.

Ewels P, Magnusson M, Lundin S, Käller M. **2016**. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics (Oxford, England)* **32**: 3047–3048.

Faivre-Rampant P, Zaina G, Jorge V, Giacomello S, Segura V, Scalabrin S, Guérin V, de Paoli E, Aluome C, Viger M, *et al.,* 2016. New resources for genetic studies in *Populus nigra*: genome-

wide SNP discovery and development of a 12k Infinium array. *Molecular Ecology Resources* **16**: 1023–1036.

Feng S, Jacobsen SE, Reik W. **2010**. Epigenetic reprogramming in plant and animal development. *Science (New York, N.Y.)* **330**: 622–627.

Fetter KC, Eberhardt S, Barclay RS, Wing S, Keller SR. **2019**. StomataCounter: a neural network for automatic stomata identification and counting. *New Phytologist* **223**: 1671–1681.

Fojtová M, Kovařík A, Matyášek R. **2001**. Cytosine methylation of plastid genome in higher plants. Fact or artefact? *Plant science : an international journal of experimental plant biology* **160**: 585–593.

Frels K, Chopra R, Dorn KM, Wyse DL, Marks MD, Anderson JA. **2019**. Genetic Diversity of Field Pennycress (Thlaspi arvense) Reveals Untapped Variability and Paths Toward Selection for Domestication. *Agronomy 2019, Vol. 9, Page 302* **9**: 302.

de Frenne P, Graae BJ, Rodríguez-Sánchez F, Kolb A, Chabrerie O, Decocq G, de Kort H, de Schrijver A, Diekmann M, Eriksson O, *et al.,* 2013. Latitudinal gradients as natural laboratories to infer species' responses to temperature. *Journal of Ecology* **101**: 784–795.

Fu YB. **2015**. Understanding crop genetic diversity under modern plant breeding. *Theoretical and Applied Genetics* **128**: 2131–2142.

Fultz D, Choudury SG, Slotkin RK. **2015**. Silencing of active transposable elements in plants. *Current Opinion in Plant Biology* **27**: 67–76.

Ganguly DR, Crisp PA, Eichten SR, Pogson BJ. **2017**. The Arabidopsis DNA Methylome Is Stable under Transgenerational Drought Stress. *Plant physiology* **175**: 1893–1912.

Gent JI, Ellis NA, Guo L, Harkess AE, Yao Y, Zhang X, Dawe RK. **2013**. CHH islands: de novo DNA methylation in near-gene chromatin regulation in maize. *Genome Research* **23**: 628–637.

Gornall JL, Guy RD. **2007**. Geographic variation in ecophysiological traits of black cottonwood (*Populus trichocarpa*). *Canadian Journal of Botany* **85**: 1202–1213.

Guarino F, Cicatelli A, Brundu G, Heinze B, Castiglione S. **2015**. Epigenetic Diversity of Clonal White Poplar (*Populus alba* L.) Populations: Could Methylation Support the Success of Vegetative Reproduction Strategy? *PLOS ONE* **10**: e0131480.

Heer K, Mounger J, Boquete MT, Richards CL, Opgenoorth L. **2018**. The diversifying field of plant epigenetics. *New Phytologist* **217**: 988–992.

Howe GT, Saruul P, Davis J, Chen THH. **2000**. Quantitative genetics of bud phenology, frost damage, and winter survival in an F2 family of hybrid poplars. *Theoretical and Applied Genetics* **101**: 632–642.

Hughes AR, Inouye BD, Johnson MTJ, Underwood N, Vellend M. **2008**. Ecological consequences of genetic diversity. *Ecology Letters* **11**: 609–623.

Ibáñez C, Kozarewa I, Johansson M, Ogren E, Rohde A, Eriksson ME. **2010**. Circadian clock components regulate entry and affect exit of seasonal dormancy as well as winter hardiness in *Populus* trees. *Plant physiology* **153**: 1823–33.

Ingvarsson PK, Dahlberg H. **2018**. The effects of clonal forestry on genetic diversity in wild and domesticated stands of forest trees. *https://doi.org/10.1080/02827581.2018.1469665* **34**: 370–379.

Jablonka EVA, Raz GAL. **2009**. Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *Quarterly Review of Biology* **84**: 131–176.

Jansson S, Douglas CJ. **2007**. *Populus* : A Model System for Plant Biology. *Annual Review of Plant Biology* **58**: 435–458.

Jühling F, Kretzmer H, Bernhart SH, Otto C, Stadler PF, Hoffmann S. **2016**. metilene: fast and sensitive calling of differentially methylated regions from bisulfite sequencing data. *Genome research* **26**: 256–262.

Jung S, Lee T, Cheng CH, Buble K, Zheng P, Yu J, Humann J, Ficklin SP, Gasic K, Scott K, *et al.,* 2019. 15 years of GDR: New data and functionality in the Genome Database for Rosaceae. *Nucleic Acids Research* **47**: D1137.

Karger DN, Conrad O, Böhner J, Kawohl T, Kreft H, Soria-Auza RW, Zimmermann NE, Linder HP, Kessler M. **2017**. Climatologies at high resolution for the earth's land surface areas. *Scientific Data* **4**: 170122.

De Kort H, Panis B, Helsen K, Douzet R, Janssens SB, Honnay O. **2020**. Pre-adaptation to climate change through topography-driven phenotypic plasticity. *Journal of Ecology* **108**: 1465–1474.

De Kort H, Toivainen T, Van Nieuwerburgh F, Andrés J, Hytönen TP, Honnay O. **2022**. Signatures of polygenic adaptation align with genome-wide methylation patterns in wild strawberry plants. *New Phytologist*.

Kovaka S, Zimin A V., Pertea GM, Razaghi R, Salzberg SL, Pertea M. **2019**. Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biology* **20**: 1–13.

Lang Z, Wang Y, Tang K, Tang D, Datsenka T, Cheng J, Zhang Y, Handa AK, Zhu JK. **2017**. Critical roles of DNA demethylation in the activation of ripening-induced genes and inhibition of ripening-repressed genes in tomato fruit. *Proceedings of the National Academy of Sciences of the United States of America* **114**: E4511–E4519.

Latzel V, Klimešová J. **2010a**. Transgenerational plasticity in clonal plants. *Evolutionary Ecology* **24**: 1537–1543.

Latzel V, Klimešová J. **2010b**. Transgenerational plasticity in clonal plants. *Evolutionary Ecology* **24**: 1537–1543.

Latzel V, Rendina González AP, Rosenthal J. **2016a**. Epigenetic memory as a basis for intelligent behavior in clonal plants. *Frontiers in Plant Science* **7**.

Latzel V, Rendina González AP, Rosenthal J. **2016b**. Epigenetic Memory as a Basis for Intelligent Behavior in Clonal Plants. *Frontiers in Plant Science* **7**: 1354.

Li Y, Pi M, Gao Q, Liu Z, Kang C. **2019**. Updated annotation of the wild strawberry *Fragaria vesca* V4 genome. *Horticulture Research* **6**.

Li X, Zhu J, Hu F, Ge S, Ye M, Xiang H, Zhang G, Zheng X, Zhang H, Zhang S, *et al.,* 2012. Single-base resolution maps of cultivated and wild rice methylomes and regulatory roles of DNA methylation in plant gene expression. *BMC Genomics* **13**: 1–15.

Linnaeus C, Rudberg D. **1744**. Dissertatio botanica de Peloria. *Amoenitates academicae* **1**: 280–297.

López M-E, Roquis D, Becker C, Denoyes B, Bucher E. **2022**. DNA methylation dynamics during stress response in woodland strawberry (*Fragaria vesca*). *Horticulture Research* **9**.

Lu W, Xiao L, Quan M, Wang Q, El-Kassaby YA, Du Q, Zhang D. **2020**. Linkage-linkage disequilibrium dissection of the epigenetic quantitative trait loci (epiQTLs) underlying growth and wood properties in *Populus*. *New Phytologist* **225**: 1218–1233.

Luquez V, Hall D, Albrectsen BR, Karlsson J, Ingvarsson P, Jansson S. **2008**. Natural phenological variation in aspen (*Populus tremula*): The SwAsp collection. *Tree Genetics and Genomes* **4**: 279–292.

Ma X-F, Hall D, Onge KR st., Jansson S, Ingvarsson PK. **2010**. Genetic Differentiation, Clinal Variation and Phenotypic Associations With Growth Cessation Across the *Populus tremula* Photoperiodic Pathway. *Genetics* **186**: 1033–1044.

Maher B. **2008**. Personal genomes: The case of the missing heritability. *Nature* **456**: 18–21.

Malyshev A v., Henry HAL, Bolte A, Arfin Khan MAS, Kreyling J. **2018**. Temporal photoperiod sensitivity and forcing requirements for budburst in temperate tree seedlings. *Agricultural and Forest Meteorology* **248**: 82–90.

Martin GT, Seymour DK, Gaut BS. **2021**. CHH Methylation Islands: A Nonconserved Feature of Grass Genomes That Is Positively Associated with Transposable Elements but Negatively Associated with Gene-Body Methylation. *Genome Biology and Evolution* **13**.

McKown AD, Guy RD, Klápště J, Geraldes A, Friedmann M, Cronk QCB, El-Kassaby YA, Mansfield SD, Douglas CJ. **2014**. Geographical and environmental gradients shape phenotypic trait variation and genetic structure in *Populus trichocarpa*. *New Phytologist* **201**: 1263–1276.

Meloni M, Reid A, Caujapé-Castells J, Marrero Á, Fernández-Palacios JM, Mesa-Coelo RA, Conti E. **2013**. Effects of clonality on the genetic variability of rare, insular species: the case of Ruta microcarpa from the Canary Islands. *Ecology and Evolution* **3**: 1569–1579.

Moore LD, Le T, Fan G. **2012**. DNA Methylation and Its Basic Function. *Neuropsychopharmacology 2013 38:1* **38**: 23–38.

Mounger J, Ainouche M, Bossdorf O, Cavé-Radet A, Li B, Parepa M, Salmon A, Yang J, Richards C. **2020**. Epigenetics and the success of invasive plants.

Münzbergová Z, Hadincová V. **2016**. Transgenerational plasticity as an important mechanism affecting response of clonal species to changing climate. *Ecology and Evolution* **7**: 5236–5247.

Münzbergová Z, Latzel V, Šurinová M, Hadincová V. **2019**. DNA methylation as a possible mechanism affecting ability of natural populations to adapt to changing climate. *Oikos* **128**: 124–134.

Niederhuth CE, Schmitz RJ. **2014**. Covering Your Bases: Inheritance of DNA Methylation in Plant Genomes. *Molecular Plant* **7**: 472–480.

Nunn A, Can SN, Otto C, Fasold M, Diéz Rodríguez B, Fernández-Pozo N, Rensing SA, Stadler PF, Langenberger D. **2021**. EpiDiverse Toolkit: a pipeline suite for the analysis of bisulfite sequencing data in ecological plant epigenetics. *NAR Genomics and Bioinformatics* **3**.

Nunn A, Otto C, Fasold M, Stadler PF, Langenberger D. **2022**. Manipulating base quality scores enables variant calling from bisulfite sequencing alignments using conventional Bayesian approaches. *bioRxiv*: 2021.01.11.425926.

Oksanen Jari, Blanchet F. Guillaume, Friendly Michael, Kindt Roeland, Legendre Pierre, McGlinn Dan, Minchin Peter R., O'Hara R. B., Simpson Gavin L., Solymos Peter, *et al.,* 2020. CRAN - Package vegan.

Oliver KR, Greene WK. **2009**. Transposable elements: powerful facilitators of evolution. *BioEssays : news and reviews in molecular, cellular and developmental biology* **31**: 703–714.

Ou S, Su W, Liao Y, Chougule K, Agda JRA, Hellinga AJ, Lugo CSB, Elliott TA, Ware D, Peterson T, *et al.,* 2019. Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome biology* **20**.

Peel MC, Finlayson BL, Mcmahon TA. **2007**. Updated world map of the Köppen-Geiger climate classification.

Peter Meyer IN, Lohuis M Ten. **1994**. Evidence for cytosine methylation of non-symmetrical sequences in transgenic Petunia hybrida. *The EMBO journal* **13**: 2084–2088.

Pfenninger M, Reuss F, Kiebler A, Schönnenbeck P, Caliendo C, Gerber S, Cocchiararo B, Reuter S, Blüthgen N, Mody K, *et al.,* 2021. Genomic basis for drought resistance in european beech forests threatened by climate change. *eLife* **10**.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, De Bakker PIW, Daly MJ, *et al.,* 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics* **81**: 559–575.

Quinlan AR, Hall IM. **2010**. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics (Oxford, England)* **26**: 841–842.

Raj S, Bräutigam K, Hamanishi ET, Wilkins O, Thomas BR, Schroeder W, Mansfield SD, Plant AL, Campbell MM. **2011a**. Clone history shapes Populus drought responses. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 12521–12526.

Rajkumar MS, Shankar R, Garg R, Jain M. **2020**. Bisulphite sequencing reveals dynamic DNA methylation under desiccation and salinity stresses in rice cultivars. *Genomics* **112**: 3537–3548.

Rapp RA, Wendel JF. **2005**. Epigenetics and plant evolution. *New Phytologist* **168**: 81–91.

Rasmussen KK, Kollmann J. **2007**. Low genetic diversity in small peripheral populations of a rare European tree (Sorbus torminalis) dominated by clonal reproduction. *Conservation Genetics 2007 9:6* **9**: 1533–1539.

Rey O, Danchin E, Mirouze M, Loot C, Blanchet S. **2016**. Adaptation to Global Change: A Transposable Element-Epigenetics Perspective. *Trends in Ecology and Evolution* **31**: 514–526.

Richards EJ. **2006**. Inherited epigenetic variation — revisiting soft inheritance. *Nature Reviews Genetics 2006 7:5* **7**: 395–401.

Richards CL, Alonso C, Becker C, Bossdorf O, Bucher E, Colomé-Tatché M, Durka W, Engelhardt J, Gaspar B, Gogol-Döring A, *et al.,* 2017a. Ecological plant epigenetics: Evidence from model and non-model species, and the way forward (T Coulson, Ed.). *Ecology Letters* **20**: 1576–1590.

Richards CL, Alonso C, Becker C, Bossdorf O, Bucher E, Colomé-Tatché M, Durka W, Engelhardt J, Gaspar B, Gogol-Döring A, *et al.,* 2017b. Ecological plant epigenetics: Evidence from model and non-model species, and the way forward. *Ecology Letters* **20**: 1576–1590.

Richards CL, Schrey AW, Pigliucci M. **2012**. Invasion of diverse habitats by few Japanese knotweed genotypes is correlated with epigenetic differentiation (M Vellend, Ed.). *Ecology Letters* **15**: 1016–1025.

Riggs AD, Porter TN. **1996**. Overview of epigenetic mechanisms. *Cold Spring Harbor monograph series* **32**: 29–46.

Ríos G, Leida C, Conejero A, Badenes ML. **2014**. Epigenetic regulation of bud dormancy events in perennial plants. *Frontiers in Plant Science* **5**: 247.

Robinson KM, Ingvarsson PK, Jansson S, Albrectsen BR, Soares J, Macaya-Sanz D. **2012**. Genetic Variation in Functional Traits Influences Arthropod Community Composition in Aspen (*Populus tremula* L.) (DJ Kliebenstein, Ed.). *PLoS ONE* **7**: e37679.

Rohde A, Bastien C, Boerjan W, Thomas S. **2011**. Temperature signals contribute to the timing of photoperiodic growth cessation and bud set in poplar. *Tree Physiology* **31**: 472–482.

Rohde A, Bhalerao RP. **2007**. Plant dormancy in the perennial context. *Trends in Plant Science* **12**: 217–223.

Sammarco I, Münzbergová Z, Latzel V. **2022**. DNA Methylation Can Mediate Local Adaptation and Response to Climate Change in the Clonal Plant *Fragaria vesca*: Evidence From a European-Scale Reciprocal Transplant Experiment. *Frontiers in Plant Science* **13**: 435.

Scaglione D, Pinosio S, Marroni F, Centa E di, Fornasiero A, Magris G, Scalabrin S, Cattonaro F, Taylor G, Morgante M. **2019**. Single primer enrichment technology as a tool for massive genotyping: a benchmark on black poplar and maize. *Annals of Botany* **124**: 543–551.

Scheiner SM, Goodnight CJ. **1984**. The Comparison of Phenotypic Plasticity and Genetic Variation in Populations of the Grass Danthonia spicata. *Evolution* **38**: 845.

Schemske DW, Mittelbach GG, Cornell H v., Sobel JM, Roy K. **2009**. Is There a Latitudinal Gradient in the Importance of Biotic Interactions? *Annual Review of Ecology, Evolution, and Systematics* **40**: 245–269.

Schmidt AL, Anderson LM. **2006**. Repetitive DNA elements as mediators of genomic change in response to environmental cues. *Biological reviews of the Cambridge Philosophical Society* **81**: 531–543.

Schuldt B, Buras A, Arend M, Vitasse Y, Beierkuhnlein C, Damm A, Gharun M, Grams TEE, Hauck M, Hajek P, *et al.,* 2020. A first assessment of the impact of the extreme 2018 summer drought on Central European forests. *Basic and Applied Ecology* **45**: 86–103.

Schulze J, Rufener R, Erhardt A, Stoll P. **2012**. The relative importance of sexual and clonal reproduction for population growth in the perennial herb *Fragaria vesca*. *Population Ecology 2012 54:3* **54**: 369–380.

Scotti I, González-Martínez SC, Budde KB, Lalagüe H. **2016**. Fifty years of genetic studies: what to make of the large amounts of variation found within populations? *Annals of Forest Science* **73**: 69–75.

Secco D, Wang C, Shou H, Schultz MD, Chiarenza S, Nussaume L, Ecker JR, Whelan J, Lister R. **2015**. Stress induced gene expression drives transient DNA methylation changes at adjacent repetitive elements. *eLife* **4**.

Shi W, Chen X, Gao L, Xu CY, Ou X, Bossdorf O, Yang J, Geng Y. **2019a**. Transient stability of epigenetic population differentiation in a clonal invader. *Frontiers in Plant Science* **9**: 1851.

Shi W, Chen X, Gao L, Xu C-Y, Ou X, Bossdorf O, Yang J, Geng Y. **2019b**. Transient Stability of Epigenetic Population Differentiation in a Clonal Invader. *Frontiers in Plant Science* **9**: 1851.

Sigg CD, Buhmann JM. **2008**. Expectation-maximization for sparse and non-negative PCA. *Proceedings of the 25th International Conference on Machine Learning*: 960–967.

Simonetti G, Pucci N, Brasili E, Valletta A, Sammarco I, Carnevale E, Pasqua G, Loreti S. **2019**. In vitro antimicrobial activity of plant extracts against Pseudomonas syringae pv. actinidiae causal agent of bacterial canker in kiwifruit. *Plant Biosystems*.

Singh RK, Svystun T, AlDahmash B, Jönsson AM, Bhalerao RP. **2017**. Photoperiod- and temperature-mediated control of phenology in trees – a molecular perspective. *New Phytologist* **213**: 511–524.

Slavov GT, Leonardi S, Adams WT, Strauss SH, Difazio SP. **2010**. Population substructure in continuous and fragmented stands of *Populus trichocarpa*. *Heredity* **105**: 348–357.

Slotkin RK, Martienssen R. **2007a**. Transposable elements and the epigenetic regulation of the genome. *Nature Reviews Genetics 2007 8:4* **8**: 272–285.

Slotkin RK, Martienssen R. **2007b**. Transposable elements and the epigenetic regulation of the genome. *Nature Reviews Genetics* **8**: 272–285.

Sork VL, Aitken SN, Dyer RJ, Eckert AJ, Legendre P, Neale DB. **2013**. Putting the landscape into the genomics of trees: Approaches for understanding local adaptation and population responses to changing climate. *Tree Genetics and Genomes* **9**: 901–911.

Sow MD, le Gac A-L, Lesage-Descauses M-C, Strauss SH. **2020**. Hypomethylated poplars show higher tolerance to water deficit and highlight a dual role for DNA methylation in shoot meristem: regulation of stress response and genome integrity. *bioRxiv*: 2020.04.16.045328.

TAYLOR G. **2002**. *Populus*: Arabidopsis for Forestry. Do We Need a Model Tree? *Annals of Botany* **90**: 681–689.

Thiebaut F, Hemerly AS, Ferreira PCG. **2019**. A role for epigenetic regulation in the adaptation and stress responses of non-model plants. *Frontiers in Plant Science* **10**.

Todd LA, Hall AC, Pietrobon V, Chan JNY, Laflamme G, Mekhail K. **2020**. RNA-cDNA hybrids mediate transposition via different mechanisms. *Scientific Reports 2020 10:1* **10**: 1–12.

Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam M, Ralph S, Rombauts S, Salamov A, *et al.,* 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**: 1596–1604.

Verhoeven KJF, Preite V. **2014b**. Epigenetic variation in asexually reproducing organisms. *Evolution* **68**: 644–655.

Vitasse Y, Bottero A, Cailleret M, Bigler C, Fonti P, Gessler A, Lévesque M, Rohner B, Weber P, Rigling A, *et al.,* 2019. Contrasting resistance and resilience to extreme drought and late spring frost in five major European tree species. *Global Change Biology* **25**: 3781–3792.

Wang G, Li H, Meng S, Yang J, Ye N, Zhang J. **2020**. Analysis of Global Methylome and Gene Expression during Carbon Reserve Mobilization in Stems under Soil Drying. *Plant Physiology* **183**: 1809–1824.

Warwick SI, Francis A, Susko DJ. **2011**. The biology of Canadian weeds. 9. Thlaspi arvense L. (updated). *https://doi.org/10.4141/P01-159* **82**: 803–823.

Wibowo A, Becker C, Marconi G, Durr J, Price J, Hagmann J, Papareddy R, Putra H, Kageyama J, Becker J, *et al.,* 2016. Hyperosmotic stress memory in Arabidopsis is mediated by distinct epigenetically labile sites in the genome and is restricted in the male germline by DNA glycosylase activity. *eLife* **5**.

Wu RL. **1996**. Genetic control of macro- and micro-environmental sensitivities in *Populus*. *Theoretical and Applied Genetics* **94**: 104–114.

Xu J, Zhou S, Gong X, Song Y, van Nocker S, Ma F, Guan Q. **2018**. Single-base methylome analysis reveals dynamic epigenomic differences associated with water deficit in apple. *Plant Biotechnology Journal* **16**: 672–687.

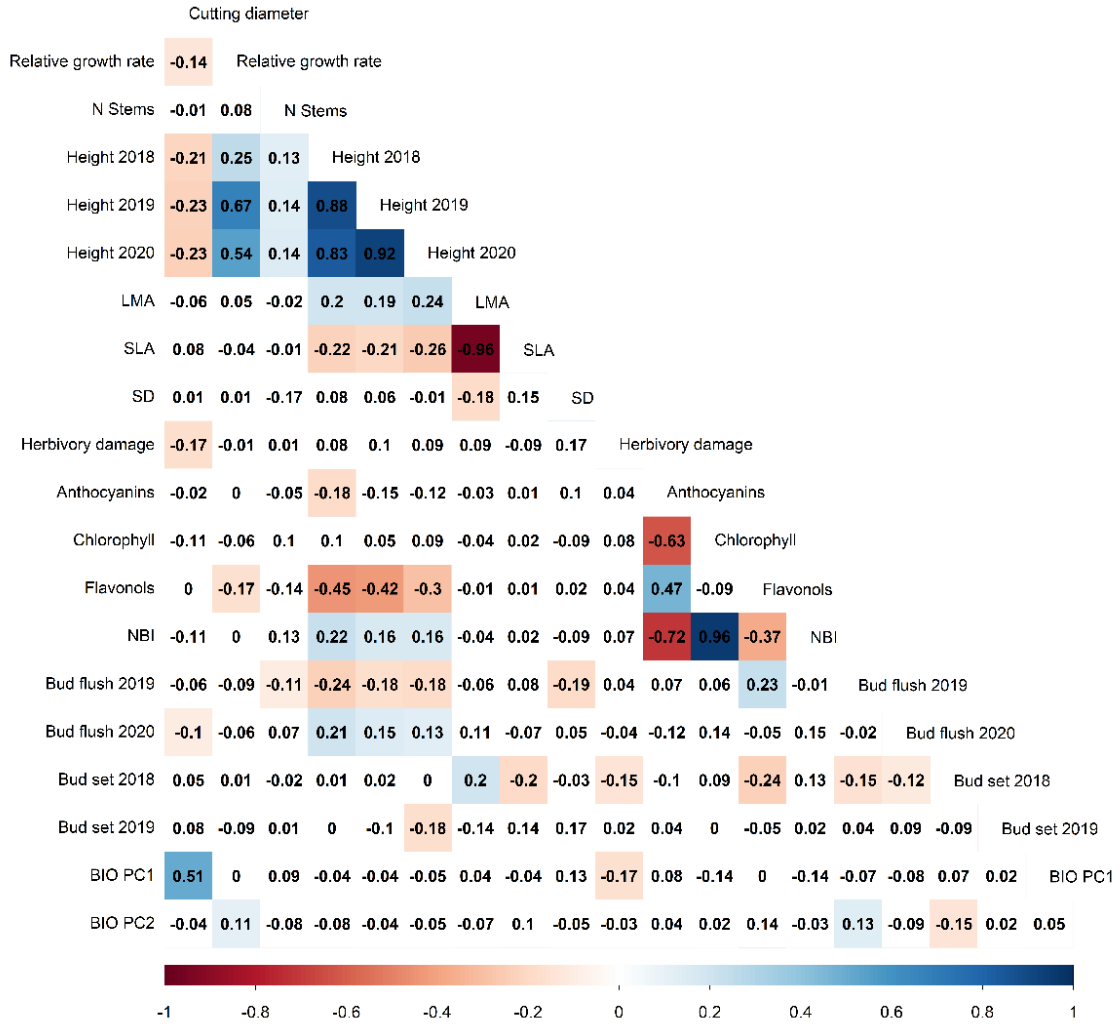
Yu G, Wang LG, Han Y, He QY. **2012**. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics : a journal of integrative biology* **16**: 284–287.

Zemach A, McDaniel IE, Silva P, Zilberman D. **2010**. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science (New York, N.Y.)* **328**: 916–919.

Supplementary material

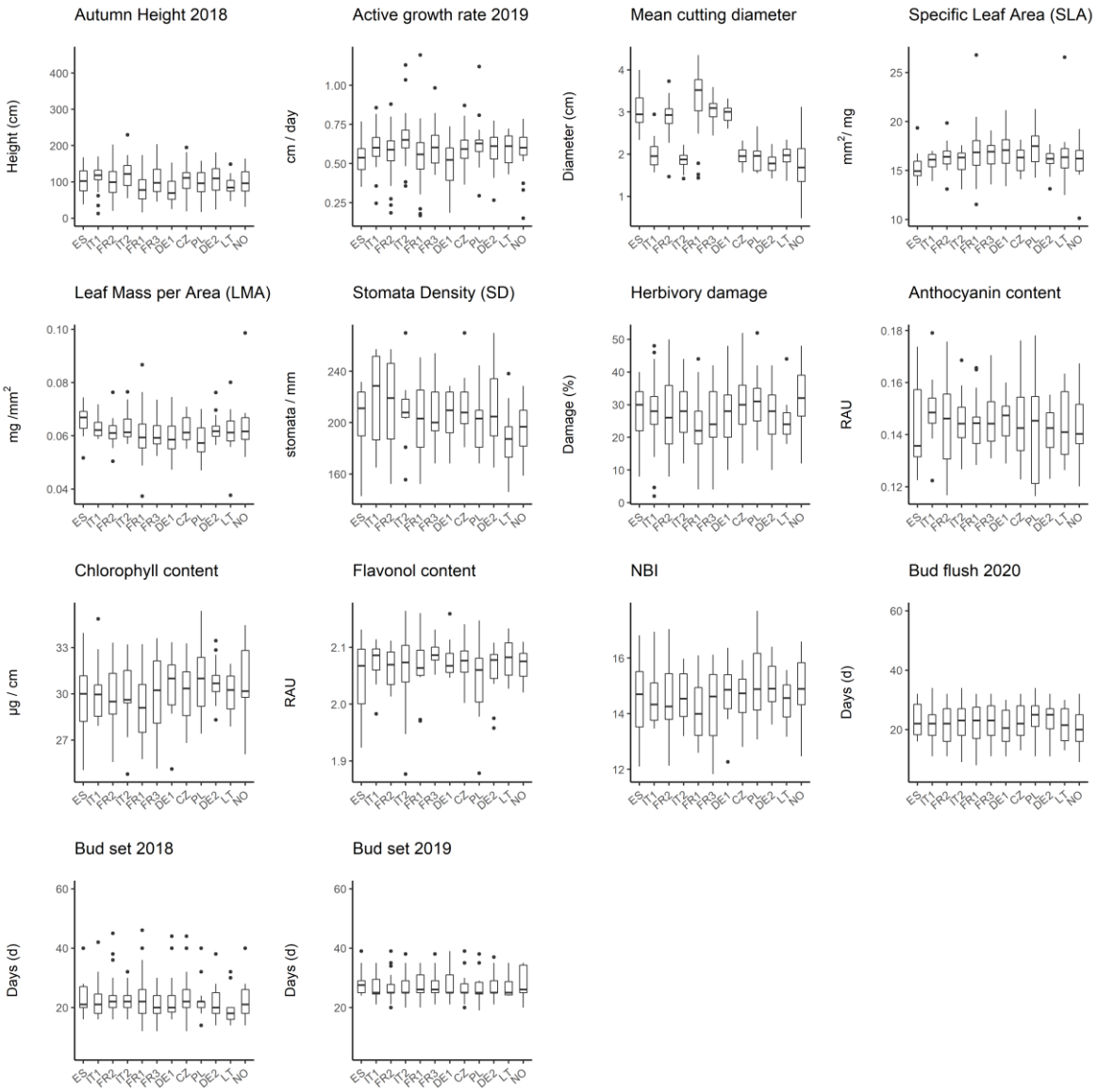
Supplementary Figure 1.

Phenotypic trait correlations. Pearson's Product-Moment Correlation coefficients (*r*) indicated among all growth, ecophysiology, biotic stress, leaf chemistry and phenology traits listed in Table 1, and climatic gradients (PC1 and PC2). Colored squares correspond with the coefficients of significant correlations.



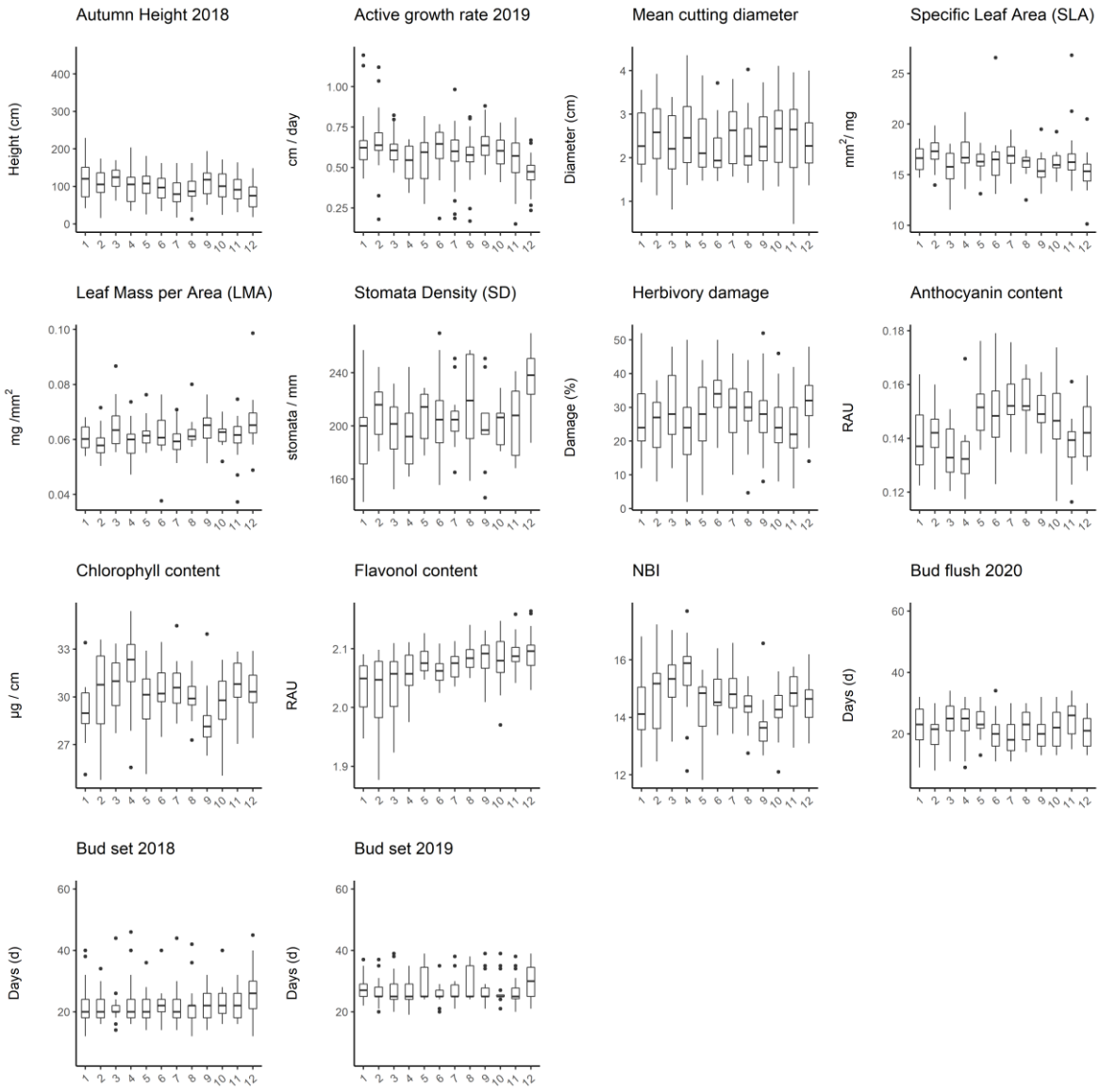
Supplementary Figure 2.

Phenotypic variation in traits where no significant differences between geographic origins were found. Boxplots are ordered by site, from south to north according to their geographic coordinates and labelled by the sample site code (ISO 3166 standard country code): 1. Spain; 2. Italy 1; 3. France 2; 4. Italy 2; 5. France 1; 6. France 3; 7. Germany 1; 8. Czech Republic; 9. Poland; 10. Germany 2; 11. Lithuania; 12. Norway.



Supplementary Figure 3.

Phenotypic variation in traits where no significant differences between geographic origins were found. Boxplots are ordered by common garden block.



Supplementary Table 1. Description and range of all bioclimatic (Bioclim) variables included in the PCA analysis, and their contributions to principal component 1 (% PC1, total variation explained = 68.1%) and Principal Component 2 (% PC2, total variation explained= 13.9%). Bioclimatic data for each of the locations of the parental clones were obtained from the CHELSA Timeseries data set.

BIOCLIM	Description	% PC1 (68.1 %)	% PC2 (13.9 %)	Range
BIO 1	Annual Mean Temperature (°C)	9.31	2.47	6.42 - 15.24
BIO 2	Mean Diurnal Range (°C)	2.98	7.05	5.45 - 9.18
BIO 3	Isothermality (%)	5.48	2.03	21.32 - 33.96
BIO 4	Temperature Seasonality (°C)	5.38	0.08	526.51 - 790.63
BIO 5	Max Temperature of Warmest Month (°C)	7.36	4.67	21.02 - 30.29
BIO 6	Min Temperature of Coldest Month (°C)	9.50	1.25	-6.64 - 4.29
BIO 7	Temperature Annual Range (°C)	0.99	2.03	22.82 - 29.79
BIO 8	Mean Temperature of Wettest Quarter (°C)	3.45	1.74	6.81 - 19.42
BIO 9	Mean Temperature of Driest Quarter (°C)	8.92	1.61	-3.61 - 24.15
BIO 10	Mean Temperature of Warmest Quarter (°C)	7.55	3.61	16.67 - 24.75
BIO 11	Mean Temperature of Coldest Quarter (°C)	9.90	1.54	-3.61 - 7.61
BIO 12	Annual Precipitation (mm)	4.56	11.82	391.45 - 951.67
BIO 13	Precipitation of Wettest Month (mm)	1.36	8.96	45.22 - 119.41
BIO 14	Precipitation of Driest Month (mm)	2.90	13.11	12.00 - 54.16
BIO 15	Precipitation Seasonality (%)	3.52	2.74	10.65 - 47.23
BIO 16	Precipitation of Wettest Quarter (mm)	1.49	7.85	128.46 - 332.24
BIO 17	Precipitation of Driest Quarter (mm)	3.76	12.40	39.00 - 172.48
BIO 18	Precipitation of Warmest Quarter (mm)	2.40	11.76	39.00 - 248.59
BIO 19	Precipitation of Coldest Quarter (mm)	9.19	3.11	63.20- 287.06

Supplementary Table 2. List and descriptions of all traits measured under common garden conditions, and the units in which each trait was measured.

Trait	Description
Tree growth	
Cutting diameter (cm)	Diameter of planted cuttings
Height Autumn (cm)	Tree height from ground to apical shoot at the end of each growing season
Growth rate (cm day^{-1})	Ratio between height gain and number of days in the growth period
Growth period (d)	Number of days between budflush stage 5 and budset stage 2.5
Stems (#)	Number of stems at the end of the growing season
Ecophysiology	
SLA (mm^2 mg^{-1})	Ratio between leaf area and leaf dry weight
LMA (mg mm^{-2})	Inverse of SLA
Stomatal Density (mm^{-2})	Number of stomata per mm^2
Biotic stress	
Herbivory damage (%)	Percentage of damage caused by herbivores
Level of rust infection (SU)	Level of infection on a scale of 1-60, calculated from leaf-level damage and tree-level damage.
Leaf Chemistry	
Chlorophyll (µg cm^{-2})	Chlorophyll content based on UV optical absorbance measurements
Flavonols (RAU)	Flavonol content based on UV optical absorbance measurements
Anthocyanins (RAU)	Anthocyanin content based on UV optical absorbance measurements
Nitrogen Balanced Index	Ratio of chlorophyll and flavonol content
Phenology	
Budset (d)	Number of days between stage 2.5 and stage 0 of the bud set period
Budflush (d)	Number of days between stage 2 and stage 5 of the bud flush period

Supplementary Table 3. Coefficients for the climatic model (PC1 + PC2 as fixed factors) and geographic model (Site of origin as fixedfactor) included in the best linear mixed effect models. Best models were chosen according to marginal and conditional R^2 values. The effect of microclimatic conditions in the common garden was included as a random effect (common garden block) in all models. Cutting diameter or tree height (in Autumn 2018 or 2019) were included as fixed effects. Marginal R^2 values provide the variance explained only by fixed effects, and conditional R^2 values provide the variance explained by the whole model. P-values for models that used "Site" as a fixed effect are not reported in the table because "Site" is a factor variable with multiple levels. Significant differences between pairs of sites can be found in Figure 2. Significance codes: $p < 0.001$ = ***; $p < 0.01$ = **; $p < 0.05$ = *

Trait Class	Phenotypic trait	Model	Fixed effects			Model R^2	
			Fixed 1	Fixed 2	Fixed 3	R^2 marginal	R^2 conditional
Biomass	Height Autumn 2018	Climatic	PC1	PC2	Cutting ***	0.058	0.139
		Geographic	Site	Cutting *	-	0.093	0.172
	Height Autumn 2019	Climatic	PC1	PC2	Cutting ***	0.061	0.170
		Geographic	Site *	Cutting	-	0.102	0.206
	Height Autumn 2020	Climatic	PC1	PC2	Cutting ***	0.064	0.192
		Geographic	Site *	Cutting		0.113	0.234
	Growth rate 2019	Climatic	PC1	PC2 *	Cutting **	0.036	0.116
		Geographic	Site	Cutting *	-	0.060	0.144
	Number of stems	Climatic	PC1 *	PC2	Cutting	0.015	0.065
		Geographic	Site *	Cutting	-	0.065	0.115

Supplementary Table 3. Continued

Trait Class	Phenotypic trait	Model	Fixed effects			Model R^2	
			Fixed 1	Fixed 2	Fixed 3	R^2 marginal	R^2 conditional
Ecophysiology	Leaf Mass Area	Climatic	PC1	PC2	Height 19 *	0.043	0.092
		Geographic	Site *	Height 19		0.083	0.121
	Specific Leaf Area	Climatic	PC1	PC2	Height 19 **	0.053	0.084
		Geographic	Site *	Height 19 *	-	0.098	0.121
	Stomata Density	Climatic	PC1	PC2	Height 19	0.030	0.139
		Geographic	Site	Height 19	-	0.077	0.168
Biotic stress	Herbivory damage	Climatic	PC1 **	PC2	Height 19 *	0.035	0.104
		Geographic	Site	Height 19	-	0.061	0.126
Leaf chemistry	Anthocyanins	Climatic	PC1	PC2	Height 19	0.021	0.245
		Geographic	Site	Height 19 *		0.043	0.258
	Chlorophyll	Climatic	PC1	PC2	Height 19	0.024	0.151
		Geographic	Site	Height 19	-	0.046	0.173
	Flavonols	Climatic	PC1	PC2	Height 19 ***	0.158	0.277
		Geographic	Site *	Height 19 ***	-	0.199	0.311

Supplementary Table 3. Continued

Trait Class	Phenotypic trait	Model	Fixed effects			Model R^2	
			Fixed 1	Fixed 2	Fixed 3	R^2 marginal	R^2 conditional
Leaf chemistry	Nitrogen Balanced Index	Climatic	PC1	PC2	Height 19 *	0.053	0.185
		Geographic	Site	Height 19 *	-	0.081	0.208
Phenology	Duration bud set 2018	Climatic	PC1	PC2	Cutting	0.008	0.036
		Geographic	Site	Cutting	-	0.022	0.053
	Duration bud set 2019	Climatic	PC1	PC2	Cutting	0.009	0.045
		Geographic	Site	Cutting	-	0.020	0.056
	Duration bud flush 2019	Climatic	PC1	PC2 *	Height 18 ***	0.058	0.144
		Geographic	Site ***	Height 18 ***		0.093	0.171
	Duration bud flush 2020	Climatic	PC1	PC2	Height 18 ***	0.052	0.112
		Geographic	Site	Height 18 ***		0.065	0.126

CHAPTER III

Effects of environmental conditions on DNA methylation

Cristian Peña Pontón, Bárbara Díez Rodríguez, Paloma Pérez Bello, Claude Becker, Lauren M. McIntyre, Wim van der Putten, Emanuele De Paoli, Katrin Heer, Lars Opgenoorth and Koen J.F. Verhoeven

This chapter has been submitted to *New Phytologist*

Abstract

DNA methylation is environment-sensitive and can mediate plant stress responses. In long-lived trees, changing environments might cumulatively shape the methylome landscape over their lifetime. However, because high-resolution methylome studies usually focus on single environments, it remains unclear to what extent the methylation responses are generic or stress-specific, and how this relates to their long-term stability. Here, we studied the methylome plasticity of a single poplar genotype, *Populus nigra* cv. 'Italica'. Adult poplar trees with diverse environmental histories were clonally propagated, and the ramets exposed to experimental cold, heat, drought, herbivory, rust infection and salicylic acid treatments. Then, we identified and compared stress-induced vs. naturally occurring DNA methylation changes using whole genome bisulfite sequencing data. Methylation changes mainly targeted transposable elements and when occurring in CG/CHG contexts, the same regions were often affected by multiple stresses, indicating a generic response. Drought triggered a unique CHH hypermethylation response in transposable elements, affecting entire superfamilies and often occurring near drought-responsive genes. Stress-induced methylation variation in CG/CHG contexts showed striking overlap with methylation differences observed between trees from distinct geographical locations. Altogether, our results indicate that generic methylome stress responses can persist as epialleles in nature while some environments trigger more transient but large and specific responses, with possible functional consequences.

Key words: *Populus nigra* cv. 'Italica' (Lombardy poplar), whole genome bisulfite sequencing (WGBS), differentially methylated region (DMR), abiotic stress, biotic stress, drought responsive transposable elements, short interspersed nuclear element (SINE), drought, heat, cold, salicylic acid, rust infection, *Melampsora larici-populina* Kleb, caterpillar, *Lymantria dispar* L. (gypsy moth)

Introduction

Plants are challenged by abiotic and biotic stresses that affect their survival, growth, and fitness. Long-lived trees must acclimate to simultaneous and seasonal stress exposures every year by employing diverse genetic and epigenetic strategies for regulation of plant growth, development, and reproduction. However, although the role of epigenetic mechanisms in stress responses is receiving increasing attention (Deleris *et al.,* 2016; Lämke & Bäurle, 2017; H. Zhang *et al.,* 2018), most knowledge comes from short-lived, annual species (Hagmann *et al.,* 2015; Kenchanmane Raju *et al.,* 2018; Wibowo *et al.,* 2016). Moreover, as mitotically stable epigenetic marks have the potential to mediate plant responses to environmental changes (Becker *et al.,* 2011; Boyko *et al.,* 2010; López Sánchez *et al.,* 2016; Schmitz *et al.,* 2011), epigenetic research in perennials may be the key for understanding such roles over various time scales.

DNA methylation is the most abundantly studied epigenetic modification; it generally refers to cytosine methylation (5mC), the addition of a methyl group to the fifth position of the pyrimidine ring of a cytosine base. In plant genomes, 5mC occurs frequently in all three sequence contexts: the symmetric CpG and CHG along with the asymmetric CHH contexts (where H = A, T or C) (X. Zhang *et al.,* 2006). Insights about the functionality of DNA methylation have been described in promoters, gene bodies and transposable elements (TEs). In promoter regions, methylation usually inhibits transcription initiation, while its function within the gene body is less clear (Bewick & Schmitz, 2017; Paszkowski & Whitham, 2001; H. Zhang *et al.,* 2018; X. Zhang *et al.,* 2006) but may act to quantitatively impede transcript elongation (Zilberman *et al.,* 2007). TEs are enriched for DNA methylation and histone modifications, which are associated with transcriptional silencing (Matzke & Mosher, 2014).

De novo methylation in all sequence contexts is directed by small RNAs and catalysed by DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2) in a process known as RNA-Directed DNA Methylation (RdDM) (Matzke & Mosher, 2014). Maintenance of CpG, CHG and CHH methylation are performed by METHYLTRANSFERASE 1 (MET1), CHROMOMETHYLASE 3 (CMT3)/CMT2, and DRM2/CMT2, respectively. Lack of DNA methyltransferase activity or methyl donor shortage following DNA replication result in passive DNA demethylation, while active DNA demethylation involves the activity of glycosylases which excise 5mC from all cytosine sequence contexts (H. Zhang *et al.,* 2018).

While it is well-established that DNA methylation is responsive to environmental factors (Liu & He, 2020), its role in mediating environmental plasticity is less well understood.

For instance, causality between induced methylation variation and modulation of gene expression is still a matter of debate (Bewick *et al.,* 2019; Secco *et al.,* 2015; Seymour & Gaut, 2020). Moreover, many plant stress responses are mediated by systemic signalling via hormones such as salicylic acid, jasmonic acid, auxin, ethylene, and abscisic acid (Karpiński *et al.,* 2013; Wang *et al.,* 2014). However, it is still undetermined whether different stresses could induce common methylation responses that can be explained by overlapping systemic signalling responses.

The cost of high-resolution DNA methylation analyses has been a major factor limiting the sample size within studies, leading to low statistical power for identifying DNA methylation variants. More importantly, this has also led to the study of single environmental factors at a time, restricting the comparison of identified responsive loci among experiments. Hence, the present view is that DNA methylation responses can be highly specific with respect to environmental conditions, which might underestimate the possibility of overlapping responses that generally occur under a variety of stresses.

To study molecular mechanisms of trees in response to environmental cues, *Populus* species have become a choice model system due to their rapid growth, easy propagation, and available genomic resources (Jansson & Douglas, 2007; Tuskan *et al.,* 2006). Moreover, poplars are riparian species that are among the woody plants most sensitive to water stress (Larchevêque *et al.,* 2011; Rood *et al.,* 2003). This prompted much research on molecular mechanisms of drought tolerance (Jia *et al.,* 2016; Viger *et al.,* 2016; Yıldırım & Kaya, 2017), including DNA methylation (Lafon-Placette *et al.,* 2018; Sow *et al.,* 2021). *Populus. nigra* cv. 'Italica', also known as the Lombardy poplar, is one of the most widely distributed poplar cultivars that was first reported in Lombardy, Italy, at the very beginning of the eighteenth century (Chenault *et al.,* 2011). Afterwards, it was introduced into France, from where it is believed that Napoleon promoted its spread across the Empire by clonal propagation (Stettler, 2009). As a result, Europe has become colonized by a genetically homogeneous male clone of the Lombardy poplar. The clonality of this cultivar makes it an excellent system for studying epigenetic plasticity in response to different environmental conditions, as it strongly limits the confounding effects of genetic variability (Díez-Rodríguez *et al.,* 2022).

Here, we used whole genome bisulfite sequencing to capture the methylation responses of young poplars, clonally propagated from adult trees from different European locations, to a variety of biotic and abiotic stresses. With this unique approach, we aimed to characterize the environmentally induced methylome variation of the Lombardy poplar and to examine methylation variation as induced by exposure to acute stress in comparison with methylation variation that had built up naturally between trees during

the lifetime of growth in different geographic locations. Our study revealed commonalities between the methylation responses to different stresses, very specific responses to some stresses, and hints at the long-term stability of such responses.

Materials and methods

Our experimental approach involved whole genome evaluation and multiple treatments. Here, we provide a brief overview of the employed methods while a detailed methodology for each section can be found as supporting information.

Plant material

Cuttings from eight adult *Populus nigra* cv. 'Italica' clones were collected from five European countries (Table S1, Figure 1A; see (Díez-Rodríguez *et al.,* 2022). At each site, at least seven hardwood cuttings of approximately 30 cm length were sampled from each adult parental tree (ortet) and stored at 4 °C for two weeks prior to planting. Cuttings were grown under controlled conditions for 12 weeks until the start of the experiment. Growth conditions were: 22/18 °C (±2°C) at day/night, 60% relative humidity (±5% Rh), 16/8 h light/dark. Cuttings were planted first in 4-liter pots with a 3:1 sand:peat mixture (v/v) and placed in a flood table for three weeks. Rooted cuttings (ramets) of similar size were transferred to 7-liter pots with a 1:1 sand:peat mixture and maintained with regular watering. Two weeks prior to the start of the experiment, three grams of slow-release fertilizer Osmocote Exact Mini (16+8+11+2MgO+TE) were added to each pot.

Experimental design

Seven 3-month-old ramets of similar size were selected from each of the eight (ortets) for subsequent exposure to different environmental treatments (56 trees in total). The experiment consisted of three biotic and three abiotic stresses plus a single control group, with eight replicates per treatment, where each ortet contributed one replicate to each of the treatments. All treatments were implemented simultaneously during a period of 25 days, including two 10-day stress events and one 5-day stress-free period in between (Figure S1).

Control group: during the entire stress experiment, control plants were maintained in a greenhouse under controlled conditions as stated above. The soil volumetric water content (VWC) was maintained on average at 20.2% (± 3.24 SD) by daily watering to pot

capacity. VWC was monitored daily using the WET Sensor kit (Delta-T Devices). Mean VWC was calculated for all plants with two measurements per pot (Figure S2). During the entire experiment, control plants were maintained in the same greenhouse table with other stress treatments in a Latin square design unless otherwise stated (Figure S2).

Biotic stresses: Rust infection consisted in spray-inoculation of uredospores of the poplar leaf rust fungus (*Melampsora larici-populina* Kleb.). Herbivory treatment involved the use of Gypsy moth caterpillars (*Lymantria dispar* L.). Salicylic acid (SA) treatment consisted in spray application of 1 mM SA (Sigma-Aldrich). Rust spores and caterpillars were obtained from Dr. Sybille Unsicker (Max Planck Institute for Chemical Ecology, Jena, Germany).

Abiotic stresses: Drought stress was attained by withholding watering and maintaining VWC at 8%. Plants that received heat treatment were grown at high temperatures of 30-38/28°C (day/night), while cold-treated plants were grown at 4/4°C (day/night).

VWC for all treatments (except drought) was maintained close to control conditions (Fig. S3a). Plants were moved to climate chambers for heat and cold treatments (Fig. S3b). During the stress-free period, plants were moved to the same greenhouse table as the control group (Fig. S2).

Harvesting: For DNA methylation analysis, on experimental day 26, twelve circular punches (Ø 8 mm; ~ 100 mg fresh weight in total) were cut out from the eighth mature leaf (counting from the apex of the main branch, leaf plastochron index: 10) of each plant. Mid-ribs were avoided, and leaf punches were immediately frozen in liquid nitrogen and stored at −80 °C. All sampled leaves were not directly exposed to any of the biotic stresses; thus, we characterized the systemic response.

DNA extraction and Whole Genome Bisulfite Sequencing

Per sample, frozen leaf tissue was grinded and homogenized using TissueLyser II (QIAGEN), then genomic DNA was isolated using the Sodium Dodecyl Sulfate (SDS) procedure of the NucleoSpin Plant II DNA isolation kit (Macherey-Nagel, Dueren, Germany). Preparation of DNA libraries for bisulfite sequencing was performed as described in (Nunn *et al.,* 2022). All sequencing was performed by Novogene on an Illumina HiSeq X Ten sequencing system. Libraries were sequenced with 2x150-bp paired-end reads at 30X coverage. Libraries were sequenced in a total of eight sequencing lanes, trying to allocate ramets derived from the same ortet in the same lane to avoid batch effects.

Processing of bisulfite-treated reads and methylation calling

Sequenced reads were processed using the EpiDiverse Toolkit (WGBS pipeline v1.0, https://github.com/EpiDiverse/wgbs) (Nunn *et al.,* 2021). Briefly, low-quality read-ends were trimmed (minimum base quality: 20), sequencing adapters were removed (minimum overlap: 3 bp) and very short reads (< 36 bp) were discarded. The remaining high-quality reads were aligned to the *Populus nigra* var. 'Italica' *de novo* reference genome (ENA project: PRJEB44889) using erne-bs5 (http://erne.sourceforge.net) allowing for 600-bp maximum insert size, 0.05 mismatches, and unique mapping. Per-cytosine methylation metrics were calculated using MethylDackel (https://github.com/dpryan79/MethylDackel). Three bedGraph files per sample were obtained, corresponding to cytosines on each sequence context: CpG, CHG and CHH. The methylation level (%) of a particular site was calculated by:

$$\frac{methylated\ cytosine\ read\ count}{methylated\ cytosine\ read\ count\ +\ unmethylated\ cytosine\ read\ count} * 100$$

Methylation analysis

The analysis followed three stages: 1) genome-wide methylation analysis: to detect strong global methylation patterns by comparing samples and groups to each other, 2) differential methylation analysis: to identify significant differentially methylated regions (DMRs) by testing for significant methylation differences between groups throughout the entire genome, and 3) downstream analysis: to reveal potential functional implications by testing annotated DMRs for enrichment on genomic features and gene functions. The three cytosine sequence contexts were always analysed separately. Data filtering and resolution was slightly different for each analysis (Table S2). As a first general filtering step, all cytosines with low sequencing coverage (≤ 5 reads) were removed

Genome-wide methylation analysis

After the first filtering step, samples in which the retained cytosines accounted for less than 50% of the original data were considered low-coverage outliers and were excluded from genome-wide analyses unless otherwise stated (4 outliers out of 56 samples, Supplementary file 1, Table S2).

Average global methylation

Only genomic positions with methylation information across all 52 remaining samples were considered (CpG: 1'802.288 positions, CHG: 3'256.938 positions, CHH: 12'058.984 positions). For each sample, average global methylation (%) was calculated separately for each context, as follows:

$$average\ global\ methylation\ (\%) = \frac{\sum\ (per\ cytosine\ methylation\ \%)}{Total\ number\ of\ analyzed\ cytosines}$$

The effect of the stress treatments on the level of global methylation was evaluated for each context using a linear mixed model with treatment as fixed factor and parental tree (ortet) as random factor. Statistical analyses were calculated in R (version 4.0.3), the *lmer* function of the *lmerTest* package (Kuznetsova *et al.,* 2017) was used to fit the model, and multiple comparisons (Tukey's post-hoc tests) were calculated with the *glht* function of the *multcomp* package (Hothorn *et al.,* 2008).

Principal component analysis, hierarchical clustering, and correlation analysis

For CpG and CHG context, the same filtered data used for average global methylation analysis was considered for principal component analysis (PCA) and hierarchical clustering (HC). As we observed that a very large fraction of cytosines in CHH context showed very low methylation variation across samples, an extra filtering step removed CHH positions where more than 90% of the samples showed very low (0-5%) or very high (95-100%) methylation to keep the more variable positions.

Principal components (PCs) were calculated in R using the *prcomp* function of the *stats* package (R Core Team, 2022). HC (Ward's method) was computed by first calculating the corresponding distance matrix (Manhattan method) using the *dist* and *hclust* functions from the *stats* R package.

Pairwise correlation analysis was performed using genomic regions instead of single positions. First, the poplar genome was compartmentalized in 100-bp non-overlapping bins. Average methylation per bin was calculated, and only bins with methylation information across all samples were retained. As the intraclass correlation coefficient (ICC) reflects both degree of correlation and agreement between measurements (Koo & Li, 2016), ICC was calculated for all pairwise comparisons between samples.

Coefficients were calculated in R using the *icc* function of the *irr* package (Gamer *et al.,* 2019) with the ICC form: two-way random effects, absolute agreement, single measurement, according to (McGraw & Wong, 1996) convention.

Methylation profiles

All (56) samples were included in the analysis. For gene regions, only protein-coding genes with known 5'UTR and 3'UTR coordinates were considered. For transposable elements, only TEs longer than 150 bp were analyzed. Methylation profiles over the largest poplar scaffold (scaffold 1 = 33'746.648 bases) were used as a proxy for chromosome-wise methylation variation comparison. The scaffold was compartmentalized in 50-kb bins, then for each sample, per-bin average methylation was calculated. Finally, to calculate per-bin methylation for each treatment, a weighted mean was calculated accounting for the number of cytosines per bin per sample using the formula:

$$weighted\ mean\ methylation\ for\ bin\ X\ (\%) = \frac{\sum_k^m\ (methylation\ \%\ bin\ X_k) * (\#Cs\ bin\ X_k)}{\sum_k^m\ \#Cs\ bin\ X_k}$$

where ***X*** represents any 50-kb bin, ***k*** is a replicate, and ***m*** is the total number of replicates per treatment. For each treatment and context, per-bin methylation differences compared to the control group were calculated and used to plot heatmaps and simple moving averages (SMA). SMAs were calculated using the R function *geom_ma* of the *tidyquant* package (Dancho, 2022).

Differential methylation analysis

Differentially methylated regions (DMRs) induced by each stress treatment were identified by testing local methylation differences between each treatment and control group. All replicates per treatment were included in the tests and each cytosine sequence context was analysed separately using the EpiDiverse/DMR pipeline v0.9.1 (https://github.com/EpiDiverse/dmr) (Nunn *et al.,* 2021). Briefly, DMRs were identified by metilene (https://www.bioinf.uni-leipzig.de/Software/metilene), with parameters as follows. Minimum read depth per position: 6; minimum cytosine number per DMR: 10; minimum distance between two different DMRs: 146 bp; per-group minimal non-missing data for estimating missing values: 0.8; adjusted p-value (Benjamini-Hochberg) to detect

significant DMRs: 0.05. Only significant DMRs with minimum methylation difference of 10 percentage points between groups were used for downstream analyses.

DMR calling

Since the genome-wide methylation analyses revealed strong CpG and CHG methylation patterns associated to sample origin (ortet identity) irrespective of stress treatment, stress-DMRs were identified using a jack-knife approach (leave-one-out) to reduce within-treatment variation caused by individual outlying ortets. As eight different ortets were included in the experimental design, a total of eight DMR calls were performed, in which samples derived from single ortets were left out on each DMR call (Figure S4). Hence, seven replicates were included on each DMR calling. All identified DMRs were retained for further analysis. Additionally, to check if our jack-knife approach for DMR calling produced robust results, we called stress-DMRs using Methylkit (Akalin *et al.,* 2012) and intersected both results. This check indicated a large overlap between both methods and highlighted the conservative nature of our results (Table S3).

Based on overlaps among DMRs that were identified in more than one treatment, DMRs were classified as multi-stress or stress-specific DMRs. Genomic regions where a DMR was identified in more than one sequence context were labelled as multi-context DMRs.

We also performed DMR callings among ortets. Ramets derived from the same ortet and exposed to different treatments were considered replicates. Briefly, DMRs were called for all pairwise comparisons among the eight ortets (total: 28 DMR-sets per context). Next, for each context, all DMRs were classified according to the number of pairwise comparisons in which each DMR appeared in: DMRs found in a *unique comparison* or DMRs *shared by two or more comparisons*.

Downstream methylation analyses

DMR annotation

Statistically significant DMRs were annotated using the *Populus nigra* cv. 'Italica' protein-coding gene model annotation. Only the longest transcript per gene was used for this analysis. DMRs were also associated with TEs based on a TE prediction for this cultivar. Gene models and TE predictions used in this study were generated as part of the ongoing *P. nigra* cv. 'Italica' reference genome project (PRJEB44889). Short descriptions of these annotation files can be found along with their deposited versions (see Data availability statement). Short interspersed nuclear elements (SINEs) were manually added to the predicted TEs based on BLASTN results (70% similarity, 90%

coverage) using the consensus sequences of Salicaceae SINE families (Kögler *et al.,* 2020).

DMR enrichment on genomic features

For each context, all DMRs, irrespective of treatment, were tested for enrichment in gene bodies, exons, introns, gene flanking regions and TEs (Z-test for proportions). As DMRs were enriched in TEs, we also tested whether the occurrence of DMRs in gene bodies, exons, introns and gene flanking regions was conditional on the presence of TEs (Chi-square tests for independence, McNemar's test). Additionally, as drought showed the largest response associated to TEs, we calculated the relative fold enrichment for drought CHH-DMRs on each TE superfamily. P-values were obtained from the hypergeometric test and then adjusted (Bonferroni) according to the number of TE superfamilies tested. Drought-DMR-enriched TE superfamilies were referred to as Drought-Responsive TEs (DR-TEs), which included all SINE and MITE/DTHs elements.

$$fold\ enrichment = \frac{\frac{\#DMRs\ inside\ elements\ of\ TE\ superfamily\ X}{\#\ total\ TE-associated\ DMRs}}{\frac{total\ length\ of\ all\ elements\ of\ TE\ superfamily\ X\ (bp)}{total\ length\ of\ all\ TEs\ in\ the\ genome\ (bp)}}$$

Stress-induced methylation variation of DMRs and TEs

Since stress-DMRs were identified between each treatment vs. the control group, we were interested in evaluating the methylation response of such regions induced by all the other treatments. Using the genomic coordinates of the identified DMRs, we calculated average methylation levels of the corresponding regions on each sample Then, we calculated the average methylation level across treatment replicates. Finally, for each region, we computed the methylation difference between each treatment and control group. We only analysed regions with enough methylation information (≥ 8 cytosines) and replication (≥ 6 replicates).

Moreover, as we detected TE superfamilies enriched with drought-DMRs, we analysed the drought-induced methylation response of all poplar TE superfamilies to check for generalized responses of entire TE superfamilies. For each sequence context, average methylation levels of each individual TE were calculated for drought and control samples. We only analysed TEs with enough methylation information (≥20 cytosines for CHH, ≥10 cytosines for CpG and CHG) and replication (≥ 6 replicates). For each TE element, we computed the methylation difference between drought and control group. Then, to

summarize and compare results among TE superfamilies, we grouped TE elements in boxplots according to each superfamily.

Gene ontology enrichment analysis

Functional enrichment analysis has to be carefully interpreted as gene expression data was not collected in this experiment and most stresses produced very few DMRs. Therefore, our analysis was mainly focused on medium-to-large gene sets associated with drought-induced methylation responses.

Genes associated with drought CHH-DMRs were subjected to gene ontology (GO) enrichment analysis. The gene background was built with the closest Arabidopsis (*A. thaliana*) homologue of each *P. nigra* cv. 'Italica' gene, which was determined using BLAST best reciprocal hits (RBH) of the protein sequences (R package *orthologr* (Drost *et al.,* 2015). Best hits were filtered by keeping alignments covering at least 60% of both *Arabidopsis* and *P. nigra* proteins, and minimum 60% similarity. *Arabidopsis* protein sequences were extracted from phytozome V13, and functional annotations were retrieved from the PLAZA 5.0 dicots database (https://bioinformatics.psb.ugent.be/plaza/). GO enrichments were performed using clusterProfiler v4 (Wu *et al.,* 2021). P-values were adjusted for multiple testing controlling the positive false discovery rate (q-value).

Enrichments for genes associated with all SINE and MITE/DTH elements were performed in the same manner. Gene sets for functional enrichment included genes associated with either all SINEs, all MITE/DTHs, or both (DR-TEs). Additional enrichments were performed for a subset of potential Highly Drought-Responsive TEs (HDR-TEs), i.e., SINEs and MITE/DTHs that displayed at least 5% hypermethylation compared to the control group. Finally, for comparison, all enrichments were analysed in the context of the drought CHH-DMR gene set enrichment.

Results

Drought induces a large and distinctive genome-wide CHH hypermethylation response in the Lombardy poplar

Genome-wide methylation analyses were performed to identify overall strong methylation patterns among samples and treatments. Starting with average global methylation, the linear mixed models revealed significant treatment effects on DNA methylation (CpG: $p=0.048$, CHG: $p=0.043$, CHH: $p<0.01$). However, only drought treatment induced a significant global increase of CHH methylation compared to control group (Tukey's test $p<0.01$). In addition, cold treatment induced significantly higher CpG and CHG methylation levels compared to salicylic acid treatment ($p=0.0241$ and 0.0412, respectively) (Fig. S5a).

For CHH methylation, PCA and HC analyses highlighted noticeable clusters for drought and heat treatments (Fig. 1d, S6c, S7b). High correlations were observed among drought-treated samples, while the lowest correlation coefficients were found when comparing drought samples with any other sample (Figure S8c). Methylation profiles over the largest poplar scaffold confirmed the genome-wide drought-induced CHH hypermethylation and underlined a close relationship with TE content as both profiles showed peak similarities (Fig. 2, S9). Over TE regions, profiles corroborated the drought-induced CHH hypermethylation, and highlighted CHH hypomethylation induced by rust infection. Profiles over genic regions revealed that drought-induced CHH hypermethylation mainly targeted gene-flanking regions rather than gene bodies (Fig S10, S11).

The effect of other treatments was also observed in the methylation profiles. Profiles of CpG/CHG methylation along the scaffold 1 confirmed the genome-wide cold-induced hypermethylation that was already detected in the global methylation analysis. In addition, several treatments showed overlapping profiles of hypermethylation (drought, heat, and rust), and hypomethylation (SA and herbivory). Visual observation of the methylation profiles indicated a positive correlation between CpG and CHG methylation variation (Fig. 2, S9).

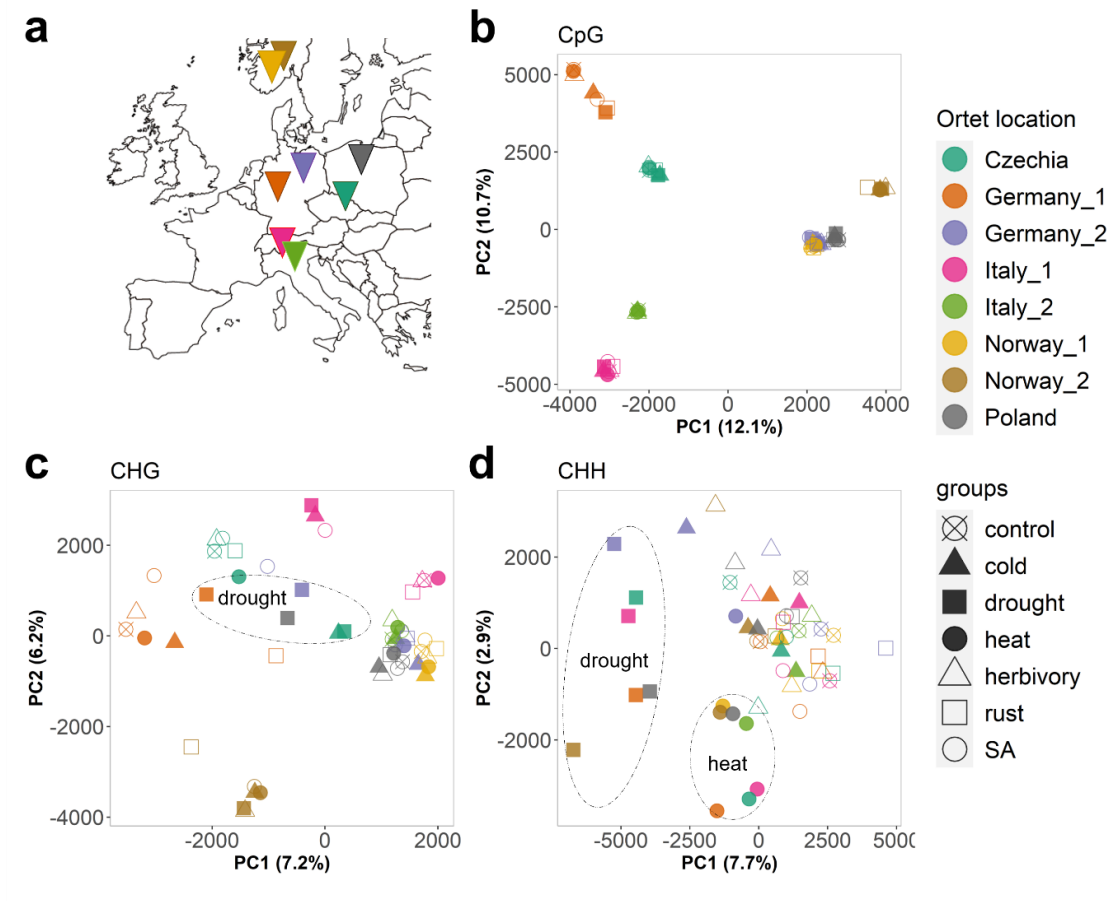


Figure 1. Analysis of genome-wide methylation patterns from stress-treated Lombardy poplar ramets. **a)** Sampling locations of the clonally propagated ortets used for the experiment. **b**, **c**, and **d)** Unsupervised principal component analysis of CpG, CHG and CHH methylation. Ramets are colored by ortet identity. Different shapes represent each experimental group. Drought and heat clusters are highlighted with dashed ovals.

Interestingly, over genic regions, the effects of stress treatments were detectable mainly in gene flanking regions while in TE regions, cold and SA induced the largest CpG/CHG methylation responses: hypermethylation and hypomethylation, respectively (Figure S11).

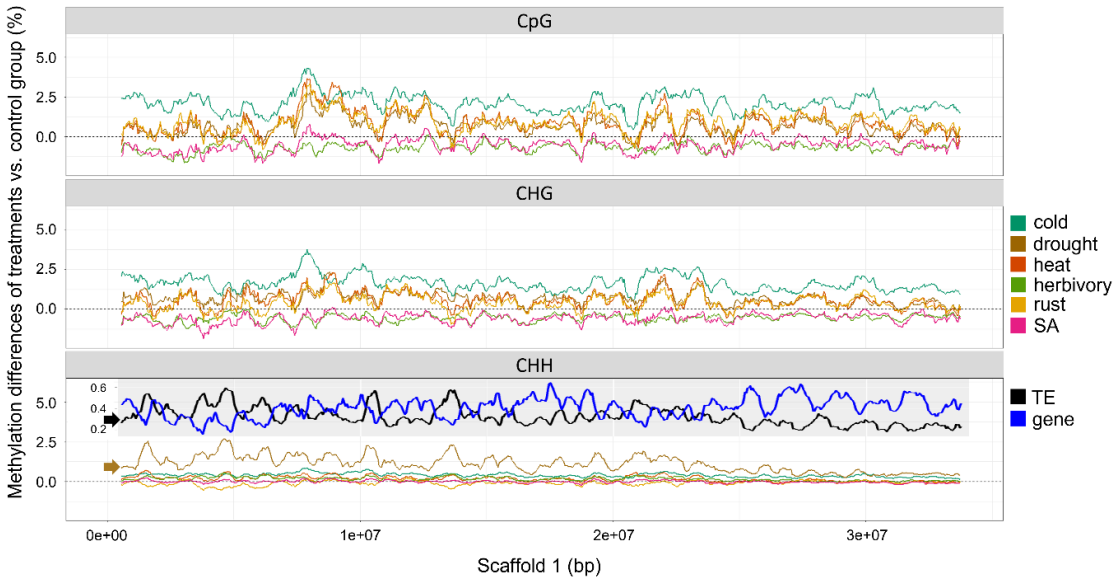


Figure 2. Metaplots of CpG, CHG and CHH methylation level differences (treatments vs control group) over the scaffold 1 of *Populus nigra* cv. Italica. Simple moving averages (SMA) over a period of ten 50-kb bins were calculated and plotted for each treatment and context. Profiles for TE and gene content were added on top of CHH metaplot (SMAs per ten 50-kb bin) to highlight the relationship between methylation variation and gene/TE content, especially obvious for drought-induced CHH methylation variation (brown arrow) and TE content profile (black arrow).

Genome-wide CpG and CHG methylation largely reflect sample origin rather than treatment effect

Linear mixed models revealed significant ortet effects on the average global CpG methylation ($p<0.01$) (Figure S5b). Detailed insights were observed on PCA, HC and correlation analysis where distances among ramets derived from the same ortet (within-ortet) were much smaller than those among ramets derived from different ortets (between-ortet), irrespective of the treatment (Figure 1b, S6a, S8a, S8d). A similar but less pronounced within-ortet clustering was found for CHG methylation, with an additional clustering of drought-treated ramets, irrespective of the ortet (Figure 1c, S6b, S7a, S8e).

Transposable elements are enriched with stress-induced DMRs

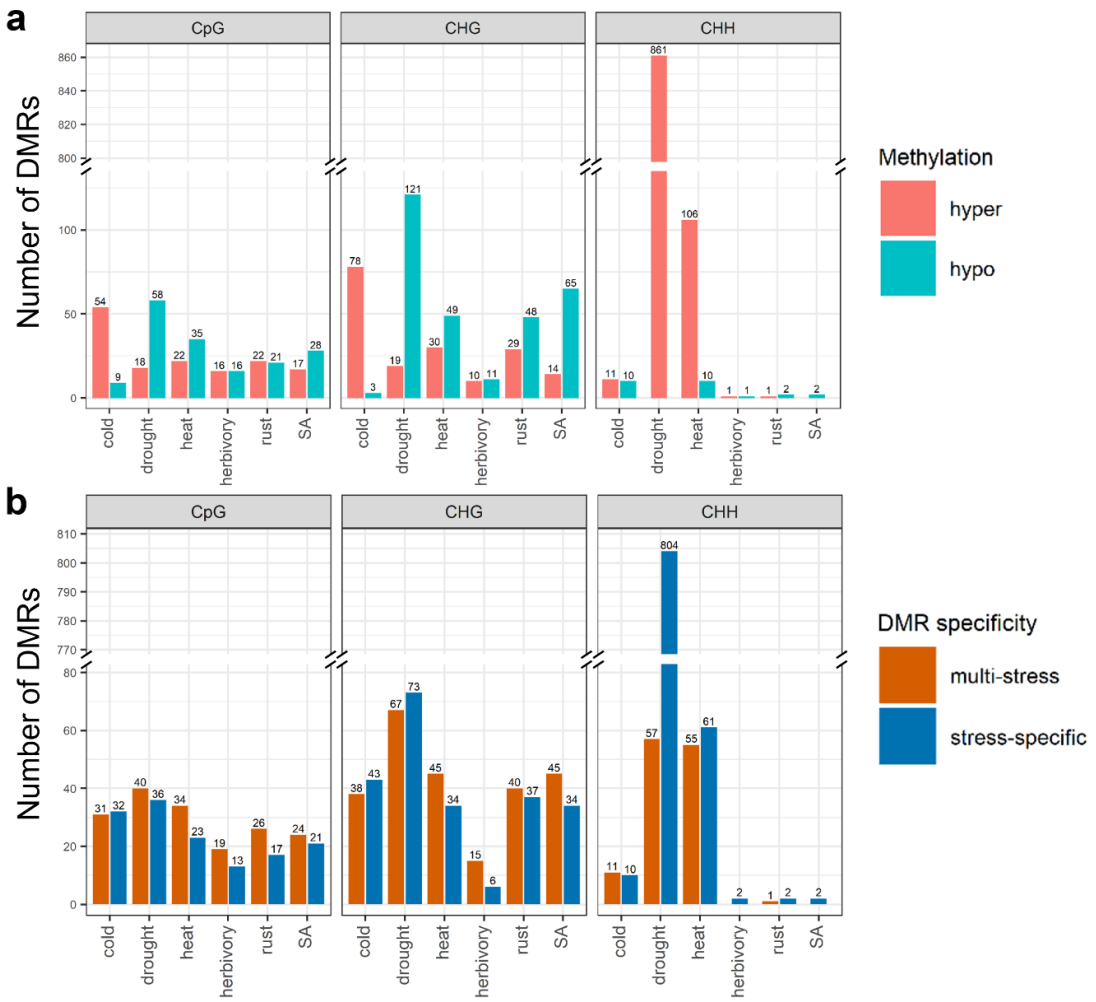


Figure 3. Summary of significant DMRs induced in the Lombardy poplar by each stress treatment versus the control group in CpG, CHG and CHH contexts. **a)** DMRs classified by methylation direction: hypermethylated (red), hypomethylated (blue) **b)** DMRs classified by specificity: multi-stress (orange), stress-specific (dark blue). DMR numbers are shown on top of each bar.

We identified a total of 1,798 DMRs across all treatments and sequence contexts (Fig. 3, table S4). Drought induced the largest number of DMRs among all treatments (mostly CHH hypermethylations), while cold induced the largest amount of hypermethylated DMRs in CpG/CHG contexts (Fig. 3a, table S4). In general, similar amounts of multi-stress and stress-specific DMRs were observed in each treatment, except for drought CHH-DMRs (Figure 3b, Table S5). Among the 203 multi-stress DMRs, drought and heat showed the largest intersection with 57 DMRs (Fig. S12).

Enrichment tests showed that all DMRs, irrespective of sequence context, were enriched in TEs. In addition, CpG-DMRs mainly targeted gene bodies, specifically exons, while CHH-DMRs were enriched over gene flanking regions (Fig. 4b, Table S6). CHG-DMRs were enriched in intergenic regions associated with TEs, while TE-associated CHH-DMRs were enriched in gene flanking regions and introns (Table S7). Moreover, we observed an increased frequency of DMRs in TE flanking regions, especially within the first 200 bp (Fig. S13), revealing TEs as a major source of methylation variation (DMRs) irrespective of treatment (Fig. S18, S19, S20).

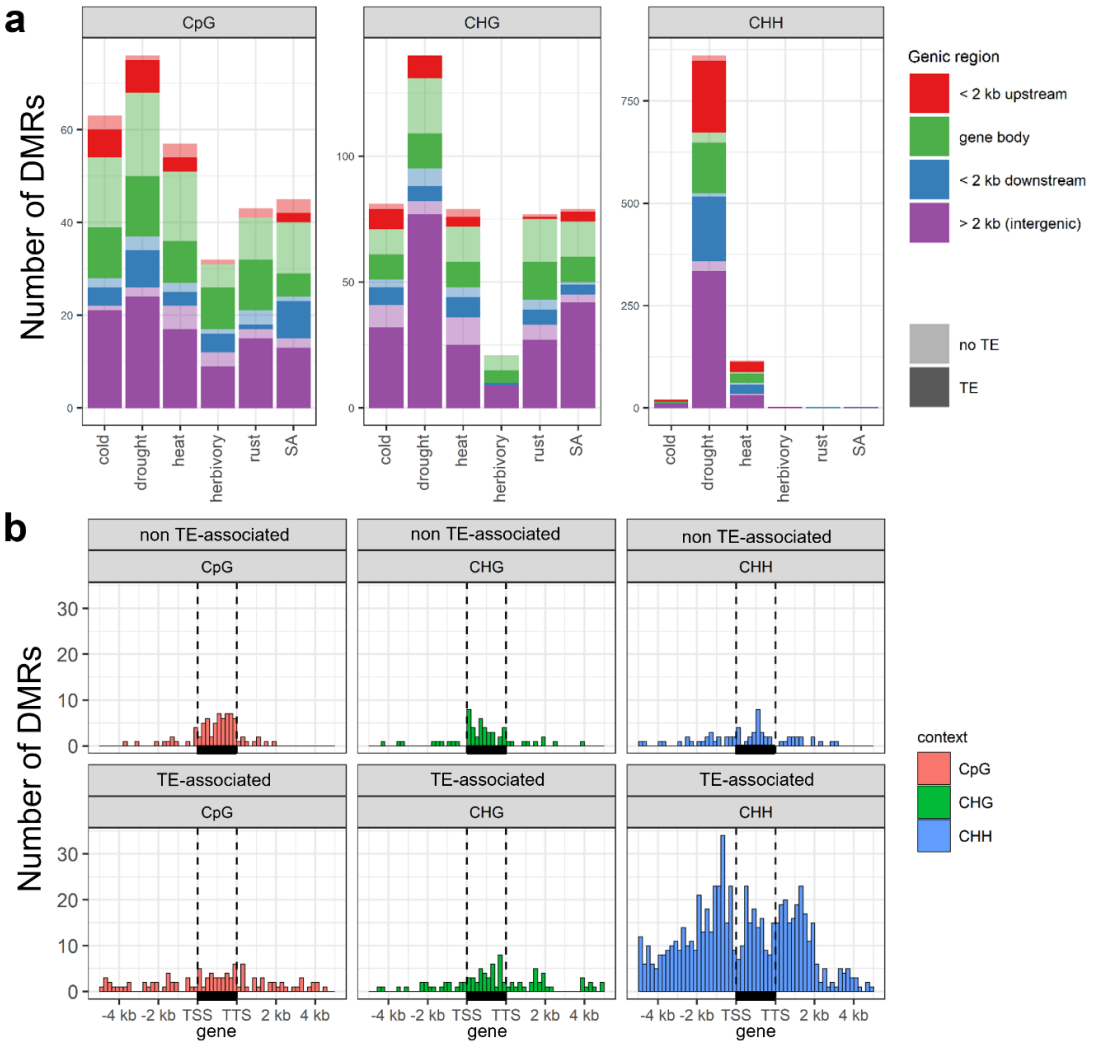


Figure 4. Distribution of stress-induced DMRs over the Lombardy poplar genome. **a)** For each treatment and context, DMR counts (irrespective of treatment) are shown for gene body, ±2kb gene flanking regions, and intergenic regions. Dark/light colors differentiate the number of DMRs associated/non-associated with TEs in the corresponding region. **b)** Detailed distribution of all

stress-induced DMRs along genic regions, per context and TE association. Vertical dashed lines indicate the gene transcription start site (TSS) and transcription termination site (TTS). Horizontal black boxes represent the gene body. Gene lengths were normalized to 2kb. Z-tests for proportions were performed based on the content of each genomic feature in the poplar genome (See Table S6 for complete results).

Drought-induced CHH hypermethylation is stronger on specific TE superfamilies

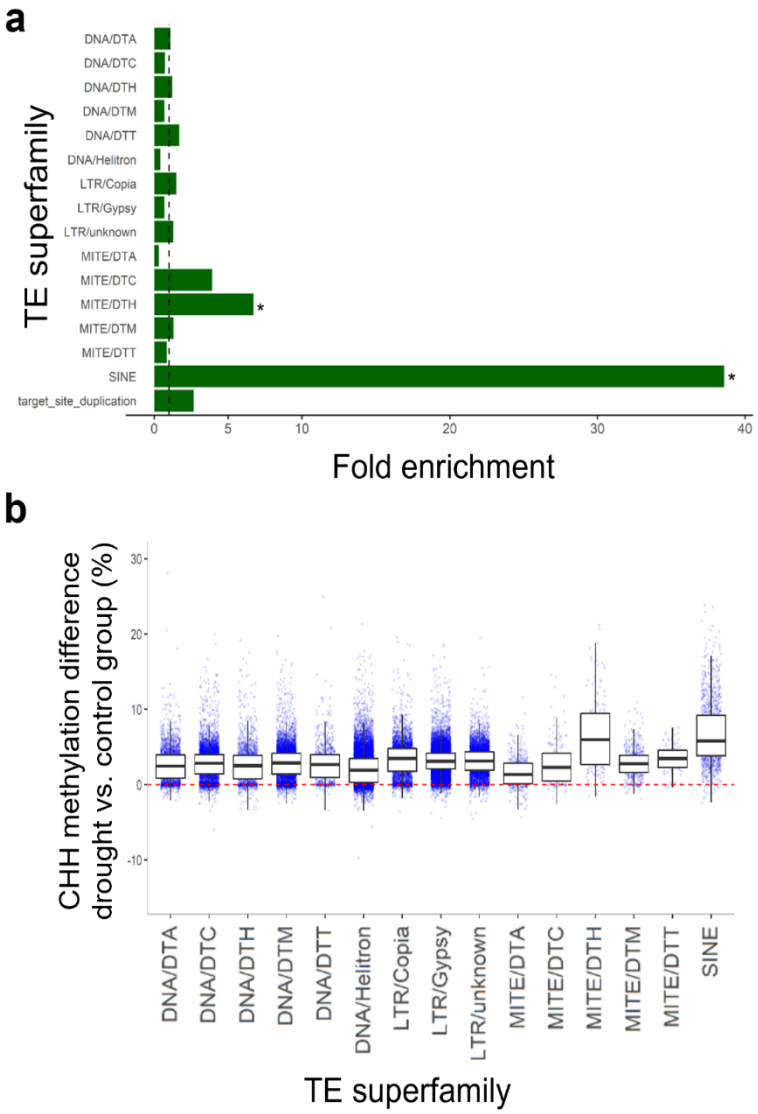


Figure 5. Analysis of drought-induced CHH hypermethylation of the Lombardy poplar TE superfamilies. **a)** Fold enrichment analysis of TE superfamilies targeted by drought CHH-DMRs. Enrichments were calculated based on the total length of each TE superfamily in the genome. Hypergeometric tests identified significant enrichments ($p<0.001$) for SINE and MITE/DTH superfamilies (Table S10). **b)** Boxplots of drought-induced CHH methylation variation (vs. control group) over individual TE elements. Each boxplot summarizes the overall methylation response of a specific TE superfamily (x-axis) to drought stress. Horizontal dotted red line depicts the zero drought-control difference.

DMR enrichments over each TE superfamily revealed that Short Interspersed Nuclear Elements (SINE) and Miniature Inverted-repeat Transposable Elements (MITE), especially MITE/DTH, showed an exceptionally strong response (Fig. 5, Table S8). Interestingly, SINE and MITE/DTH elements also display the highest CHH methylation under control conditions among all TE superfamilies (Figure S14). Methylation analysis over genic regions showed that drought induced CHH hypermethylation of SINE and MITE/DTH elements irrespective of gene proximity (Fig. S15 and S16, respectively).

CpG/CHG stress-DMRs are also multi-stress DMRs and ortet-DMRs

By examining the methylation response of stress-specific DMR regions in all other treatments, we observed that most of the stress-specific CpG/CHG-DMRs also showed a response to other treatments, usually in the same direction (either hyper- or hypo-methylation) (Fig. 6a, S17). Thus, different stresses tended to result in similar methylation responses at these genomic locations, even when statistical significance was only reached in response to some treatments. Moreover, by examining the methylation level of these responsive regions in the control group, we noticed that CpG/CHG-DMRs had intermediate CpG/CHG methylation and low CHH methylation, while CHH-DMRs showed very high methylation in all contexts (Fig. 6b).

The amount of ortet-DMRs was several orders higher than the stress-DMRs, especially in CpG/CHG context. For each comparison between two ortets, we identified on average 1,425 CpG-DMRs, 1,621 CHG-DMRs and 133 CHH-DMRs (Table S9). We detected a total of 9,840 CpG-DMRs, 7,353 CHG-DMRs and 1,141 CHH-DMRs after accounting for DMRs found in more than one pairwise comparison (Table S10). Such ortet-DMRs were considered as a product of natural methylation variation, and when intersected with stress-DMRs, the analysis revealed that most of the stress CpG-DMRs (71%) and CHG-

DMRs (85%) were also identified as ortet-DMRs. However, only 6% of stress CHH-DMRs were found in the ortet-DMR set. (Fig. 6c, Table S10).

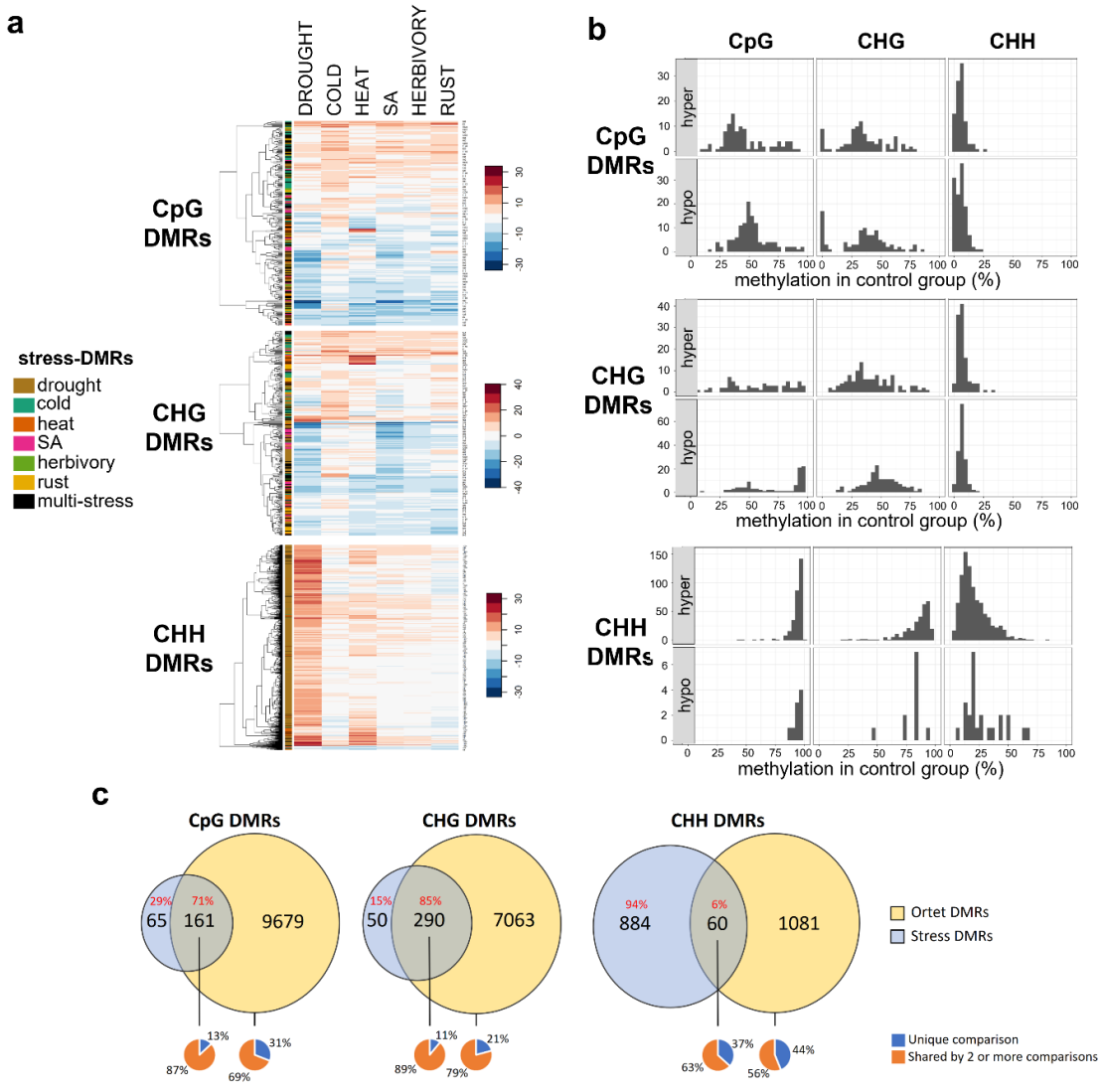


Figure 6. Detailed methylation patterns of DMRs identified in the Lombardy poplar. **a)** Heatmap and hierarchical clustering of the average difference methylation levels (compared to control) of the 1,728 identified stress-DMRs **b)** Histograms of CpG, CHG and CHH methylation level in the control group for all stress-DMRs. Histograms are shown according to DMR features (context and response: hyper/hypo). **c)** Venn diagrams of the intersections between ortet-DMRs and stress-DMRs for each sequence context. Uniqueness of ortet-DMRs (and for the intersection) is shown below the Venn diagrams.

Functional analysis of genes associated to drought-DMRs and drought-responsive TEs

Enrichment analyses revealed very few gene ontology (GO) terms that were significant after multiple testing correction (q-value). GO terms with uncorrected p-values (<0.05) suggested that genes associated with drought CHH-DMRs were enriched in processes related to response to abiotic stimulus (GO:0071214), osmotic stress (GO:0006970), and water deprivation (GO:0009414) (Supplementary file 2). Comparisons of functional enrichments of gene sets associated to drought CHH-DMRs and drought-responsive TEs (DR-TEs) highlighted considerable overlaps. Response to abscisic acid (GO:0009737) and protein kinase activity (GO:0004672) were terms that were enriched in almost all gene sets, while cellular response to water deprivation (GO:0042631) and cellular response to water stimulus (GO:0071462) were enriched only in drought-DMR, MITE/DTH and DR-TE sets. In addition, SINE-associated genes were mainly enriched in terms related to protein phosphorylation while MITE/DTH-associated genes were mostly enriched in terms related to ABA/hormone signalling and response to water stimulus. Gene sets associated with HDR-TEs showed similar enrichments than those accounting for DR-TEs (Table 1). Thus, regions and TE superfamilies that showed methylation responses to drought seem to be located close to drought-responsive genes.

Table 1. Gene Ontology (GO) enrichment analysis of genes associated with drought CHH-DMRs and Drought-Responsive TEs (DR-TEs) and Highly Drought-Responsive TEs (HDR-TEs). Only significant GO terms (p-value < 0.05) for each gene set are marked by “X”. Enriched GO terms for drought CHH-DMRs (far right) were used as a basis for comparison among all gene set enrichments.

Table 1

			DR-TEs			HDR-TEs (5% hypermethylation)			
GO ID	**GO Description**	**ontology**	**SINEs**	**MITEHs**	**both**	**HDR SINEs**	**HDR MITEHs**	**both**	**DROUGHT CHH-DMRs**
GO:0009719	response to endogenous stimulus	BP		X	X				X
GO:0009725	response to hormone	BP		X	X				X
GO:0016310	phosphorylation	BP	X		X	X		X	X
GO:0006468	protein phosphorylation	BP	X		X	X		X	X
GO:0001101	response to acid chemical	BP			X				X
GO:0097305	response to alcohol	BP	X	X	X	X	X	X	X
GO:0009737	response to abscisic acid	BP		X	X	X	X	X	X
GO:0071229	cellular response to acid chemical	BP		X	X				X
GO:0042631	cellular response to water deprivation	BP		X	X				X
GO:0071462	cellular response to water stimulus	BP		X	X				X
GO:0019853	L-ascorbic acid biosynthetic process	BP	X		X				X
GO:0009963	positive regulation of flavonoid biosynthetic process	BP					X		X
GO:0019632	shikimate metabolic process	BP						X	X
GO:0016772	transferase activity, transferring phosphorus-containing groups	MF	X			X		X	X
GO:0016301	kinase activity	MF	X			X		X	X
GO:0016773	phosphotransferase activity, alcohol group as acceptor	MF	X		X	X		X	X
GO:0004672	protein kinase activity	MF	X		X	X		X	X
GO:0004674	protein serine/threonine kinase activity	MF	X		X	X		X	X
GO:0106310	protein serine kinase activity	MF	X		X	X		X	X
GO:0106311	protein threonine kinase activity	MF	X		X	X		X	X
GO:0050660	flavin adenine dinucleotide binding	MF				X			X
GO:0000287	magnesium ion binding	MF					X		X
GO:0015144	carbohydrate transmembrane transporter activity	MF					X		X
GO:0090599	alpha-glucosidase activity	MF		X	X				X
GO:0003855	3-dehydroquinate dehydratase activity	MF			X			X	X
GO:0004764	shikimate 3-dehydrogenase (NADP+) activity	MF			X			X	X
GO:0008143	poly(A) binding	MF						X	X
GO:0070717	poly-purine tract binding	MF						X	X
GO:0030136	clathrin-coated vesicle	CC			X		X	X	X
GO:0031312	extrinsic component of organelle membrane	CC					X	X	X
GO:0031314	extrinsic component of mitochondrial inner membrane	CC					X	X	X
GO:0012510	trans-Golgi network transport vesicle membrane	CC					X		X

Discussion

In this study, we characterized the DNA methylation response to a panel of six different environmental treatments in the clonal tree *Populus nigra cv.* 'Italica'. To the best of our knowledge, this is the first WGBS study that evaluates the effects of many different environmental factors on a tree species, and with a high number of replicates (n=8), thereby enabling a robust and comprehensive comparative analysis of the DNA methylation stress response.

Global signatures of the poplar methylome response to individual stress treatments

Abiotic stresses

The global patterns of DNA hypermethylation after exposure to abiotic stresses substantiates previous studies on *Populus* that have reported global DNA methylation increases after 5-weeks of drought stress in *P. trichocarpa* (Liang *et al.,* 2014) and after 7 days of salt stress in *P. euphratica* (Su *et al.,* 2018). In *P. simonii*, methylation gradually increased during the first 24 h of either cold, heat, salinity or osmotic stress, and certain enzymes involved in (de)methylation pathways were up/downregulated in a stress-specific manner (Song *et al.,* 2016). However, the low-resolution methods (HPLC and MSAP) used to quantify DNA methylation did not allow the authors of that study to further investigate methylation at context-specific level. Here, using WGBS data, we were able to determine that genome-wide stress-induced hypermethylation can arise in a sequence context specific manner as a response to specific stresses. Together with the mentioned studies, our findings suggest that the context-specific hypermethylation patterns induced by abiotic stresses are the product of *de novo* methylation via the RdDM pathway combined with stress-specific up/downregulation of demethylation pathways.

Abscisic acid (ABA) is known to initiate stress signalling leading to physiological acclimation upon stress (Jia *et al.,* 2016, 2017; Popko *et al.,* 2010). However, only few studies have hinted its potential role in mediating global hypermethylation responses (Lafon-Placette *et al.,* 2018; Song *et al.,* 2016; Su *et al.,* 2018). For instance, ABA treatments in Arabidopsis induced hypermethylation at ABA-responsive genes (Gohlke *et al.,* 2013). Moreover, ABA-mediated upregulation of specific microRNAs can downregulate targeted demethylases (Sunkar & Zhu, 2004), which in turn may result in hypermethylation. Because increase of methylation may be associated with gene silencing (Fojtova *et al.,* 2003; Paszkowski & Whitham, 2001), stress-induced global hypermethylation may induce progressive gene silencing leading to arrested growth

under adverse conditions. Conversely, In *P. tremula* and tree peony, bud growth reactivation is preceded by a progressive reduction of genomic DNA methylation (Conde *et al.,* 2017; Y. Zhang *et al.,* 2020). Therefore, it is plausible that cold-induced hypermethylation occurs as a first response during winter, arresting growth, followed by a gradual demethylation that leads to growth reactivation in spring. Since changes in CHH methylation are less stable than those in CpG/CHG contexts (Secco *et al.,* 2015; Wibowo *et al.,* 2016), context-specific hypermethylation suggests different stabilities and thus durations of the response. This might reflect differences in duration of the environmental stresses in nature, specifically longer cold periods (winter) versus brief episodes of drought during the growing season.

Biotic stresses

The effect of biotic stresses on DNA methylation has been examined in *Arabidopsis* and other species (Dowen *et al.,* 2012; Ramos-Cruz *et al.,* 2021; H. Zhang *et al.,* 2018), however little information is available about woody plants. It is known that SA treatment, rust infection, and caterpillar attack increase the levels of SA, JA and ABA in the affected poplar leaves (Clavijo Mccormick *et al.,* 2014; Eberl, Hammerbacher, *et al.,* 2018; Eberl, Perreca, *et al.,* 2018; Li *et al.,* 2018; Ullah *et al.,* 2019). Moreover, SA can be transported from infected to uninfected sites to induce systemic acquired resistance (SAR) (Li *et al.,* 2018). Therefore, we will discuss the biotic-induced methylation patterns in the context of SAR as we sampled non-directly affected leaves that grew during the stress periods.

Consistent with our results, treatment with exogenous SA has been reported to induce DNA hypomethylation in other species, which in turn activates defense-response genes (Dowen *et al.,* 2012; Ngom *et al.,* 2017). Specifically, in *Vitis amurensis*, exogenous SA upregulated specific demethylases (Kiselev *et al.,* 2013), which induced hypomethylation and consequently enhanced production of secondary metabolites (Kiselev *et al.,* 2015). The loss of DNA methylation can either prime (upon removal of CHH methylation) or constitutively derepress (upon removal of CpG/CHG/CHH methylation) the SA-dependent defense response (Deleris *et al.,* 2016; López Sánchez *et al.,* 2016). Since we found similar CpG/CHG hypomethylation patterns upon both herbivory and SA treatment, we speculate that such response mainly de-repress the SA-dependent defense response, while the TE-associated CHH hypomethylation response after rust infection could have a priming effect.

Diminished rust infection has been observed in drought-affected poplars, explained to some extent by increased stomatal closure mediated by ABA (Ullah *et al.,* 2019), but disregarding the interplay between methylation responses. Here, we showed that rust

infection also induced CpG/CHG hypermethylation profiles very similar to those observed under drought and heat, but not cold. These similarities suggest a possible overlap between the responses to drought and rust infection, as it has been suggested by results on the poplar apoplast proteome (Pechanova *et al.,* 2010).

Hotspots of environmental-induced methylation variation

Experiments for studying stress effects on DNA methylation usually analyse DMRs induced by single stresses. Thus, intersection of several DMR sets detected from different stresses allows the capture of more generic responses. Though the latter approach can pinpoint multi-stress DMRs (Song *et al.,* 2016; Xue *et al.,* 2013), it likely underestimates commonalities in the methylation response to different stresses because stringent significance thresholds in DMR detection can leave most of the responding loci undetected. Here, we found that CpG/CHG-DMRs often showed a similar response irrespective of treatment, suggesting that much of the stress response in poplar is generic, rather than stress-specific. Based on these observations, we suspect that many of the reported stress-specific DMRs in other species likely have also a multi-stress nature, which would imply a more careful interpretation of DMR results in the future.

Our results resemble the observations of epimutational hotspots in nearly isogenic *Arabidopsis* lines under greenhouse and natural environmental conditions (Becker *et al.,* 2011; Hagmann *et al.,* 2015; Schmitz *et al.,* 2011). Such epimutation hotspots are characterized by steady-state intermediate methylation levels (Hazarika *et al.,* 2022), which was also observed in the control methylation levels of CG/CHG DMRs. However, where the intermediate methylation level of *A. thaliana* hotspots is due to sparse cytosine methylation (only a subset of CpGs is methylated), in our case it is the result of individual cytosines being partially methylated. Since such CpG/CHG-DMRs are often located on TE flanking regions, we hypothesize that TE-mediated stress-induced (de)methylation is the source of methylation variation on the TE edges, here identified as multi-stress DMRs.

Stress-induced methylation variation as a source of epialleles under natural conditions

In this clonal system, the accumulation of methylation variation can be attributed mostly to spontaneous and environmentally induced variation. In other species, transient stress-responsive epigenetically labile regions have been identified to also overlap with naturally occurring DMRs, suggesting a non-random stress-triggered epigenetic reprogramming (Miryeganeh *et al.,* 2022; Wibowo *et al.,* 2016). Here, we identified many CpG/CHG-DMRs among ortets, which are thought to be mitotically stable and hence

clonally transmissible. More interestingly, a large proportion of stress-induced DMRs overlapped with ortet-DMRs. This result indicates that at least part of the natural methylation variation of the clonal system at a European scale is induced by changing environments. Consequently, environment-induced methylation variants in CpG/CHG contexts could be fixed and appear as natural epialleles detectable across the tree lifespan and maybe next clonal generations.

In contrast, induced CHH-DMRs showed only a minor overlap with ortet-DMRs, even though such DMRs largely arose in response to drought and heat. This observation supports the idea that CHH methylation variation quickly disappears after the stress is gone, preventing induced CHH-DMRs to persist as natural epialleles. This capability of CpG/CHG methylation to track long-term environmental variation seems to be supported by recent observations in other trees (Heer *et al.,* 2018; Miryeganeh *et al.,* 2022).

Functionality of the poplar methylome response to drought

As poplar is a fast-growing riparian tree whose high productivity requires high water availability (Monclus *et al.,* 2006; Vanden Broeck, 2003), methylation responses to drought are potentially relevant for the ecology of this species. Recent reports in the species have found significant genotypic variation involved in drought tolerance (Viger *et al.,* 2016) as well as for drought escape (Yıldırım & Kaya, 2017). Therefore, efficient finetuning of the drought escape and tolerance responses is likely a strong selection pressure in this clonal cultivar, which may have promoted the evolution of DNA-methylation-based regulatory mechanisms.

Even though the study of the functionality of DNA methylation would require at the very least quantification of gene expression, some patterns that we observed in the methylome response to drought suggest a functional consequence. Here, we reported TE-associated CHH hypermethylation mostly in gene flanking regions, which has been also described in *P. trichocarpa* (Liang *et al.,* 2014). However, our analysis also revealed hypermethylation enriched on specific TE superfamilies: SINE and MITE/DTH.

TE activity can be triggered by biotic and abiotic stress conditions (Lanciano & Mirouze, 2018; Seibt *et al.,* 2016), which in turn may lead to a rapid and extensive TE amplification followed by inactivity and drift (Jiang *et al.,* 2004). Such is the case of SINEs and MITEs irrespective of their inherent differences: retrotransposons vs. DNA transposons, respectively. Both superfamilies are relatively short elements frequently inserted close to and within genes (Kögler *et al.,* 2020; Seibt *et al.,* 2016), likely due to their tendency to integrate in hypomethylated DNA regions (Arnaud *et al.,* 2000). Also, both are

preferential targets for de novo methylation, which can then spread into flanking sequences and may affect the expression of nearby genes (Arnaud *et al.,* 2000; Chen *et al.,* 2014). TE proximity to genes may suggest that stress-induced TE hypermethylation could be a by-product of highly expressed nearby genes, as previously suggested by (Secco *et al.,* 2015). However, we observed that CHH hypermethylation occurred irrespective of their distance to genes, indicating that such response may not be a consequence of nearby gene expression.

Hypermethylation of entire TE superfamilies in response to stress has not been previously reported in other species. We found that SINE and MITE/DTH elements were already highly methylated (and presumably silenced) under control conditions. Therefore, as drought seems to reinforce such hypermethylation, we speculate that the selective silencing of these elements can have regulation consequences of nearby drought-responsive genes as hinted by GO enrichment results.

Genes associated with drought CHH-DMRs seemed enriched in general responses to drought, such as ABA signalling, protein kinase activity, and response to water deprivation. Remarkably, genes associated with SINE and MITE/DTH elements were also enriched in similar functional responses. Evidence of MITE-derived small RNAs regulating abscisic acid signalling and abiotic stress responses in rice (Yan *et al.,* 2011) suggests that these elements may have been selected and retained close to specific genes that play a role in the rapid response to drought. Therefore, we speculate that the triggered hypermethylation response is regulating, via specific TE superfamilies, genes and pathways involved in functional responses to drought. To test this functional hypothesis, it will be important in follow-up studies to monitor expression and methylation of SINEs, MITE/DTHs, and the nearby genes prior to and during drought stress.

In conclusion, our evaluation of the poplar methylome plasticity upon abiotic and biotic treatments allowed the discovery of multi-stress hotspots that are partially shaping the natural methylation variation. Moreover, we were able to identify specific TE superfamilies whose response to drought may have been selected to cope with extreme conditions. Our study furthermore highlights the importance of analyzing the effect of multiple factors in the same experiment to avoid overstatements of individual effects.

Data availability

The bisulfite sequencing raw data is deposited in the ENA Sequence Read Archive Repository (www.ebi.ac.uk/ena/) under study accession number: PRJEB51831. Methylation files for the three contexts and the list of annotated differentially methylated regions are deposited in zenodo (https://zenodo.org/) under DOI: 10.5281/zenodo.7193978. Gene models and TE predictions for the poplar clonal cultivar used in this study are deposited in zenodo under DOI: (to be uploaded)

Supporting information

https://www.biorxiv.org/content/10.1101/2022.10.18.512698v1.supplementary-material

Acknowledgements

We thank Slavica Milanovic-Ivanovic and Gregor Disveld for their technical help in the molecular lab and greenhouse settings, respectively. We thank Morgane van Antro and Haymanti Saha for the fruitful discussions and sampling assistance. We thank Bhumika Dubay for her work on the reference genome, gene model and TE predictions. We are grateful for all the input and discussions with all the members of the Epidiverse Consortium. We also express our gratitude to Sybille Unsicker for providing the poplar rust spores and gypsy moth caterpillars.

This work was supported by the European Training Network “EpiDiverse” and received funding from the EU Horizon 2020 program under Marie SkłodowskaCurie grant agreement No 764965.

Conflict of interest

The authors have no conflict of interest to declare.

References

Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F. E., Figueroa, M. E., Melnick, A., & Mason, C. E. 2012. methylKit: A comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biology*, *13*(10), R87.

Arnaud, P., Goubely, C., Pélissier, T., & Deragon, J. M. 2000. SINE retroposons can be used in vivo as nucleation centers for de novo methylation. *Molecular and Cellular Biology*, *20*(10), 3434–3441.

Becker, C., Hagmann, J., Müller, J., Koenig, D., Stegle, O., Borgwardt, K., & Weigel, D. 2011. Spontaneous epigenetic variation in the Arabidopsis thaliana methylome. *Nature*, *480*(7376), Article 7376.

Bewick, A. J., & Schmitz, R. J. 2017. Gene body DNA methylation in plants. *Current Opinion in Plant Biology*, *36*, 103–110.

Bewick, A. J., Zhang, Y., Wendte, J. M., Zhang, X., & Schmitz, R. J. 2019. Evolutionary and Experimental Loss of Gene Body Methylation and Its Consequence to Gene Expression. *G3 Genes|Genomes|Genetics*, *9*(8), 2441–2445.

Boyko, A., Blevins, T., Yao, Y., Golubov, A., Bilichak, A., Ilnytskyy, Y., Hollander, J., Meins, F., & Kovalchuk, I. 2010. Transgenerational Adaptation of Arabidopsis to Stress Requires DNA Methylation and the Function of Dicer-Like Proteins. *PLoS ONE*, *5*(3), e9514.

Chen, J., Hu, Q., Lu, C., & Kuang, H. 2014. Evolutionary Genomics of Miniature Inverted-Repeat Transposable Elements (MITEs) in Plants. In P. Pontarotti (Ed.), *Evolutionary Biology: Genome Evolution, Speciation, Coevolution and Origin of Life* (pp. 157–168). Springer International Publishing.

Chenault, N., Arnaud-Haond, S., Juteau, M., Valade, R., Almeida, J.-L., Villar, M., Bastien, C., & Dowkiw, A. 2011. SSR-based analysis of clonality, spatial genetic structure and introgression from the Lombardy poplar into a natural population of *Populus nigra* L. along the Loire River. *Tree Genetics & Genomes*, *7*(6), 1249–1262

Clavijo Mccormick, A., Irmisch, S., Reinecke, A., Boeckler, G. A., Veit, D., Reichelt, M., Hansson, B. S., Gershenzon, J., Köllner, T. G., & Unsicker, S. B. 2014. Herbivore-induced volatile emission in black poplar: Regulation and role in attracting herbivore enemies. *Plant, Cell & Environment*, *37*(8), 1909–1923.

Conde, D., Le Gac, A.-L., Perales, M., Dervinis, C., Kirst, M., Maury, S., González-Melendi, P., & Allona, I. 2017. Chilling-responsive DEMETER-LIKE DNA demethylase mediates in poplar bud break. *Plant, Cell & Environment*, *40*(10), 2236–2249.

Dancho, M. 2022. *R package tidyquant: Tidy Quantitative Financial Analysis* (R package version 1.0.5). https://github.com/business-science/tidyquant

Deleris, A., Halter, T., & Navarro, L. 2016. DNA Methylation and Demethylation in Plant Immunity. *Annual Review of Phytopathology*, *54*, 579–603.

Díez-Rodríguez, B., Peña-Ponton, C., Pérez-Bello, P., Bette, J., Lerbs, L., Mackenbach, T., Wulle, S., Paoli, E. D., Verhoeven, K. J. F., Heer, K., & Opgenoorth, L. 2022. *An uncommon garden experiment: Microenvironment has stronger influence on phenotypic variation than epigenetic memory in the clonal Lombardy poplar* (p. 2022.03.22.485169). bioRxiv.

Dowen, R. H., Pelizzola, M., Schmitz, R. J., Lister, R., Dowen, J. M., Nery, J. R., Dixon, J. E., & Ecker, J. R. 2012. Widespread dynamic DNA methylation in response to biotic stress. *Proceedings of the National Academy of Sciences*, *109*(32), E2183–E2191.

Drost, H.-G., Gabel, A., Grosse, I., & Quint, M. 2015. Evidence for Active Maintenance of Phylotranscriptomic Hourglass Patterns in Animal and Plant Embryogenesis. *Molecular Biology and Evolution*, *32*(5), 1221–1231.

Eberl, F., Hammerbacher, A., Gershenzon, J., & Unsicker, S. B. 2018. Leaf rust infection reduces herbivore-induced volatile emission in black poplar and attracts a generalist herbivore. *New Phytologist*, *220*(3), 760–772.

Eberl, F., Perreca, E., Vogel, H., Wright, L. P., Hammerbacher, A., Veit, D., Gershenzon, J., & Unsicker, S. B. 2018. Rust Infection of Black Poplar Trees Reduces Photosynthesis but Does Not Affect Isoprene Biosynthesis or Emission. *Frontiers in Plant Science*, *9*.

Fojtova, M., Van Houdt, H., Depicker, A., & Kovarik, A. 2003. Epigenetic Switch from Posttranscriptional to Transcriptional Silencing Is Correlated with Promoter Hypermethylation. *Plant Physiology*, *133*(3), 1240–1250.

Gamer, M., Lemon, J., Fellows, I., & Singh, P. 2019. *R package irr: Various Coefficients of Interrater Reliability and Agreement* (R package version 0.84.1). https://CRAN.R-project.org/package=irr

Gohlke, J., Scholz, C.-J., Kneitz, S., Weber, D., Fuchs, J., Hedrich, R., & Deeken, R. 2013. DNA Methylation Mediated Control of Gene Expression Is Critical for Development of Crown Gall Tumors. *PLOS Genetics*, *9*(2), e1003267.

Hagmann, J., Becker, C., Müller, J., Stegle, O., Meyer, R. C., Wang, G., Schneeberger, K., Fitz, J., Altmann, T., Bergelson, J., Borgwardt, K., & Weigel, D. (2015). Century-scale Methylome Stability in a Recently Diverged Arabidopsis thaliana Lineage. *PLOS Genetics*,

Hazarika, R. R., Serra, M., Zhang, Z., Zhang, Y., Schmitz, R. J., & Johannes, F. 2022. Molecular properties of epimutation hotspots. *Nature Plants*, *8*(2), Article 2.

Heer, K., Ullrich, K. K., Hiss, M., Liepelt, S., Schulze Brüning, R., Zhou, J., Opgenoorth, L., & Rensing, S. A. 2018. Detection of somatic epigenetic variation in Norway spruce via targeted bisulfite sequencing. *Ecology and Evolution*, *8*(19), 9672–9682.

Hothorn, T., Bretz, F., & Westfall, P. 2008. Simultaneous inference in general parametric models. *Biometrical Journal. Biometrische Zeitschrift*, *50*(3), 346–363.

Jansson, S., & Douglas, C. J. 2007. Populus: A Model System for Plant Biology. *Annual Review of Plant Biology*, *58*(1), 435–458.

Jia, J., Li, S., Cao, X., Li, H., Shi, W., Polle, A., Liu, T.-X., Peng, C., & Luo, Z.-B. 2016. Physiological and transcriptional regulation in poplar roots and leaves during acclimation to high temperature and drought. *Physiologia Plantarum*, *157*(1), 38–53.

Jia, J., Zhou, J., Shi, W., Cao, X., Luo, J., Polle, A., & Luo, Z.-B. 2017. Comparative transcriptomic analysis reveals the roles of overlapping heat-/drought-responsive genes in poplars exposed to high temperature and drought. *Scientific Reports*, *7*(1), Article 1.

Jiang, N., Feschotte, C., Zhang, X., & Wessler, S. R. 2004. Using rice to understand the origin and amplification of miniature inverted repeat transposable elements (MITEs). *Current Opinion in Plant Biology*, *7*(2), 115–119.

Karpiński, S., Szechyńska-Hebda, M., Wituszyńska, W., & Burdiak, P. 2013. Light acclimation, retrograde signalling, cell death and immune defences in plants. *Plant, Cell & Environment*, *36*(4), 736–744.

Kenchanmane Raju, S. K., Shao, M.-R., Wamboldt, Y., & Mackenzie, S. 2018. Epigenomic plasticity of Arabidopsis msh1 mutants under prolonged cold stress. *Plant Direct*, *2*(8), e00079.

Kiselev, K. V., Tyunin, A. P., & Karetin, Y. A. 2013. Influence of 5-azacytidine and salicylic acid on demethylase gene expression in cell cultures of *Vitis amurensis* Rupr. *Acta Physiologiae Plantarum*, *35*(6), 1843–1851.

Kiselev, K. V., Tyunin, A. P., & Karetin, Y. A. 2015. Salicylic acid induces alterations in the methylation pattern of the VaSTS1, VaSTS2, and VaSTS10 genes in *Vitis amurensis* Rupr. Cell cultures. *Plant Cell Reports*, *34*(2), 311–320.

Kögler, A., Seibt, K. M., Heitkam, T., Morgenstern, K., Reiche, B., Brückner, M., Wolf, H., Krabel, D., & Schmidt, T. 2020. Divergence of 3′ ends as a driver of short interspersed nuclear element (SINE) evolution in the Salicaceae. *The Plant Journal*, *103*(1), 443–458.

Koo, T. K., & Li, M. Y. 2016. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*, 1–26.

Lafon-Placette, C., Le Gac, A.-L., Chauveau, D., Segura, V., Delaunay, A., Lesage-Descauses, M.-C., Hummel, I., Cohen, D., Jesson, B., Le Thiec, D., Bogeat-Triboulot, M.-B., Brignolas, F., & Maury, S. 2018. Changes in the epigenome and transcriptome of the poplar shoot apical meristem in response to water availability affect preferentially hormone pathways. *Journal of Experimental Botany*, *69*(3), 537–551.

Lämke, J., & Bäurle, I. (2017). Epigenetic and chromatin-based mechanisms in environmental stress adaptation and stress memory in plants. *Genome Biology*, *18*, 124.

Lanciano, S., & Mirouze, M. 2018. Transposable elements: All mobile, all different, some stress responsive, some adaptive? *Current Opinion in Genetics & Development*, *49*, 106–114.

Larchevêque, M., Maurel, M., Desrochers, A., & Larocque, G. R. 2011. How does drought tolerance compare between two improved hybrids of balsam poplar and an unimproved native species? *Tree Physiology*, *31*(3), 240–249.

Li, Y., Zhang, W., Dong, H., Liu, Z., Ma, J., & Zhang, X. 2018. Salicylic acid in Populus tomentosa is a remote signalling molecule induced by Botryosphaeria dothidea infection. *Scientific Reports*, *8*(1), Article 1.

Liang, D., Zhang, Z., Wu, H., Huang, C., Shuai, P., Ye, C.-Y., Tang, S., Wang, Y., Yang, L., Wang, J., Yin, W., & Xia, X. 2014. Single-base-resolution methylomes of populus trichocarpa reveal the association between DNA methylation and drought stress. *BMC Genetics*, *15*(1), S9.

Liu, J., & He, Z. 2020. Small DNA Methylation, Big Player in Plant Abiotic Stress Responses and Memory. *Frontiers in Plant Science*, *11*.

López Sánchez, A., Stassen, J. H. M., Furci, L., Smith, L. M., & Ton, J. 2016. The role of DNA (de)methylation in immune responsiveness of Arabidopsis. *The Plant Journal*, *88*(3), 361–374.

Matzke, M. A., & Mosher, R. A. 2014. RNA-directed DNA methylation: An epigenetic pathway of increasing complexity. *Nature Reviews. Genetics*, *15*(6), 394–408.

McGraw, K. O., & Wong, S. P. 1996. Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*(1), 30–46.

Miryeganeh, M., Marlétaz, F., Gavriouchkina, D., & Saze, H. 2022. De novo genome assembly and in natura epigenomics reveal salinity-induced DNA methylation in the mangrove tree Bruguiera gymnorhiza. *New Phytologist*, *233*(5), 2094–2110.

Monclus, R., Dreyer, E., Villar, M., Delmotte, F. M., Delay, D., Petit, J.-M., Barbaroux, C., Le Thiec, D., Bréchet, C., & Brignolas, F. 2006. Impact of drought on productivity and water use efficiency in 29 genotypes of Populus deltoides x Populus nigra. *The New Phytologist*, *169*(4), 765–777.

Ngom, B., Sarr, I., Kimatu, J., Mamati, E., & Kane, N. A. 2017. Genome-wide analysis of cytosine DNA methylation revealed salicylic acid promotes defense pathways over seedling development in pearl millet. *Plant Signaling & Behavior*, *12*(9), e1356967.

Nunn, A., Can, S. N., Otto, C., Fasold, M., Díez Rodríguez, B., Fernández-Pozo, N., Rensing, S. A., Stadler, P. F., & Langenberger, D. 2021. EpiDiverse Toolkit: A pipeline suite for the analysis of bisulfite sequencing data in ecological plant epigenetics. *NAR Genomics and Bioinformatics*, *3*(4), lqab106.

Nunn, A., Rodríguez-Arévalo, I., Tandukar, Z., Frels, K., Contreras-Garrido, A., Carbonell-Bejerano, P., Zhang, P., Ramos Cruz, D., Jandrasits, K., Lanz, C., Brusa, A., Mirouze, M., Dorn, K., Galbraith, D. W., Jarvis, B. A., Sedbrook, J. C., Wyse, D. L., Otto, C., Langenberger, D., … Chopra, R. 2022. Chromosome-level Thlaspi arvense genome provides new tools for translational research and for a newly domesticated cash cover crop of the cooler climates. *Plant Biotechnology Journal*, *20*(5), 944–963.

Paszkowski, J., & Whitham, S. A. 2001. Gene silencing and DNA methylation processes. *Current Opinion in Plant Biology*, *4*(2), 123–129.

Pechanova, O., Hsu, C.-Y., Adams, J. P., Pechan, T., Vandervelde, L., Drnevich, J., Jawdy, S., Adeli, A., Suttle, J. C., Lawrence, A. M., Tschaplinski, T. J., Séguin, A., & Yuceer, C. (2010). Apoplast proteome reveals that extracellular matrix contributes to multistress response in poplar. *BMC Genomics*, *11*, 674.

Popko, J., Hänsch, R., Mendel, R.-R., Polle, A., & Teichmann, T. 2010. The role of abscisic acid and auxin in the response of poplar to abiotic stress. *Plant Biology*, *12*(2), 242–258.

R Core Team. 2022. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/

Ramos-Cruz, D., Troyee, A. N., & Becker, C. 2021. Epigenetics in plant organismic interactions. *Current Opinion in Plant Biology*, *61*, 102060.

Rood, S. B., Braatne, J. H., & Hughes, F. M. R. 2003. Ecophysiology of riparian cottonwoods: Stream flow dependency, water relations and restoration†. *Tree Physiology*, *23*(16), 1113–1124.

Schmitz, R. J., Schultz, M. D., Lewsey, M. G., O'Malley, R. C., Urich, M. A., Libiger, O., Schork, N. J., & Ecker, J. R. 2011. Transgenerational Epigenetic Instability Is a Source of Novel Methylation Variants. *Science*, *334*(6054), 369–373.

Secco, D., Wang, C., Shou, H., Schultz, M. D., Chiarenza, S., Nussaume, L., Ecker, J. R., Whelan, J., & Lister, R. (2015). Stress induced gene expression drives transient DNA methylation changes at adjacent repetitive elements. *ELife*, *4*, e09343.

Seibt, K. M., Wenke, T., Muders, K., Truberg, B., & Schmidt, T. 2016. Short interspersed nuclear elements (SINEs) are abundant in Solanaceae and have a family-specific impact on gene structure and genome organization. *The Plant Journal*, *86*(3), 268–285.

Seymour, D. K., & Gaut, B. S. 2020. Phylogenetic Shifts in Gene Body Methylation Correlate with Gene Expression and Reflect Trait Conservation. *Molecular Biology and Evolution*, *37*(1), 31–43.

Song, Y., Ci, D., Tian, M., & Zhang, D. 2016. Stable methylation of a non-coding RNA gene regulates gene expression in response to abiotic stress in Populus simonii. *Journal of Experimental Botany*, *67*(5), 1477–1492.

Sow, M. D., Le Gac, A.-L., Fichot, R., Lanciano, S., Delaunay, A., Le Jan, I., Lesage-Descauses, M.-C., Citerne, S., Caius, J., Brunaud, V., Soubigou-Taconnat, L., Cochard, H., Segura, V., Chaparro, C., Grunau, C., Daviaud, C., Tost, J., Brignolas, F., Strauss, S. H., … Maury, S. 2021. RNAi suppression of DNA methylation affects the drought stress response and genome integrity in transgenic poplar. *New Phytologist*, *232*(1), 80–97.

Stettler, R. F., Reinhard F. 2009. *Cottonwood and the river of time: On trees, evolution, and society*. Seattle: University of Washington Press.

Su, Y., Bai, X., Yang, W., Wang, W., Chen, Z., Ma, J., & Ma, T. 2018. Single-base-resolution methylomes of Populus euphratica reveal the association between DNA methylation and salt stress. *Tree Genetics & Genomes*, *14*(6), 86.

Sunkar, R., & Zhu, J.-K. 2004. Novel and Stress-Regulated MicroRNAs and Other Small RNAs from Arabidopsis[W]. *The Plant Cell*, *16*(8), 2001–2019.

Tuskan, G. A., DiFazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., Aerts, A., Bhalerao, R. R., Bhalerao, R. P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., … Rokhsar, D. 2006. The Genome of Black Cottonwood, Populus trichocarpa (Torr. & Gray). *Science*, *313*(5793), 1596–1604.

Ullah, C., Tsai, C.-J., Unsicker, S. B., Xue, L., Reichelt, M., Gershenzon, J., & Hammerbacher, A. 2019. Salicylic acid activates poplar defense against the biotrophic rust fungus Melampsora larici-populina via increased biosynthesis of catechin and proanthocyanidins. *New Phytologist*, *221*(2), 960–975.

Vanden Broeck. (2003). *EUFORGEN Technical Guidelines for genetic conservationand use for black poplar (Populus nigra)* (pp. 1–6). International Plant Genetic Resources Institute. http://www.euforgen.org/publications/publication/ipopulus-nigrai-technical-guidelines-for-genetic-conservation-and-use-for-black-poplar/

Viger, M., Smith, H. K., Cohen, D., Dewoody, J., Trewin, H., Steenackers, M., Bastien, C., & Taylor, G. 2016. Adaptive mechanisms and genomic plasticity for drought tolerance identified in European black poplar (Populus nigra L.). *Tree Physiology*, *36*(7), 909–928.

Wang, C., El-Shetehy, M., Shine, M. B., Yu, K., Navarre, D., Wendehenne, D., Kachroo, A., & Kachroo, P. 2014. Free radicals mediate systemic acquired resistance. *Cell Reports*, *7*(2), 348–355.

Wibowo, A., Becker, C., Marconi, G., Durr, J., Price, J., Hagmann, J., Papareddy, R., Putra, H., Kageyama, J., Becker, J., Weigel, D., & Gutierrez-Marcos, J. 2016. Hyperosmotic stress memory in Arabidopsis is mediated by distinct epigenetically labile sites in the genome and is restricted in the male germline by DNA glycosylase activity. *ELife*, *5*, e13546.

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., Fu, X., Liu, S., Bo, X., & Yu, G. 2021. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation*, *2*(3).

Xue, L.-J., Guo, W., Yuan, Y., Anino, E. O., Nyamdari, B., Wilson, M. C., Frost, C. J., Chen, H.-Y., Babst, B. A., Harding, S. A., & Tsai, C.-J. 2013. Constitutively Elevated Salicylic Acid Levels Alter Photosynthesis and Oxidative State but Not Growth in Transgenic Populus[C][W]. *The Plant Cell*, *25*(7), 2714–2730.

Yan, Y., Zhang, Y., Yang, K., Sun, Z., Fu, Y., Chen, X., & Fang, R. 2011. Small RNAs from MITE-derived stem-loop precursors regulate abscisic acid signaling and abiotic stress responses in rice. *The Plant Journal*, *65*(5), 820–828.

Yıldırım, K., & Kaya, Z. 2017. Gene regulation network behind drought escape, avoidance and tolerance strategies in black poplar (Populus nigra L.). *Plant Physiology and Biochemistry: PPB*, *115*, 183–199.

Zhang, H., Lang, Z., & Zhu, J.-K. 2018. Dynamics and function of DNA methylation in plants. *Nature Reviews Molecular Cell Biology*, *19*(8), Article 8.

Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S. W.-L., Chen, H., Henderson, I. R., Shinn, P., Pellegrini, M., Jacobsen, S. E., & Ecker, J. R. 2006. Genome-wide High-Resolution Mapping and Functional Analysis of DNA Methylation in Arabidopsis. *Cell*, *126*(6), 1189–1201.

Zhang, Y., Si, F., Wang, Y., Liu, C., Zhang, T., Yuan, Y., & Gai, S. 2020. Application of 5-azacytidine induces DNA hypomethylation and accelerates dormancy release in buds of tree peony. *Plant Physiology and Biochemistry: PPB*, *147*, 91–100.

Zilberman, D., Gehring, M., Tran, R. K., Ballinger, T., & Henikoff, S. 2007. Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. *Nature Genetics*, *39*(1), 61–69.

CHAPTER IV

Natural patterns of DNA methylation

Bárbara Díez Rodríguez, Dario Galanti, Adam Nunn, Cristian Peña-Ponton, Paloma Pérez-Bello, Iris Sammarco, Katharina Jandrasits, Claude Becker, Emanuele de Paoli, Koen J.F Verhoeven, Lars Opgenoorth and Katrin Heer

This chapter is in preparation for submission to *New Phytologist*

Abstract

Environmental changes can trigger phenotypic variation in plants through epigenetic mechanisms, but strong genetic influences make it difficult to isolate and study epigenetic effects. Clonal trees with low genetic variation, such as the Lombardy poplar *(Populus nigra* cv. 'Italica' Duroi), offer a unique system to study epigenetic variation associated with the environment. We collected cuttings (ramets) of Lombardy poplar along a wide geographical range in Europe. We performed whole-genome-bisulfite sequencing of 164 ramets grown in a common garden and of a subset of 35 of the original parental individuals. Using historical bioclimatic data, we tested the relationship between DNA methylation and climatic gradients. We found that average methylation levels in TEs and promoter regions correlate with biologically relevant climatic variables. Furthermore, we observed that DNA methylation was transmitted to the next clonal generation, but a fraction of the methylome changed relatively fast when comparing the parental individuals with the clonal offspring. Our results suggest that the poplar methylome is a dynamic layer of information that can be transmitted to the clonal offspring and potentially affect how poplars acclimate to new environmental conditions.

Keywords: *Populus nigra* cv 'Italica', DNA methylation, whole genome bisulfite sequencing (WGBS), differentially methylated regions (DMR), acclimation, adaptation

Introduction

In the last couple of decades, extreme weather events have been increasing, often exceeding plants' and animals' tolerance thresholds, and driving mass mortalities in many species (IPCC, 2022). Understanding how plants respond to such weather events and other environmental conditions has thus become crucial for conservation policies and forest management programs. In studies on plant natural populations, intraspecific genetic diversity has been shown to contribute to the resistance and resilience of populations (Hughes et al., 2008). Genetic variation provides the baseline for phenotypic variation on which evolutionary processes can act, and plays an important role in plant adaptation (Fisher, 1958; Hughes *et al.*, 2008). However, advances in molecular biology and genomics have shown that phenotypic variation among individuals is not only determined by genetic variation (Rapp & Wendel, 2005). One additional cause of phenotypic variation is epigenetic variation (Cubas *et al*., 1999; Manning *et al*., 2006; Xie *et al.*, 2015). Several studies have shown that epigenetic variation can be spatially structured among and within plant populations, and that such a structure can be associated with environmental variation and phenotypic differentiation (Lira-Medeiros et al., 2010; Medrano et al., 2014; Avramidou et al., 2015; Kawakatsu et al., 2016; de Kort et al., 2020; Boquete et al., 2021; Galanti et al., 2022, Sammarco et al., 2022). Although causal relationships remain to be studied, such observations suggest that epigenetic variation could contribute to the acclimation of plants to changes in environmental conditions.

There are several molecular mechanisms involved in epigenetic variation, such as histone modifications, DNA methylation and small RNA-mediated processes (reviewed in Lloyd and Lister, 2022). Among these, DNA cytosine methylation (mC), is currently the most widely studied and best characterized modification (Zemach *et al.,* 2013; Matzke & Mosher, 2014; Zhang et al., 2018; Lloyd & Lister, 2022) and consists of a base alteration in which a methyl group is added to the 5th carbon of a cytosine (Moore *et al.*, 2012). In plants, cytosine methylation occurs at three different sequence contexts: CG, CHG and CHH, where H = A, T or C. Methylation at the CG and CHG contexts is usually symmetrical across both DNA strands, whereas methylation at CHH sites is asymmetrical (Meyer et al., 1994; Finnegan et al., 2003; Zhang et al., 2006; Lister et al., 2008). As a result of different mechanisms involved in DNA methylation maintenance, different sequence contexts differ in their degrees of *mitotic stability*, which are mainly dictated by their symmetry. In the symmetrical contexts, methylation maintenance is guided by the complementary DNA strand, and thus stably inherited across mitotic

divisions (Niederhuth & Schmitz 2014). On the other hand, methylation in the asymmetrical context is maintained mainly by *de novo* establishment and thus less stable across cell divisions (Peter Meyer & Lohuis, 1994). In addition, depending on the genomic feature context, DNA methylation has different roles. For example, methylation in all sequence contexts is associated with silencing of transposable elements (TEs), while CG methylation is found in promoters of transcriptionally inactive genes and in the gene body of active genes (reviewed in Niederhuth & Schmitz, 2017). Variation in DNA methylation can be under genetic control (Zhang *et al.*, 2018; Johannes & Schmitz, 2019) and arise stochastically as a result of imperfect DNA methylation maintenance (Becker *et al.*, 2011; Schmitz *et al*., 2011; Johannes & Schmitz, 2019), or be induced by environmental conditions (Raj *et al.*, 2011; Bräutigam *et al.*, 2013; Lämke & Bäurle, 2017). Furthermore, some of these methylation marks can be transmitted from parental individuals to offspring (Johannes *et al*., 2009; Becker & Weigel, 2012; Herman & Sultan, 2016; Gáspár *et al.*, 2019; Boquete *et al.*, 2021). If DNA methylation can be induced by environmental conditions, we would expect patterns of DNA methylation to be associated with geographic or climatic gradients beyond what can be explained by the underlying genetic structure of the studied population. Several studies indeed found correlations between methylation patterns and habitat or climate in different plant species. However, almost all these studies were conducted on sexually reproducing plant species, were constrained to small-scale geographic gradients, or used low-resolution molecular methods (Lira-Medeiros *et al.*, 2010; Nicotra *et al.*, 2015; Avramidou *et al.*, 2015; Gugger *et al.*, 2016; Herrera *et al.*, 2017; Gáspár *et al.*, 2019). With the continuous decrease of sequencing costs, recent studies based on whole genome bisulfite sequencing (WGBS) have provided more detailed methylation data (Dubin *et al.*, 2015; Kawakatsu *et al.*, 2016; de Kort *et al.*, 2020; Galanti *et al.*, 2022). With WGBS we can now quantify methylation at the scale of whole genomes and accurately map methylated cytosines at a single-base resolution (Lister and Ecker, 2009). Nevertheless, the extent to which genetic variation influences epigenetic variation is still not clear (Richards *et al.*, 2010, 2017). Studing epigenetic variation in asexually (i.e. clonally) reproducing species allows focusing on epigenetic variation in the absence of confounding genetic variation. Moreover, during sexual reproduction, some proportion of the methylation patterns might be reset (Wibowo et al., 2016), whereas we assume that they are faithfully transmitted during clonal propagation. Thus epigenetic marks have therefore the potential to be stably transmitted across clonal generations and may thus create heritable phenotypic variation (Verhoeven & Preite, 2014).

Since the first assembly of the *P. trichocarpa* genome in 2006, the amount of available genetic, genomic, and biochemical resources has increased considerably, and *Populus* species have become a model for studying plant adaptation (Taylor, 2002; Tuskan et al., 2006; Jansson & Douglas, 2007). The Lombardy poplar *(Populus nigra* cv. 'Italica*'* Duroi) is a widely distributed tree clone. This variety likely originated in the 18th century from one single male tree of *P. nigra,* located in central Asia (Elwes & Henry, 1913), and was spread by cuttings worldwide from Italy. It is assumed that most Lombardy poplars originate from artificial propagation performed by humans (CABI, 2022).

Here, we present the first study investigating DNA methylation variation in a clonal tree species. We collected poplar cuttings from a wide climatic and geographic gradient across Europe and planted them in a common garden in Central Germany. We analyzed methylation variation among trees in the field and in the common garden. Thus, we were able to address two questions: (1) given a uniform genetic background, do different environmental conditions result in differences in DNA methylation in Lombardy poplar? If so, (2) do these differences persist over time after clonal propagation in a common environment?

Materials and methods

Plant material and common garden design

Between February and March 2018, we sampled cuttings from *Populus nigra cv* 'Italica' clones in Europe across geographical gradients that spanned from 41° to 60° N and -5° to 25° E approximately, at twelve sampling sites that covered seven different Köppen-Geiger climate subtypes (Peel *et al.*, 2007). We tagged and georeferenced the source trees (hereafter referred to as "ortets"). During the first week of May 2018, we planted the cuttings (hereafter referred to as "ramets") in a common garden in the Marburg Botanical Garden (Germany) under a random block design. The common garden area was not shaded in any way, allowing the ramets to grow under direct sunlight. No herbicides, pesticides, or fertilizers were used in the common garden. We planted the ramets with 1 m between trees and watered them frequently for a period of five months until the end of summer. A more detailed description of sampling and the common garden set-up can be found in Díez Rodríguez et al., (2022).

Whole genome bisulfite sequencing

Of the 375 individuals considered to belong to the same genotype by Díez Rodríguez *et al*. (2022), we selected a subset for WGBS. We chose 14 ramets from 12 sampling sites from the common garden, except for those from Lithuania, of which only 10 ramets had survived in the garden, resulting in a total of 164 individuals (Supplementary Figure 1, right panel). From the original set of ortets, we chose 5 individuals from seven out of the 12 sampling sites, with a total of 35 individuals (Supplementary Figure 1, left panel). We collected leaf material from individuals, both in the field and in the common garden, at approximately the same time in July 2018. We extracted DNA from leaf tissue obtained from mature, healthy leaves dried in silica gel using the PeqGOLD Plant DNA mini kit (PEQLAB Biotechnologie GmbH, Erlangen, Germany). We used the NEBNext Ultra II DNA Library Prep Kit for sequencing library preparation, combined with EZ-96 DNA Methylation-Gold MagPrep (ZYMO) for bisulfite libraries. The protocol involved: i) end repair and 3' adenylation of sonicated DNA fragments, ii) NEBNext adaptor ligation and U excision, iii) size selection with AMPure XP Beads (Beckman Coulter, Brea, CA), iv) bisulfite treatment and cleanup of libraries, v) PCR enrichment and index ligation using Kapa HiFi Hot Start Uracil+ Ready Mix (Agilent) for bisulfite libraries (14 cycles), vi) final size selection and cleanup. Finally, we sequenced paired-end for 150 cycles on a HiSeq X Ten instrument (Illumina, San Diego, CA). All sequenced raw fastq files are available at the European Nucleotide Archive (ENA) database, under project number PRJEB44879.

Methylation data and DMR calling

For the methylation analysis we used the EpiDiverse toolkit (version 1.0), a pipeline suite for WGBS data analysis in non-model plant species (Nunn *et al.*, 2021). For alignment, quality control, and methylation extraction we used the WGBS pipeline. This pipeline uses FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to perform quality control, erne-bs5 (Prezza *et al*., 2012; http://erne.sourceforge.net/) to map raw reads, Picard MarkDuplicates (https://broadinstitute.github.io/picard/) to filter PCR duplicates and MethylDackel (https://github.com/dpryan79/MethylDackel) to perform the methylation calling. We mapped the samples to the *Populus nigra* cv 'Italica' reference genome, freely available at the European Nucleotide Archive (ENA) under project

number PRJEB44889. We only retained uniquely-mapping reads longer than 36 bp. On average, around 80% of the total number of reads were mapped to the reference genome. We calculated the bisulfite non-conversion rate using the mitochondrial genome, and found a mean rate of 0.005. Mapping stats and conversion rates for each individual sample are shown in Supplementary Table 1. Methylation levels for each called position were calculated according to Schultz et al. (2012) and using the following formula (*C* = reads supporting methylated cytosine, *T* = reads supporting unmethylated cytosine, *i* = position of cytosine):

$$(Ci /(Ci + Ti)) * 100$$

We obtained individual bedGraph files for each sample and context. We filtered out positions with a coverage lower than 6. For five ramet samples, less than 60% of the initial positions remained after filtering, and were thus excluded from the data set. We then merged the individual files into multisample bed files using custom scripts based on the *unionbedg* command from the BEDTools suite (Quinlan & Hall, 2010), retaining positions that were called in at least 80% of the samples. To directly compare only positions with methylation calls common to all samples, we obtained three different files per context. The first file contained 35 ortet samples (as mentioned in the plant material section); the second file contained 158 ramet samples; and lastly, the third file contained 64 paired ortet and ramet samples (32 samples from ortets and 32 from their respective ramets). A summary of the number of samples and the number of positions retained in each file is shown in Table 1. To study the epigenetic structure of the poplar clones, we ran Principal Component Analysis (PCA) per context using the *prcomp* function of the *stats* package (ver. 4.1.3; R Core Team, 2022).

Table 1. Summary of number of samples and positions included in each file used for methylation analysis

Type of file	N Samples	N positions per context		
		CpG	CHG	CHH
Ortet	35	8,318,522	13,678,685	76,501,469
Ramet	158	7,820,008	12,961,553	72,754,297
Paired	64	8,139,896	13,412,560	75,215,708

The EpiDiverse toolkit (Nunn *et al.*, 2021) includes a DMR pipeline that uses metilene (Jühling *et al.*, 2016) to call Differentially Methylated Regions (DMRs) between all possible pre-defined pairwise comparisons between sites for each sequence context. We used the default parameters of the DMR pipeline to define DMRs. In this study, each sampling site where the ramets were collected was considered as an individual group and compared to all the other sites. DMRs were called among three different group sets. First, we ran the DMR pipeline using only groups containing ortet samples in each pairwise comparison; second, we compared groups containing only ramet samples; and third, we compared ortet samples with their paired ramet samples. We then used custom scripts to summarize the results of the pipeline, and obtained a single file for each context and each run with a list of all DMRs, their genomic coordinates, and the specific pairwise comparison they belonged to. Supplementary Figure 1 shows a schematic description of the pairwise comparison design.

Variant calling, filtering and imputation

We used the EpiDiverse SNP pipeline (Nunn *et al.*, 2021, 2022) with default parameters to infer Single Nucleotide Polymorphisms (SNPs) from WGBS data. We combined the output of individual Variant Call Format (VCF) files from the ramet samples into a multisample VCF file using BCFtools (*v1.9*, Danecek et al., 2011). We filtered for variants successfully genotyped in at least 90% of individuals, with a minimum quality score of 30 and a minimum mean depth of 3. For the PCA analysis, we retained only biallelic SNPs and removed SNPs with more than 10% missing values and a Minor Allele Frequency (MAF) < 0.01. The remaining missing values were imputed with BEAGLE v 5.1 (Browning, Zhou, and Browning 2018). We also removed SNPs that were heterozygous in more than 95% of the samples. To reduce the number of SNPs for downstream analysis, we filtered redundant SNPs by pruning for Linkage Disequilibrium (LD) with a maximum LD of 0.8 between SNP pairs in a sliding window of 50 SNPs. After filtering and imputing, we were able to retain 343,977 SNPs. We performed the PCA analysis with PLINK (*v1.90b6.12*, Purcell et al., 2007) and plotted the results with custom scripts in R (https://github.com/EpiDiverse/scripts).

Correlation between methylation and bioclims

To assess correlations between methylome variation and climatic variables, we obtained bioclimatic data for each of the locations of the ortets from the CHELSA time-series data set (Karger *et al.*, 2017). The CHELSA data set covers the period between 1979 and 2013 and provides gridded data at a resolution of 30 arcsec (~ 1km). We included all 19 bioclimatic variables, as described in the CHELSA web page: https://chelsa-climate.org/bioclim/. Bioclimatic data for all sequenced individuals is available in Zenodo at https://doi.org/10.5281/zenodo.5995424. The methylation data for specific genomic regions used in the correlation analysis was obtained using the BEDTools *intersect* command (Quinlan and Hall, 2010) and a custom structural annotation. The annotations are available at the European Nucleotide Archive (ENA) under project number PRJEB44889. We correlated average global methylation levels with CHELSA bioclims using the Spearman method. The analysis was performed with the *corr.test* function of the *psych* package (ver. 2.2.5, Revelle, 2022) and plotted using the *heatmap.2* function of the *gplots* package (ver. 3.1.3, Warnes et al., 2022).

Mantel tests

To investigate if epigenetic distance between individual ramets was correlated with geographic, climatic and/or genetic distance, we performed mantel tests, using the *mantel* function of the *vegan* package (ver. 2.5-7; Oksanen et al., 2013). As input for the geographic and climatic distance matrices, we used the original geographic coordinates and the bioclimatic data of the ramets. We calculated two types of epigenetic distance matrices. The first matrix was based on the methylation levels of single methylated positions (MPs). In the second matrix, we used the BEDTools suit to merge the DMRs called from multiple pairwise comparisons in order to obtain a union set of candidate regions, variable between two or more populations of ramets. We then calculated mean methylation levels (according to Schultz et al. 2012) in each region. For the genetic distance matrix we used the same SNPs that were used for the genetic structure analysis. To standardize the data and make it comparable, we then conducted a PCA and calculated the first three PCs for each type of input. We then created Euclidean distance matrices using the *dist* function of the R *stats* package (Version 4.2.1, R core team, 2022). Finally, we ran the mantel tests with the Pearson correlation method and 9999 permutations.

Persistence of DNA methylation patterns

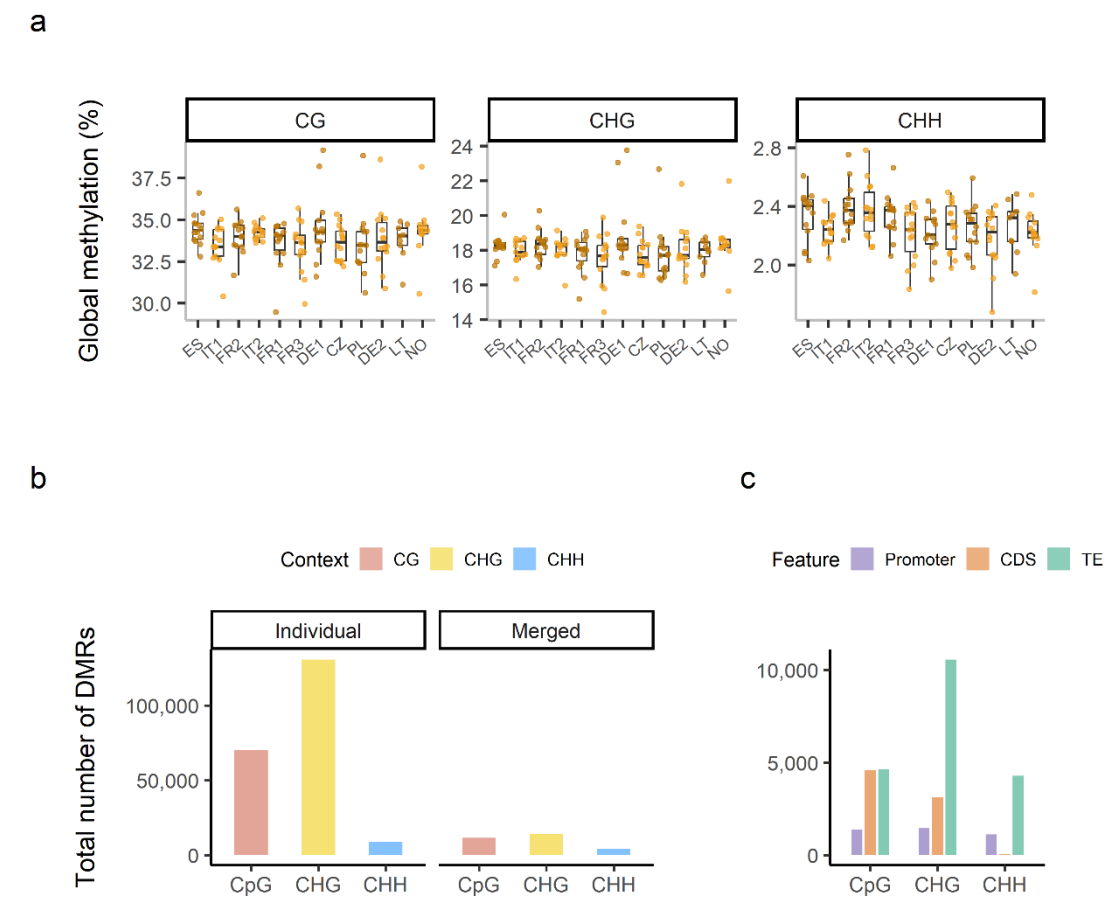
To study if methylation patterns were conserved across clonal generations, we focused on the seven sites for which we had collected samples from ortets and ramets. We called DMRs between sites for ortets, for ramets, and between ortets and ramets from each site. Supplementary Figure S2 shows the total number of DMRs for each pairwise comparison among ortets (A) and ramets (B), respectively, ordered according to latitude of origin from South to North. If methylation patterns are conserved in the next clonal generation we assumed we would be able to find the same DMRs when comparing the same sampling-site pairs between ortets and between ramets. We therefore intersected the bed files with all the DMRs called using the BEDTools intersect command. Specifically, we intersected a file containing DMRs called from group A vs group B ortets with a file containing DMRs called from ramets belonging to the same groups (i.e. corresponding to the clonal offspring). We then repeated the analysis for each of the 21 possible pairwise comparisons between sites. Supplementary Figure S4 shows a detailed count of hypermethylated and hypomethylated DMRs for each pairwise comparison. After running the intersections, we created individual files containing all the regions found among ramets that overlap with regions found among ortets.

Results

Methylation profiles in the Lombardy poplar

Average global methylation in ramets of the Lombardy poplar from 12 different sampling sites ranged from 30 to 40% in the CG context, 15-25% in the CHG context and 1-3% in the CHH context (Figure 1A). We did not find any statistically significant differences among methylation levels from different sites in any of the contexts, and variation within each group seemed to be higher than the variation among groups. We found the highest number of DMRs in the CHG context (~130,000 DMRs), followed by the CG context (~ 70,000 DMRs) and the CHH context, where only ~ 9,100 DMRs were called among all pairwise comparisons among sites (Figure 1B). However, most of these DMRs were common to two or more comparisons. When common DMRs were merged into unique regions, we found around 11,400 CG-DMRs, 14,400 CHG-DMRs and 4,100 CHH-DMRs. The length of the merged DMRs ranged from 10 to around 5,000 bases (Supplementary

figure S3). Of these DMRs, a considerable fraction overlapped with annotated transposable elements (TEs) in all sequence contexts (~4,600; ~ 10,500 and 4,200, respectively for CG, CHG and CHH). Interestingly, only 31 DMRs in the CHH context overlapped with coding sequences (CDS), while around 4,600 CG- and 3,100 CHG-



DMRs overlapped with these regions (Figure 2c).

Figure 1. Methylation profiles in the Lombardy poplar (ramets). **a**. Variation in methylation levels among ramets across geographical gradients in all sequence contexts. Sites are ordered from South to North according to their geographic coordinates and labeled by the sample site code (ISO 3166 standard country code): ES: Spain, n = 14 ; IT1: Italy 1, n = 13; FR2: France 2, n = 13; IT2: Italy 2, n = 14; FR1: France 1, n = 14; FR3: France 3, n = 14; DE1: Germany 1, n = 13; CZ: Czech Republic, n = 14; PL: Poland, n = 14; DE2: Germany 2, n = 14; LT: Lithuania, n = 9; NO: Norway, n = 12. Note the different scales in the Y axes (n = 158). **b.** Total number of DMRs in each sequence context, called from all pairwise comparisons (n = 158). Number of individual DMRs corresponds to every single DMR. Total number of merged DMRs corresponds to merged regions found in two or more comparisons **C.** Total number of merged DMRs overlapping specific genomic features in each sequence context (n = 158)

Genetic and epigenetic structure

To investigate a potential relationship between genetic and epigenetic structure in the Lombardy poplar, we conducted a Principal Component Analysis (PCA) based on methylated positions (MPs) and SNPs inferred from WGBS of the ramet samples. Among the sequenced 'Italica' clones, we did not find any clear genetic or epigenetic structure that could be associated with the geographic origin of the ramets (Figure 2). As explained in Díez Rodríguez et al. (2022), the ramets that belonged to the 'Italica' cluster had a mean number of pairwise differences among individual ramets of around 96 SNPs out of the 4.906 investigated remaining positions. We targeted 4,906 loci equally distributed across the 19 *P. nigra* chromosomes selected from a larger set identified in Scaglione et. al (2019), which should allow for accurate and effective genotyping of population groups. To further assess if the loci targeted were actually sufficient for genotyping the populations analyzed, we called SNPs from the WGBS data. In this way, we increased the number of SNPs available for the study to 986,948 SNPs, mostly reflecting heterozygosity of the clonal genotype, not genetic differences between samples. After we removed SNPs heterozygous in > 95% of the samples and performed the pruning step, 343,977 SNPs remained for the analysis. Still, we did not find any structure that could be associated with geographic patterns.

Furthermore, when running PCA with MPs inside CDS (Figure S4), we observed some grouping, but this was not explained by any of the environmental variables that we tested (such as habitat type, elevation or habitat disturbance level). Despite the lack of epigenetic structure, some individuals with the same site of origin seemed to group together (Figures 2 and S5), indicating similar methylation profiles

.

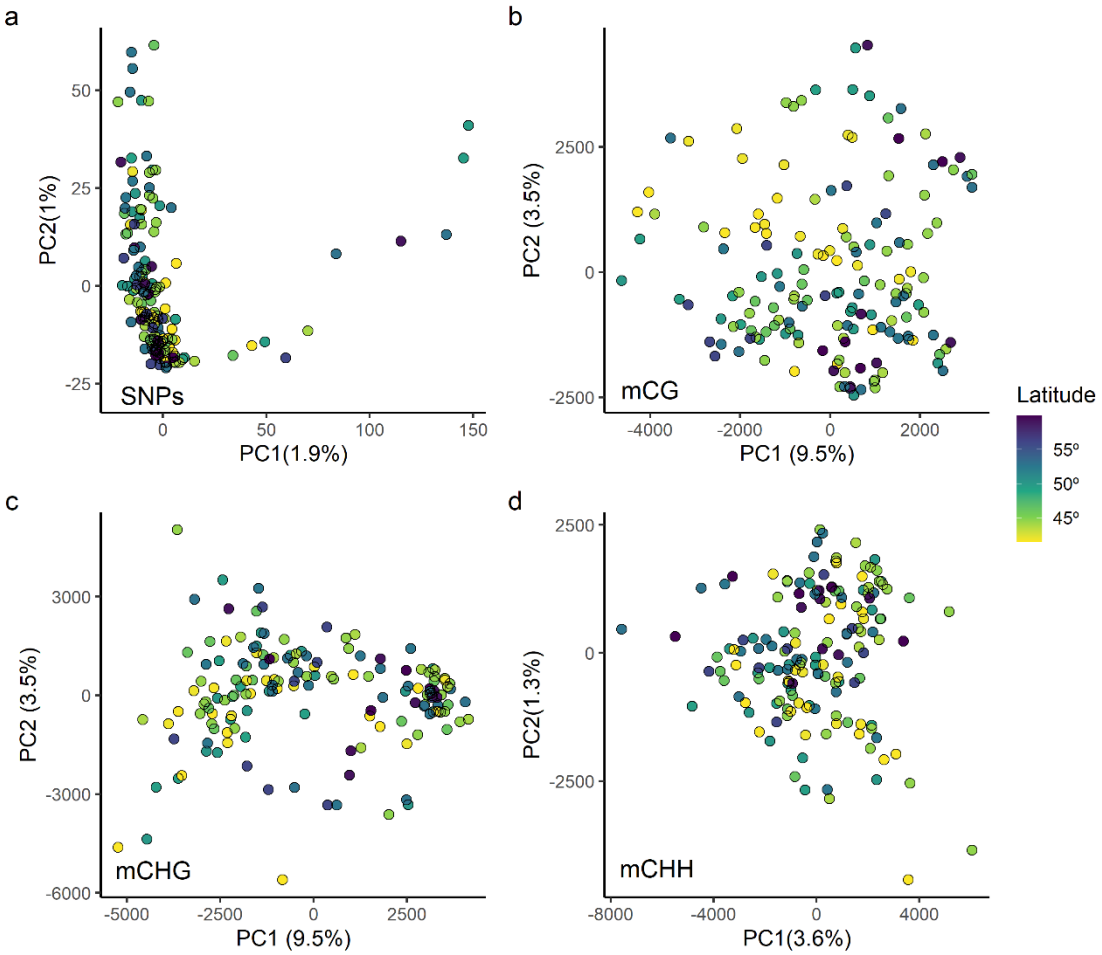


Figure 2. Genetic and epigenetic structure of the poplar ramets, colored according to latitude of origin. **A:** Genetic structure based on the SNPs called from WGBS data. **B-D**: Epigenetic structure for the CG (B), CHG (C) and CHH (D) sequence contexts.

Relationship between methylation, geographic origin, and climate

To assess if there was any relationship between epigenetic variation, genetic variation, geographic origin, and climatic conditions, we analyzed the correlation between epigenetic distance and genetic, geographic, and climatic distance in ramets using mantel tests. We first correlated geographic with climatic distance, and genetic with both geographic and climatic distance. We found that climatic distance correlated with geographic distance (R = 0.7, p = 0.001), but genetic distance was not correlated with geographic distance or climatic distance (R = -0.03, p = ns, in both tests). We created epigenetic distance matrices based on MPs and DMRs. We did not find any correlation

between epigenetic and genetic distance in any case, except for the MPs in the CHH context (Table 2). However, epigenetic distance significantly correlated with geographic and climatic distance in almost all cases. The highest correlation coefficients were found in the CG context between DMR-based epigenetic distance and both geographic and climatic distance (R=0.164 and p < 0.001, and R=0.141 and p < 0.001, respectively). Because geographic distance and climatic distance were strongly correlated, we ran partial mantel tests between epigenetic distance and climatic distance accounting for the geographic distance. In this case, most of the significant correlations disappeared, except for MPs in the CHH context.

Table 2. Mantel test coefficients for the correlation between epigenetic distance and genetic, geographic, and climatic distance in ramets. Epigenetic distance was tested both as individual methylated positions (MPs) and differentially methylated regions (DMRs). Significant correlations are highlighted in bold font.

	Context	Genetic distance		Geographic distance		Climatic distance		Climatic distance (partial)	
		R	p	R	p	R	p	R	p
MPs	CG	0.001	0.445	**0.093**	**0.002**	**0.082**	**0.001**	0.015	0.258
	CHG	0.009	0.330	**0.079**	**0.011**	**0.068**	**0.005**	0.012	0.307
	CHH	0.035	0.231	0.043	0.115	**0.067**	**0.012**	**0.054**	**0.036**
DMRs	CG	0.043	0.202	**0.164**	**<0.001**	**0.141**	**<0.001**	0.023	0.192
	CHG	0.005	0.381	**0.080**	**0.003**	**0.072**	**0.001**	0.015	0.237
	CHH	-0.008	0.536	**0.063**	**0.021**	**0.064**	**0.005**	0.025	0.152

To study the association between methylation patterns and climate of origin in more detail, we conducted a correlation analysis between global methylation levels in specific genomic features (i.e., promoters, coding sequences (CDS) and TEs) and bioclimatic variables (Figure 3). We found significant correlations in all sequence contexts, with the highest number of correlations observed in the CHH context. In fact, for the CHH context, we found correlations between all three genomic features and most temperature-related bioclimatic variables, such as maximum temperature and mean temperature related variables. Additionally, methylation levels in promoters and TEs in this context were

negatively correlated with both latitude and longitude. On the other hand, methylation levels in the CG and CHG contexts showed no correlation with climatic variables, except for methylation in promoter and CDS regions and three precipitation variables (precipitation in the wettest month and wettest quarter, and precipitation in the warmest quarter). Furthermore, variables in CHH were grouped in a separate cluster while CG and CHG variables grouped mainly by genomic features.

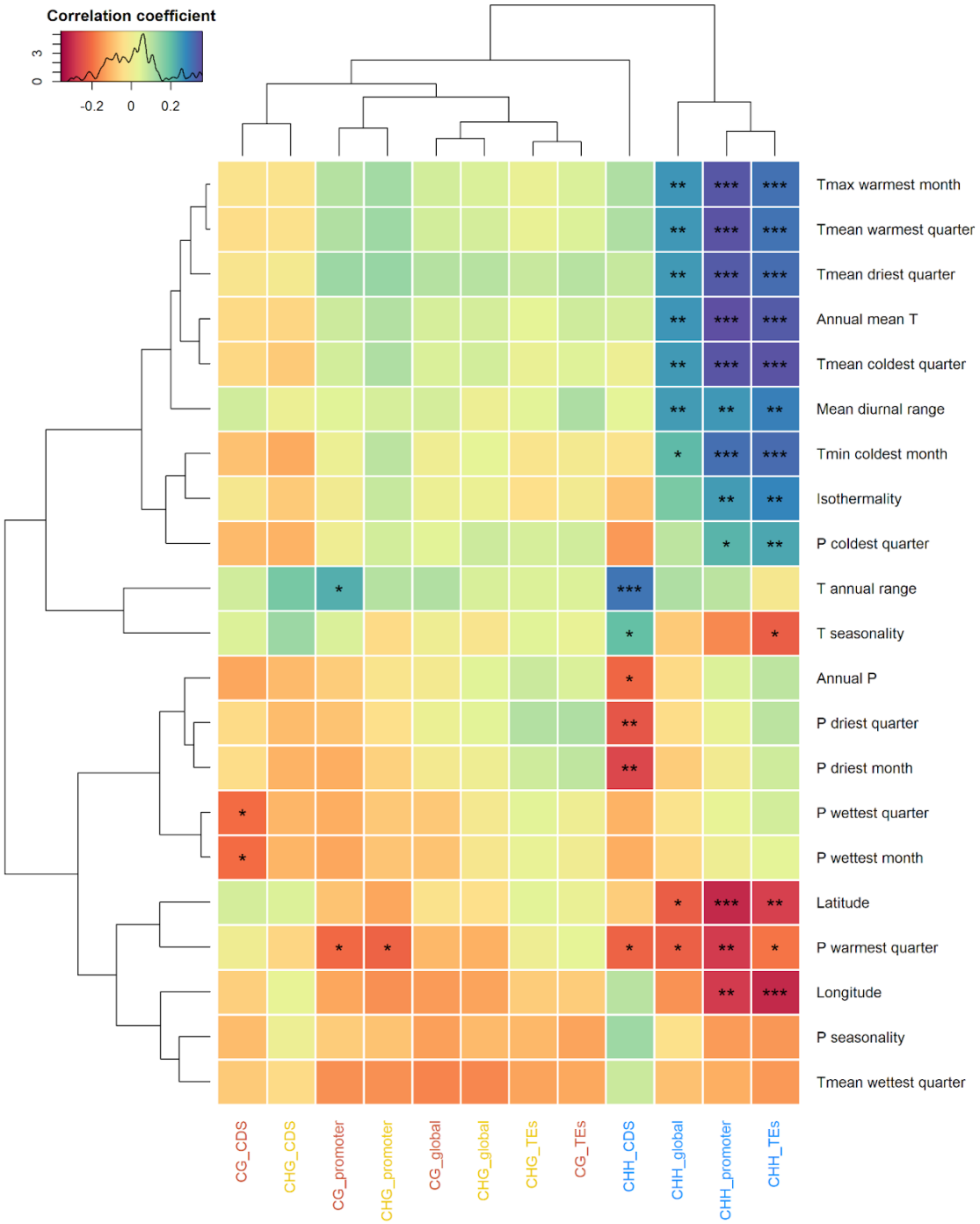


Figure 3. Spearman correlation analysis between global methylation levels in different genomic features and bioclimatic variables extracted from the CHELSA database. P = precipitation, T = temperature. P-values are adjusted for multiple pairwise comparisons using the "BH" method. Statistically significant correlations are labeled with the following code: $p < 0.001$ = ***; $p < 0.01$ = **; $p < 0.05$ = *. Correlations are grouped using the hierarchical clustering method.

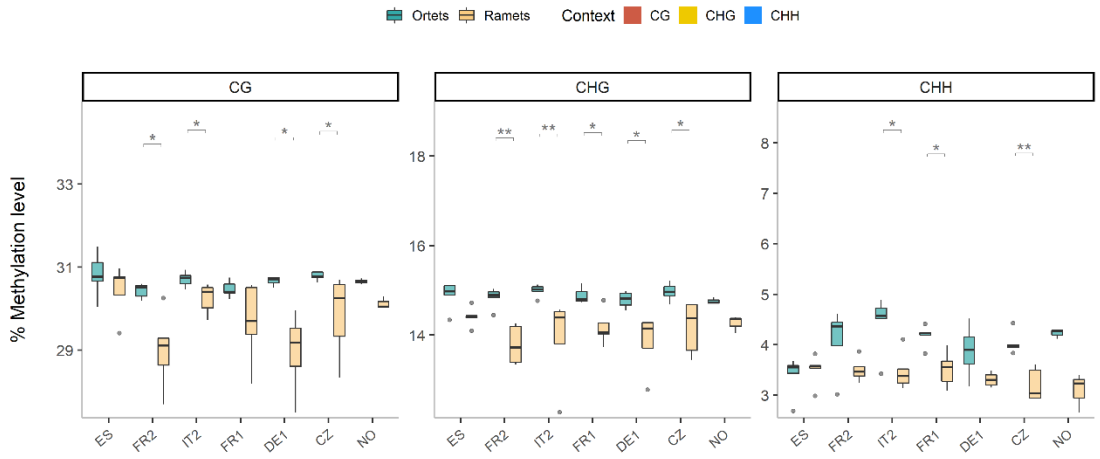
Persistence of DNA methylation patterns across clonal generations

To investigate if methylation patterns can be transmitted to the next clonal generation, we first compared average global methylation levels between ortets (parental individuals) and ramets (clonal offspring). In the ortets, methylation levels were consistently higher in all contexts (Figure 4A). The difference in global methylation levels between ortets and ramets was further evidenced by the number of hypermethylated ortet-vs-ramet DMRs (Figure 4B). When comparing ortets with their ramets, the number of DMRs in the CG context was considerably low for some groups (e.g., ES, IT2, FR1, CZ, NO), and the lowest of all contexts (10,180 total DMRs vs. 31,600 and 13,601 for CHG and CHH, respectively). On the other hand, the number of DMRs in the CHG and CHH contexts was more variable among different sites. Additionally, we conducted a PCA analysis using the paired clones (Figure S6) and found that pairs tended to group together, especially in the CG context.

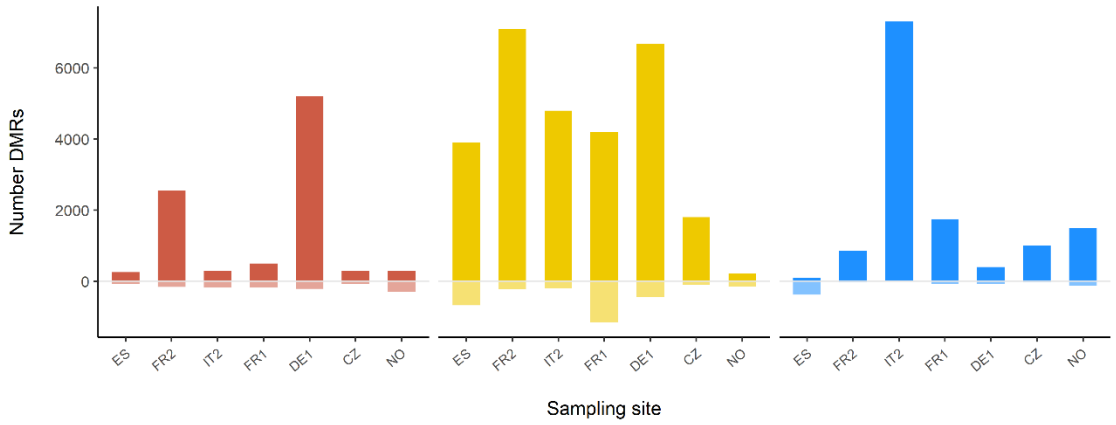
To further assess if methylation patterns persisted across clonal generations, we then intersected the DMRs found between pairwise comparisons in the ortets and the DMRs found between the ramets (Figure 4C). Between 25% and 50% of the ortet DMRs in CG and CHG overlapped with ramet DMRs. This percentage was considerably lower in the case of the CHH context, where less than 10% of the DMRs were also found in the ramets.

Figure 4. A. Differences in global methylation levels between ortets (green) and their paired ramets (orange), for each sequence context. Statistically significant correlations are labeled with the following code: $p < 0.001$ = ***; $p < 0.01$ = **; $p < 0.05$ = *. P values were adjusted for multiple pairwise testing using the "BH" method. B. Total number of hypermethylated (above the 0 line) and hypomethylated (below the 0 line) DMRs between ortets and their paired ramets. C. Percentage of DMRs among ramet pairwise comparisons that overlap with DMRs among ortet pairwise comparisons. Each bar represents a pairwise comparison between ortets from each sampled site in Europe and the ramets of the same individuals. The dashed line indicates the threshold for 50% of ramet DMRs that overlap with ortet DMRs.

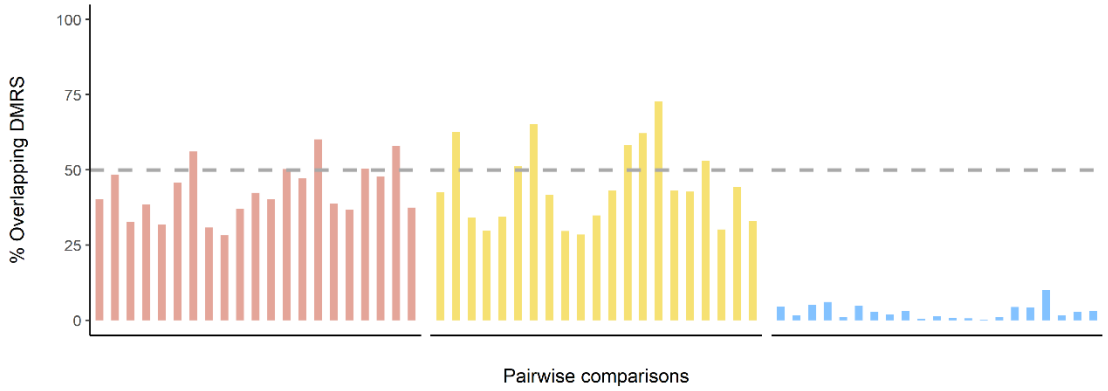
a


Ortets
Ramets
Context
CG
CHG
CHH
% Methylation level
ES
FR2
IT2
FR1
DE1
CZ
NO

b


Number DMRs
Sampling site

c


% Overlapping DMRS
Pairwise comparisons

Discussion

So far only few studies have used epigenomics to investigate the effects of environmentally induced epigenetic variation at a landscape level. Here, we present the first landscape-scale investigation of DNA methylation patterns in a system that has been almost exclusively clonally propagated. We found that average methylation levels were significantly correlated with climatic variables and persisted across at least one clonal generation, despite the lack of evident genetic or epigenetic structure.

The lack of genetic structure can be explained by the very low genetic diversity found by genotyping the poplar clones (ramets) established in our common garden (Figure 2a). This was expected, given the clonal history of the 'Italica' cultivar. The 'Italica' cultivar likely originated from a single male clone in Central Asia, from where it spread to Europe. It is widely accepted that this clone was further artificially propagated from an individual or group of individuals found in Lombardy, Italy (Elwes and Henry, 1913). Our results suggest that a major fraction of the clones across Europe do indeed share a common line.

In a similar fashion, we did not find any clear epigenetic population structure but there appears to be some grouping in the CG context (Figure 2b) and epigenetic distance was positively correlated with geographic distance (Table 2). Furthermore, MPs inside CDS regions do show a pattern, but it was not explained by any of the environmental variables used in the analysis. This evidence points to the importance of other sources of epigenetic variation, such as genetic somatic mutations or stochastic epimutations. Several studies have reported age-related changes in the levels of cytosine methylation due to spontaneous methylation changes (Fraga *et al.*, 2002; Dubrovina & Kiselev, 2016). Furthermore, Hofmeister *et al*. (2020) found evidence that spontaneous methylation changes are cumulative across somatic development in the close relative *Populus trichocarpa*, and that they have a higher rate than genetic mutations. Considering that the 'Italica' cultivar has been artificially propagated for the last two centuries, stochastic epimutations have likely accumulated across several clonal generations, confounding any environmentally induced epigenetic population structure.

Previous studies on population epigenomics have found that epigenetic variation is associated with genetic variation in Brassicaceae (Dubin *et al*, 2015, Kawakatsu *et al*., 2016; Galanti *et al*., 2022), thus hindering the study of the relationship between environmental epigenetic variation and climatic conditions. The use of a clonal cultivar circumvents this problem. We used mantel tests to investigate if epigenetic distance,

measured as the distance between both single methylated variants (MPs) and differentially methylated regions (DMRs), was correlated with genetic, geographic and/or climatic distance (Table 2). We found that epigenetic distance did not correlate with genetic distance in all cases except one (MPs in the CHH context) but correlated with both geographic distance and climatic distance in almost all cases (see also Figure S2). However, when accounting for geographic distance, the correlations with climatic distance disappeared, except for MPs in the CHH context. As suggested above, if stochastic epimutations are contributing to a major fraction of the epigenetic variation, the correlation between epigenetic distance and geographic distance could be explained by isolation-by-distance processes, since this cultivar was gradually propagated across Europe (Slatkin, 1993). This evidence thus suggests that epigenetic variation of the individuals analyzed might be both under environmental and stochastic control.

To assess whether the methylation profiles under climatic control could potentially have a functional role, we extracted the methylation levels of specific genomic features (gene promoters, gene body and transposable elements, specifically). We then correlated methylation levels with individual bioclimatic variables (Figure 3). Methylation levels were strongly correlated with most temperature variables, particularly in the case of gene promoters and TEs in the CHH context, which would also explain the correlation with latitude and longitude. Our results are in line with previous studies that have reported the potential effects of temperature on DNA methylation in several plant organisms (Dubin et al., 2015; Conde et al., 2017; Zhang et al., 2018; Galanti et al., 2022; Sammarco et al., 2022). On the other hand, methylation levels correlated with very few precipitation variables but, as opposed to temperature variables, we observed more significant correlations in the CG and CHG context. It is conceivable that a certain degree of environmental information regarding water availability might be encoded in more stable methylation contexts and transmitted to the clonal offspring, since *Populus nigra* is a riparian species that depends on river flooding regimes for successful seed and cutting dispersal (Smulders *et al.*, 2008). Nevertheless, our results indicate that methylation patterns in CHH might be highly dynamic and rapidly respond to new environmental cues. This assumption is further supported by the changes in global methylation levels observed between ortet-ramet pairs (Figure 4A). Although there were almost no differences in methylation levels between individuals from different geographic origins in any of the contexts, methylation levels were significantly higher in the ortets than in the corresponding ramets for many locations. In poplar, methylation levels have been shown to increase under drought conditions (Raj *et al*., 2011; Peña Pontón *et al,* 2022). Given that 2018 was a year characterized by particularly extreme drought events in Europe,

and the ramets were well watered during the whole summer, it is possible that the differences in methylation levels between ortets and ramets are the result of differences in water availability. Furthermore, we observed a considerable decrease in the number of DMRs found among ramets (Supplementary Figure 2), suggesting that methylation profiles in leaves in the CHH context might have already adjusted to the new conditions of the common garden.

Despite these dynamic changes in CHH methylation, a considerable fraction of the methylation patterns appeared to be transmitted to the clonal offspring, particularly in the CG and CHG contexts. We found that approximately 25% of the DMRs in CG and CHG called from pairwise comparisons among the ramets of different sampled sites overlapped with the DMRs found among the ramets of the same pairwise comparison (Figure 4C). The fact that we could find these specific regions both in the ortets and the ramets provides further evidence that methylation patterns in the CG and CHG contexts can potentially be transmitted to the clonal offspring. Conversely, less than 10% of the DMRs found in the CHH context were transmitted to the next clonal generation. This further supports our conclusion that methylation in the CHH context is highly dynamic.. It is, however, challenging to determine if there was an active change in the methylome as a result of new environmental cues, or if these patterns are established *de novo* every year in leaf tissue. If in fact leaf CHH methylation patterns are determined in every new season, this could possibly explain the low number of DMRs observed in the CHH context, both among the ortets and the ramets (Supplementary Figure S2). If the environmental conditions in the common garden resemble those of the original sites, then the methylome in CHH in the ramets would also resemble the methylome of the ortets. If the conditions are nothing alike, then a higher number of DMRs would be expected. Based on the total number of DMRs, the latter might be true. The number of DMRs was considerably higher when comparing ortets sampled in Spain with ortets sampled in Northern European sites (Figure S2), while only a few DMRs were found between sites that belong to similar Köppen climatic areas (e.g., FR1 vs FR2). In the common garden, however, where the environmental conditions were the same for all the individuals, the number of total DMRs between ramets from different sites was very low, suggesting that the ramets might have rapidly adjusted to common garden conditions. As proposed by Ito and colleagues (2019), DNA methylation in natural environments might have two components, genomic regions that might change dynamically and epigenetic marks for stable gene expression that are rather fixed. If this is the case, it opens interesting new research possibilities, if a certain fraction of epigenetic information is stored in symmetrical stable contexts, but some of it can rapidly shift to reflect new

environments. In practical terms, this would imply that methylation variation is partitioned in distinct “modules”, and further experiments should target individual sources of environmentally induced epigenetic variation.

In summary, our study is the first landscape-scale investigation of DNA methylation patterns in a system that has been almost exclusively clonally propagated. We found that methylation patterns in the Lombardy poplar are independent of genetic structure, but that methylation profiles are associated with climatic conditions. Furthermore, we have shown that a fraction of DMRs is transmitted to the next clonal generation, and that methylation in the CHH levels is highly dynamic and might rapidly adjust to new environmental conditions. Our results suggest that the CHH context is the most responsive to changing environments and that the stability of induced changes across clonal generations is stronger in CG and CHG. We have shown that the Lombardy poplar is a valuable system to study environmentally induced epigenetic variation in a naturally occurring near-isogenic population, with limited confounding genetic variation. Our study provides further insight into how methylation patterns in natural populations might vary along geographic and climatic gradients. However, further research is necessary to assess whether DNA methylation can have an effect on phenotypic plasticity. The high resolution methylome data generated in our experiment is a significant resource for Epigenome Wide Association Studies (EWAS), and can considerably contribute to our understanding of how methylation variation affects plant acclimation and adaptation.

Acknowledgements

We thank members of the EpiDiverse consortium (www.epidiverse.eu) for valuable inputs during preparation and execution of the study, and reviewers for useful comments. For computing, we acknowledge Prof. Peter Stadler at the University of Leipzig and David Langenberger from ecSeq, for hosting the EpiDiverse servers

Funding

This work was supported by the European Training Network “EpiDiverse” and received funding from the EU Horizon 2020 program under Marie SkłodowskaCurie grant agreement No 764965; the work was further supported by the Austrian Academy of Sciences (ÖAW).

Data availability

The *Populus nigra* cv 'Italica' reference genome and the genome annotations are freely available at the European Nucleotide Archive (ENA) under project number PRJEB44889. All sequenced raw fastq files are available under project number PRJEB44879. Bioclimatic data for all sequenced individuals is available in Zenodo at https://doi.org/10.5281/zenodo.5995424.

References

Ally D, Ritland K, Otto SP. **2008**. Can clone size serve as a proxy for clone age?An exploration using microsatellite divergence in Populus tremuloides. *Molecular Ecology* **17**: 4897–4911.

Avramidou E v., Ganopoulos I v., Doulis AG, Tsaftaris AS, Aravanopoulos FA. **2015**. Beyond population genetics: natural epigenetic variation in wild cherry (Prunus avium). *Tree Genetics and Genomes* **11**: 1–9.

Balloux F, Lehmann L, de Meeûs T. **2003**. The Population Genetics of Clonal and Partially Clonal Diploids. *Genetics* **164**: 1635–1644.

Bates D, Mächler M, Bolker B, Walker S. **2015**. Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software* **67**.

Becker C, Hagmann J, Müller J, Koenig D, Stegle O, Borgwardt K, Weigel D. **2011**. Spontaneous epigenetic variation in the Arabidopsis thaliana methylome. *Nature 2011 480:7376* **480**: 245–249.

Becker C, Weigel D. **2012**. Epigenetic variation: origin and transgenerational inheritance. *Current Opinion in Plant Biology* **15**: 562–567.

Boquete MT, Muyle A, Alonso C. **2021**. Plant epigenetics: phenotypic and functional diversity beyond the DNA sequence. *American Journal of Botany* **108**: 553–558.

Bräutigam K, Vining KJ, Lafon-Placette C, Fossdal CG, Mirouze M, Marcos JG, Fluch S, Fraga MF, Guevara MÁ, Abarca D, *et al.* 2013. Epigenetic regulation of adaptive responses of forest tree species to the environment. *Ecology and Evolution* **3**: 399–415.

vanden Broeck A, Cox K, Brys R, Castiglione S, Cicatelli A, Guarino F, Heinze B, Steenackers M, vander Mijnsbrugge K. **2018**. Variability in DNA Methylation and Generational Plasticity in the Lombardy Poplar, a Single Genotype Worldwide Distributed Since the Eighteenth Century. *Frontiers in Plant Science* **9**: 1635.

CABI: ***Populus Nigra (Black Poplar)*****.** https://www.cabi.org/isc/datasheet/43535. Accessed 24 Jan. 2022.

Chenault N, Arnaud-Haond S, Juteau M, Valade R, Almeida JL, Villar M, Bastien C, Dowkiw A. **2011**. SSR-based analysis of clonality, spatial genetic structure and introgression from the Lombardy poplar into a natural population of Populus nigra L. along the Loire River. *Tree Genetics & Genomes 2011 7:6* **7**: 1249–1262.

Climate Change 2022: Impacts, Adaptation and Vulnerability | Climate Change 2022: Impacts, Adaptation and Vulnerability.

Conde D, le Gac AL, Perales M, Dervinis C, Kirst M, Maury S, González-Melendi P, Allona I. **2017**. Chilling-responsive DEMETER-LIKE DNA demethylase mediates in poplar bud break. *Plant, Cell & Environment* **40**: 2236–2249.

Core Team R. **2017**. R: A language and environment for statistical computing. R Foundation for Statistical Computing. *Vienna, Austria: URL https://www. R-project. org/.[Google Scholar]*.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, *et al.* 2011. The variant call format and VCFtools. *Bioinformatics* **27**: 2156–2158.

Díez Rodríguez B, Peña C, Pérez-Bello P, Bette J, Lerbs L, Mackenbach T, Wulle S, de Paoli E, Verhoeven KJF, Heer K, *et al.* 2022. An Uncommon Garden Experiment: Microenvironment Has Stronger Influence on Phenotypic Variation than Epigenetic Memory in the Clonal Lombardy Poplar. *SSRN Electronic Journal*.

Dubin MJ, Zhang P, Meng D, Remigereau MS, Osborne EJ, Casale FP, Drewe P, Kahles A, Jean G, Vilhjálmsson B, *et al.* 2015. DNA methylation in Arabidopsis has a genetic basis and shows evidence of local adaptation. *eLife* **4**.

Dubrovina AS, Kiselev K v. 2016. Age-associated alterations in the somatic mutation and DNA methylation levels in plants. *Plant Biology* **18**: 185–196.

Elwes HJ, Henry A. **1913**. The trees of Great Britain & Ireland. R. & R. Clark, Limited, Edinburgh, 2022 pp.

Finnegan EJ, Genger RK, Peacock WJ, Dennis ES. **2003**. DNA METHYLATION IN PLANTS. *http://dx.doi.org/10.1146/annurev.arplant.49.1.223* **49**: 223–247.

Fisher RA. **1958**. *The genetical theory of natural selection*. Рипол Классик.

Fraga MF, Rodriguez R, Canal MJ. **2002**. Genomic DNA methylation-demethylation during aging and reinvigoration of Pinus radiata. *Tree Physiology* **22**: 813–816.

Fu YB. **2015**. Understanding crop genetic diversity under modern plant breeding. *Theoretical and Applied Genetics* **128**: 2131–2142.

Galanti D, Ramos-Cruz D, Nunn A, Rodríguez-Arévalo I, Scheepens JF, Becker C, Bossdorf O. **2022**. Genetic and environmental drivers of large-scale epigenetic variation in Thlaspi arvense (NM Springer, Ed.). *PLOS Genetics* **18**: e1010452.

Gáspár B, Bossdorf O, Durka W. **2019**. Structure, stability and ecological significance of natural epigenetic variation: a large-scale survey in Plantago lanceolata. *New Phytologist* **221**: 1585–1596.

Gugger PF, Fitz-Gibbon S, Pellegrini M, Sork VL. **2016**. Species-wide patterns of DNA methylation variation in Quercus lobata and their association with climate gradients. *Molecular Ecology* **25**: 1665–1680.

Herman JJ, Sultan SE. **2016**. DNA methylation mediates genetic variation for adaptive transgenerational plasticity. *Proceedings of the Royal Society B: Biological Sciences* **283**.

Herrera CM, Medrano M, Bazaga P. **2017**. Comparative epigenetic and genetic spatial structure of the perennial herb *Helleborus foetidus*: Isolation by environment, isolation by distance, and functional trait divergence. *American Journal of Botany* **104**: 1195–1204.

Hofmeister BT, Denkena J, Colomé-Tatché M, Shahryary Y, Hazarika R, Grimwood J, Mamidi S, Jenkins J, Grabowski PP, Sreedasyam A, *et al.* 2020. A genome assembly and the somatic genetic and epigenetic mutation rate in a wild long-lived perennial Populus trichocarpa. *Genome Biology* **21**: 1–27.

Hughes AR, Inouye BD, Johnson MTJ, Underwood N, Vellend M. **2008**. Ecological consequences of genetic diversity. *Ecology Letters* **11**: 609–623.

Ingvarsson PK, Dahlberg H. **2018**. The effects of clonal forestry on genetic diversity in wild and domesticated stands of forest trees. *https://doi.org/10.1080/02827581.2018.1469665* **34**: 370–379.

Ito, Nishio, Tarutani, Emura, Honjo, Toyoda, Fujiyama, Kakutani, Kudoh. **2019**. Seasonal Stability and Dynamics of DNA Methylation in Plants in a Natural Environment. *Genes* **10**: 544.

Jansson S, Douglas CJ. **2007**. *Populus*: A Model System for Plant Biology. *Annual Review of Plant Biology* **58**: 435–458.

Johannes F, Porcher E, Teixeira FK, Saliba-Colombani V, Simon M, Agier N, Bulski A, Albuisson J, Heredia F, Audigier P, *et al.* 2009. Assessing the Impact of Transgenerational Epigenetic Variation on Complex Traits. *PLOS Genetics* **5**: e1000530.

Johannes F, Schmitz RJ. **2019**. Spontaneous epimutations in plants. *New Phytologist* **221**: 1253–1259.

Jühling F, Kretzmer H, Bernhart SH, Otto C, Stadler PF, Hoffmann S. **2016**. metilene: fast and sensitive calling of differentially methylated regions from bisulfite sequencing data. *Genome Research* **26**: 256–262.

Karger DN, Conrad O, Böhner J, Kawohl T, Kreft H, Soria-Auza RW, Zimmermann NE, Linder HP, Kessler M. **2017**. Climatologies at high resolution for the earth's land surface areas. *Scientific Data* **4**: 170122.

Kawakatsu T, Huang S shan C, Jupe F, Sasaki E, Schmitz RJJ, Urich MAA, Castanon R, Nery JRR, Barragan C, He Y, *et al.* 2016. Epigenomic Diversity in a Global Collection of Arabidopsis thaliana Accessions. *Cell* **166**: 492–505.

de Kort H, Panis B, Deforce D, van Nieuwerburgh F, Honnay O. **2020**. Ecological divergence of wild strawberry DNA methylation patterns at distinct spatial scales. *Molecular Ecology* **29**: 4871–4881.

Lämke J, Bäurle I. **2017**. Epigenetic and chromatin-based mechanisms in environmental stress adaptation and stress memory in plants. *Genome Biology* **18**.

Lira-Medeiros CF, Parisod C, Fernandes RA, Mata CS, Cardoso MA, Ferreira PCG. **2010**. Epigenetic Variation in Mangrove Plants Occurring in Contrasting Natural Environment. *PLOS ONE* **5**: e10326.

Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR. **2008**. Highly Integrated Single-Base Resolution Maps of the Epigenome in Arabidopsis. *Cell* **133**: 523–536.

Matzke MA, Mosher RA. **2014**. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nature Reviews Genetics 2014 15:6* **15**: 394–408.

Medrano M, Herrera CM, Bazaga P. **2014**. Epigenetic variation predicts regional and local intraspecific functional diversity in a perennial herb. *Molecular Ecology* **23**: 4926–4938.

Meloni M, Reid A, Caujapé-Castells J, Marrero Á, Fernández-Palacios JM, Mesa-Coelo RA, Conti E. **2013**. Effects of clonality on the genetic variability of rare, insular species: the case of Ruta microcarpa from the Canary Islands. *Ecology and Evolution* **3**: 1569–1579.

Moore LD, Le T, Fan G. **2012**. DNA Methylation and Its Basic Function. *Neuropsychopharmacology 2013 38:1* **38**: 23–38.

Nicotra AB, Segal DL, Hoyle GL, Schrey AW, Verhoeven KJF, Richards CL. **2015**. Adaptive plasticity and epigenetic variation in response to warming in an Alpine plant. *Ecology and Evolution* **5**: 634–647.

Niederhuth CE, Schmitz RJ. **2017**. Putting DNA methylation in context: from genomes to gene expression in plants. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1860**: 149–156.

Nunn A, Can SN, Otto C, Fasold M, Diéz Rodríguez B, Fernández-Pozo N, Rensing SA, Stadler PF, Langenberger D. **2021**. EpiDiverse Toolkit: a pipeline suite for the analysis of bisulfite sequencing data in ecological plant epigenetics. *NAR Genomics and Bioinformatics* **3**.

Nunn A, Otto C, Fasold M, Stadler PF, Langenberger D. **2022**. Manipulating base quality scores enables variant calling from bisulfite sequencing alignments using conventional bayesian approaches. *BMC Genomics* **23**: 1–10.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H. **2013**. Package 'vegan'. *Community ecology package, version* **2**: 1–295.

Peel MC, Finlayson BL, McMahon TA. **2007**. Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences* **11**: 1633–1644.

Peña-Ponton, C., Diez-Rodriguez, B., Perez-Bello, P., Becker, C., McIntyre, L. M., Putten, W. van der, Paoli, E. de, Heer, K., Opgenoorth, L., & Verhoeven, KJF. (2022). High-resolution methylome analysis in the clonal Populus nigra cv. 'Italica' reveals environmentally sensitive hotspots and drought-responsive TE superfamilies. *BioRxiv*, 2022.10.18.512698.

Peter Meyer IN, Lohuis M ten. **1994**. Evidence for cytosine methylation of non-symmetrical sequences in transgenic Petunia hybrida. *The EMBO Journal* **13**: 2084–2088.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, *et al.* 2007. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics* **81**: 559–575.

Quinlan AR, Hall IM. **2010**. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841–842.

Raj S, Bräutigam K, Hamanishi ET, Wilkins O, Thomas BR, Schroeder W, Mansfield SD, Plant AL, Campbell MM. **2011**. Clone history shapes Populus drought responses. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 12521–12526.

Rapp RA, Wendel JF. **2005**. Epigenetics and plant evolution. *New Phytologist* **168**: 81–91.

Rasmussen KK, Kollmann J. **2007**. Low genetic diversity in small peripheral populations of a rare European tree (Sorbus torminalis) dominated by clonal reproduction. *Conservation Genetics 2007 9:6* **9**: 1533–1539.

Reinders J, Wulff BBH, Mirouze M, Marí-Ordóñez A, Dapp M, Rozhon W, Bucher E, Theiler G, Paszkowski J. **2009**. Compromised stability of DNA methylation and transposon immobilization in mosaic Arabidopsis epigenomes. *Genes & development* **23**: 939–950.

Revelle, W. 2022. psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych Version = 2.2.5.

Richards CL, Alonso C, Becker C, Bossdorf O, Bucher E, Colomé-Tatché M, Durka W, Engelhardt J, Gaspar B, Gogol-Döring A, *et al.* 2017. Ecological plant epigenetics: Evidence from model and non-model species, and the way forward. *Ecology Letters* **20**: 1576–1590.

Richards CL, Bossdorf O, Verhoeven KJF. **2010**. Understanding natural epigenetic variation. *New Phytologist* **187**: 562–564.

Sammarco I, Münzbergová Z, Latzel V. **2022**. DNA Methylation Can Mediate Local Adaptation and Response to Climate Change in the Clonal Plant *Fragaria vesca*: Evidence From a European-Scale Reciprocal Transplant Experiment. *Frontiers in Plant Science* **13**: 435.

Scaglione D, Pinosio S, Marroni F, Centa E di, Fornasiero A, Magris G, Scalabrin S, Cattonaro F, Taylor G, Morgante M. **2019**. Single primer enrichment technology as a tool for massive genotyping: a benchmark on black poplar and maize. *Annals of Botany* **124**: 543–551.

Schmitz RJ, He Y, Valdés-López O, Khan SM, Joshi T, Urich MA, Nery JR, Diers B, Xu D, Stacey G, *et al.* 2013. Epigenome-wide inheritance of cytosine methylation variants in a recombinant inbred population. *Genome research* **23**: 1663–74.

Schultz MD, Schmitz RJ, Ecker JR. **2012**. 'Leveling' the playing field for analyses of single-base resolution DNA methylomes. *Trends in Genetics* **28**: 583–585.

Slatkin M. **1993**. ISOLATION BY DISTANCE IN EQUILIBRIUM AND NON-EQUILIBRIUM POPULATIONS. *Evolution* **47**: 264–279.

Smulders MJM, Cottrell JE, Lefèvre F, van der Schoot J, Arens P, Vosman B, Tabbener HE, Grassi F, Fossati T, Castiglione S, *et al.* 2008. Structure of the genetic diversity in black poplar (Populus nigra L.) populations across European river systems: Consequences for conservation and restoration. *Forest Ecology and Management* **255**: 1388–1399.

Taylor G. **2002**. Populus: Arabidopsis for Forestry. Do We Need a Model Tree? *Annals of Botany* **90**: 681–689.

Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam M, Ralph S, Rombauts S, Salamov A, *et al.* 2006. The genome of black cottonwood, Populus trichocarpa (Torr. & Gray). *Science* **313**: 1596–1604.

Tuskan GA, Francis KE, Russ SL, Romme WH, Turner MG. **1996**. RAPD markers reveal diversity within and among clonal and seedling stands of aspen in Yellowstone National Park, U.S.A. *Canadian Journal of Forest Research* **26**: 2088–2098.

Verhoeven KJF, Preite V. **2014**. Epigenetic variation in asexxually reproducing organisms. *Evolution* **68**: 644–655.

Warnes G, Bolker B, Bonebakker L, Gentleman R, Huber W, Liaw A, Lumley T, Maechler M, Magnusson A, Moeller S, Schwartz M, Venables B. 2022. _gplots: Various R Programming Tools for Plotting Data_. R package version 3.1.3, https://CRAN.R-project.org/package=gplots.

Wibowo, A., Becker, C., Marconi, G., Durr, J., Price, J., Hagmann, J., Papareddy, R., Putra, H., Kageyama, J., Becker, J., Weigel, D., & Gutierrez-Marcos, J. (2016). Hyperosmotic stress memory in arabidopsis is mediated by distinct epigenetically labile sites in the genome and is restricted in the male germline by dna glycosylase activity. *ELife*, *5*(MAY2016). https://doi.org/10.7554/ELIFE.13546

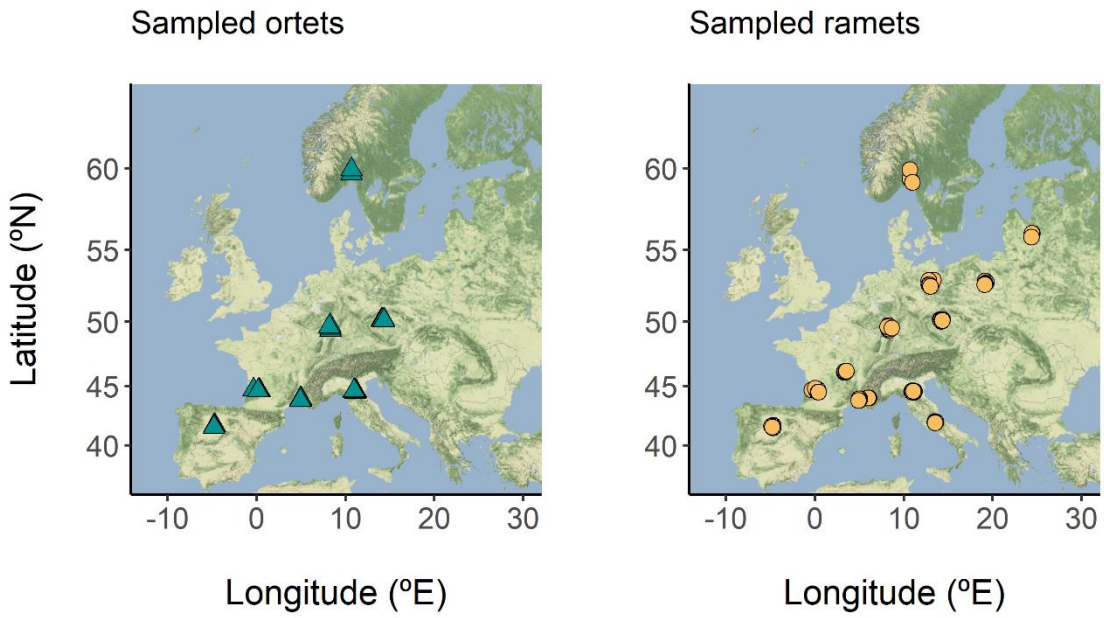
Zemach A, Kim MY, Hsieh P-H, Coleman-Derr D, Eshed-Williams L, Thao K, Harmer SL, Zilberman D. **2013**. The Arabidopsis Nucleosome Remodeler DDM1 Allows DNA Methyltransferases to Access H1-Containing Heterochromatin. *Cell* **153**: 193–205.

Zhang YY, Fischer M, Colot V, Bossdorf O. **2013**. Epigenetic variation creates potential for evolution of plant phenotypic plasticity. *New Phytologist* **197**: 314–322.

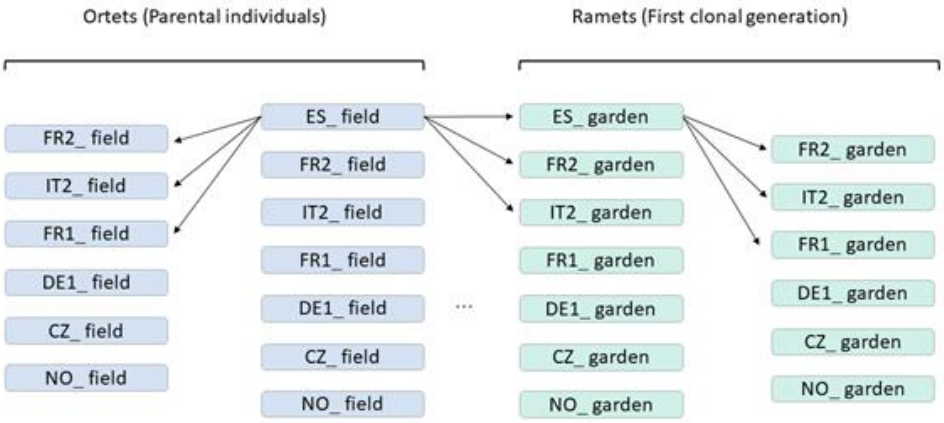
Zhang H, Lang Z, Zhu JK. **2018**. Dynamics and function of DNA methylation in plants. *Nature Reviews Molecular Cell Biology 2018 19:8* **19**: 489–506.

Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SWL, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE, *et al.* 2006. Genome-wide High-Resolution Mapping and Functional Analysis of DNA Methylation in Arabidopsis. *Cell* **126**: 1189–1201.

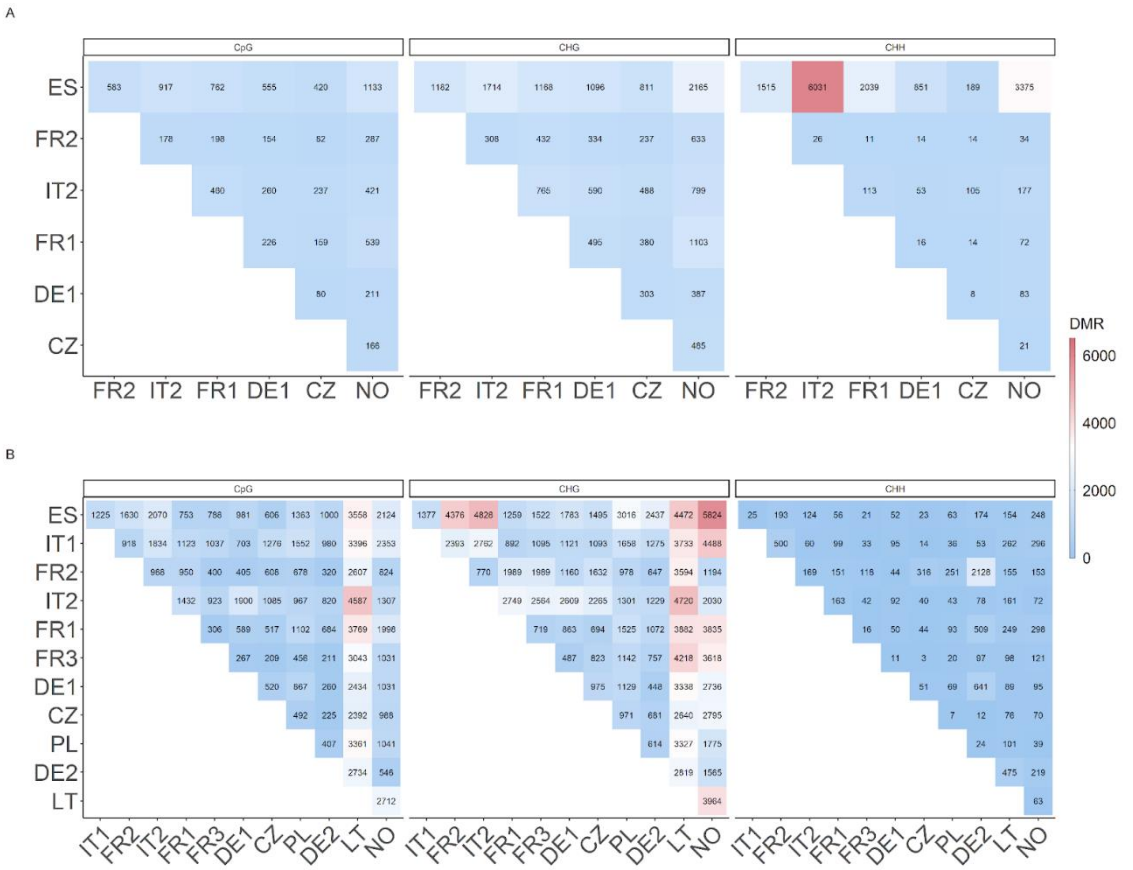
Supplementary Information



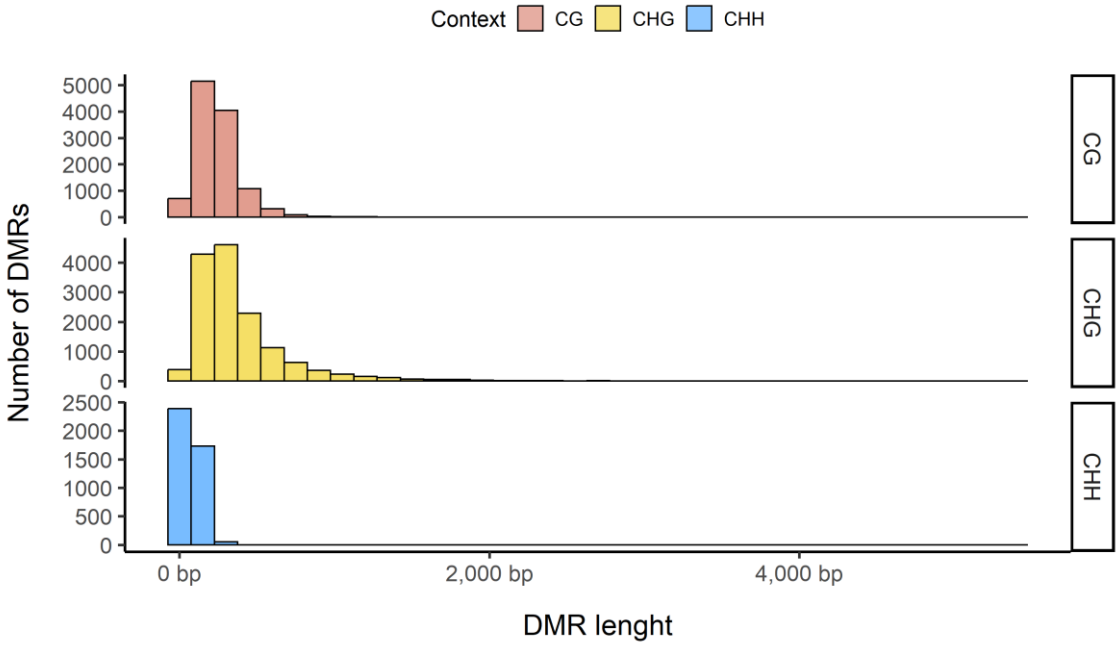
Supplementary figure 1. Sampling locations of ortets (green triangles, N = 35) and ramets (orange circles, N = 162).



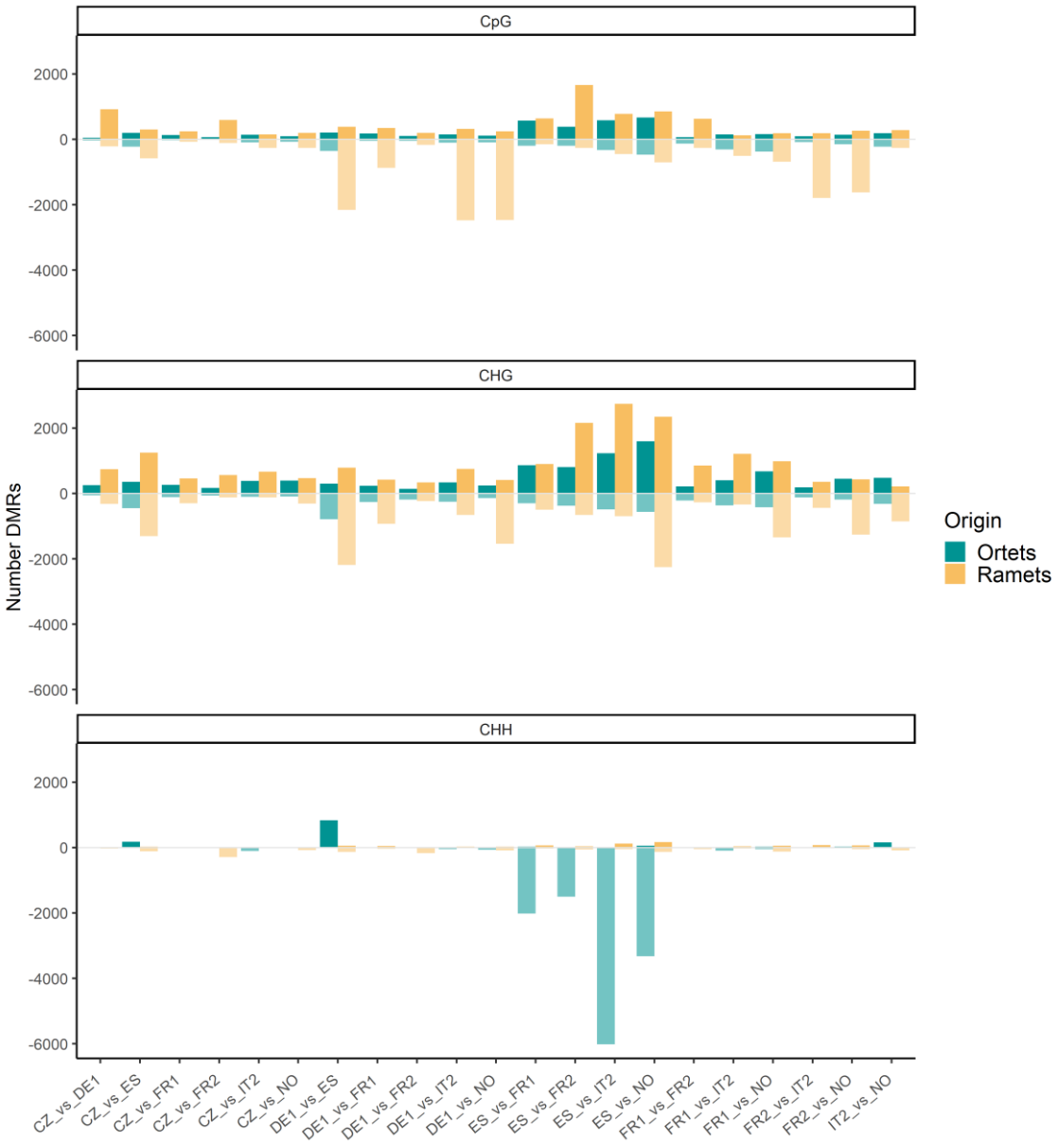
Supplementary Figure 2. Multiple pairwise comparison design for DMR calling groups. DMRs where called in three different instances. First, we run the pipeline using only groups containing ortet (field) samples in each pairwise comparison; second, we compared groups containing only ramet (garden) samples, and third, we compared ortet samples with their paired ramet samples (field vs garden).



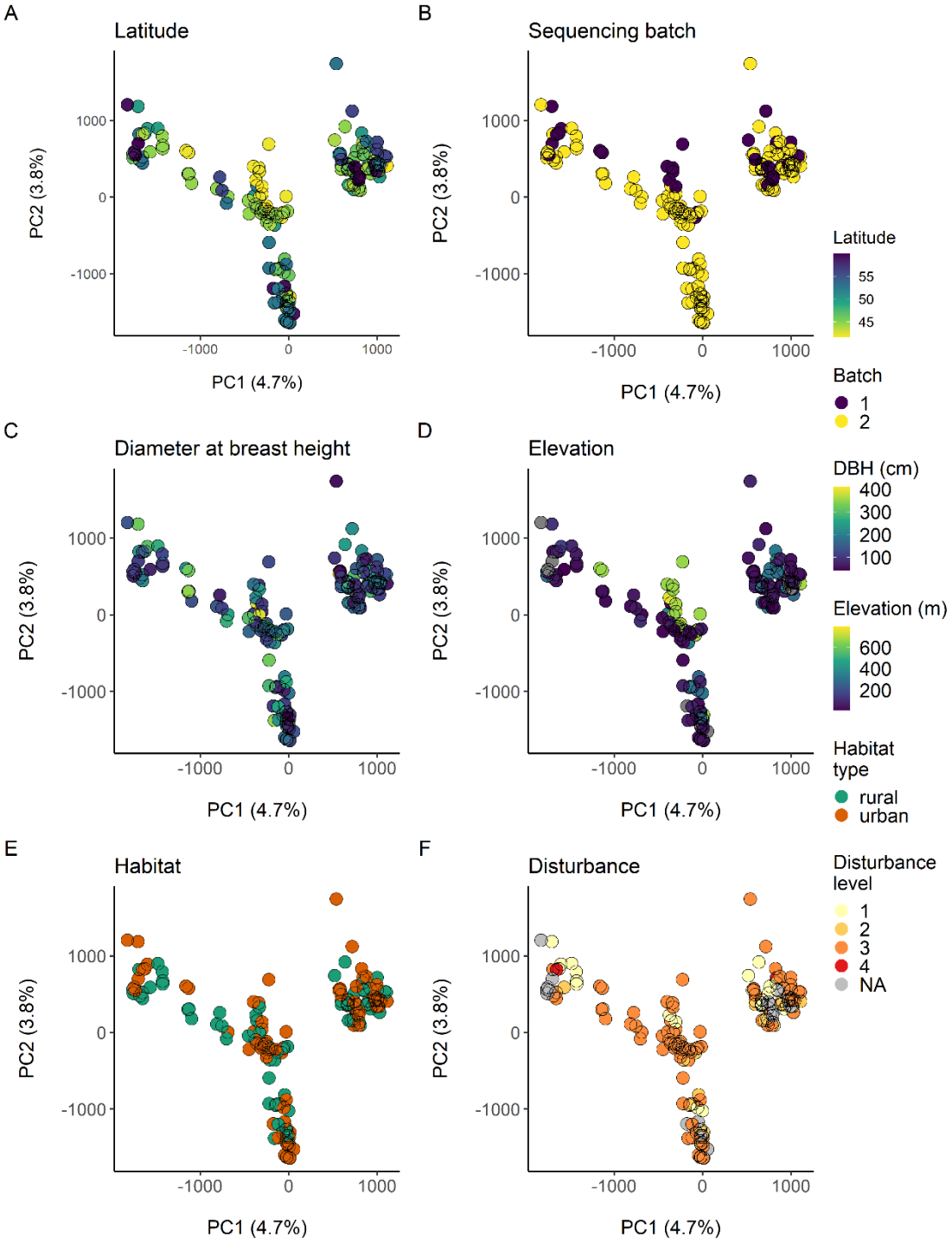
Supplementary Figure 3. Total number of DMRs between pairwise comparisons for A) ortets and B) ramets. Sampling sites are ordered from south to north, according to latitude.



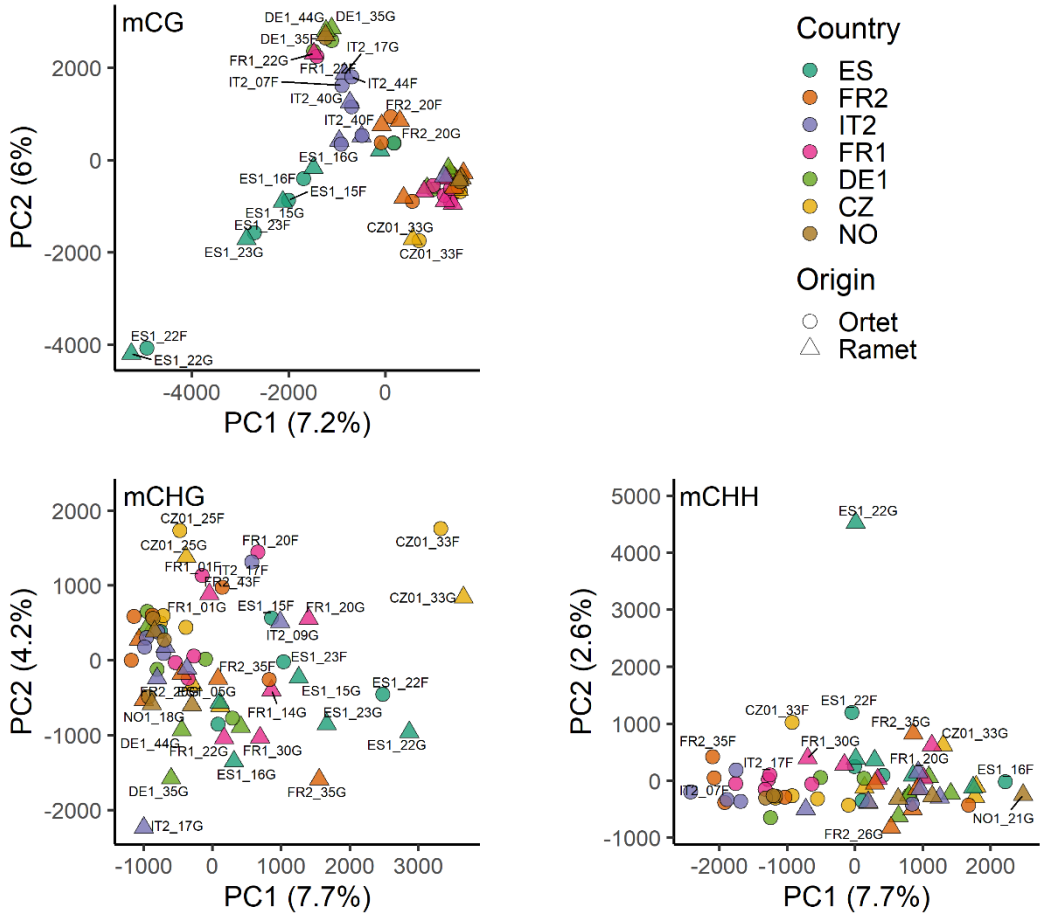
Supplementary figure 4. Histogram of the length of merged DMRs for each context.



Supplementary Figure 5. Total number of hyper (right side of the 0 line) and hypomethylated (left side of the 0 line) DMRs in all possible pairwise comparisons among paired ortets (dark green) and ramets (orange)



Supplementary Figure 6. Epigenetic structure in the Lombardy poplar using only methylated positions found inside coding sequences, colored according to different environmental variables.



Supplementary Figure 7. PCAs of epigenetic structure using methylated positions inside coding regions (CDS), comparing ortets and paired ramets, for CG, CHG and CHH contexts.

CHAPTER V

Epigenetic transgenerational inheritance

Iris Sammarco, Zuzana Münzbergová, Claude Becker, Oliver Bossdorf, Bárbara Díez Rodríguez, Dario Galanti, Adam Nunn, Vít Latzel

This chapter is in preparation for submission to *Nature Plants*

Abstract

Global warming is advancing at unprecedented rates, and whether plants will be able to survive future climatic conditions is still debated. It is thus crucial to better understand how plants adapt to rapid environmental changes. One mechanism allowing plants to quickly adapt to environmental changes includes epigenetic mechanisms, such as DNA methylation, which can alter phenotypes without changing the underlying DNA sequence. DNA methylation can in fact be under environmental control and can potentially regulate heritable phenotypic variation. The inheritance of DNA methylation variants seems to be particularly prominent across clonal generations, suggesting that DNA methylation variation might be crucial for the success and survival of clonal species, which present often low standing genetic variation. However, there are currently very few studies investigating the importance of genome-wide DNA methylation in the adaptation of wild clonal populations. To improve our understanding on the adaptive potential of DNA methylation, we need to study DNA methylation variation in plants grown in their natural environments, the extent of heritability of such variation across clonal generations and its function on gene expression, since this can ultimately alter phenotypes and thus plant fitness. To fill this knowledge gap, we studied the methylomes of plants from 21 natural populations of the clonal species *Fragaria vesca* (the wild strawberry). We selected the populations from three European countries and performed whole-genome bisulfite sequencing (WGBS) for plants grown in their natural habitats and clones of these plants of at least the third generation grown in common garden conditions (N = 84 per condition). We found that both field and garden individuals presented an epigenetic structure related to their geographic origin. Furthermore, the climate of origin of the populations was partly responsible for the epigenetic changes found between populations. The majority of these changes were inherited across at least three clonal generations, especially in the symmetric sequence contexts (CG and CHG). Finally, a subset of these epigenetic changes affected gene expression, suggesting that environmentally induced epigenetic variation can have a functional role. We conclude that DNA methylation variation in the wild is common and can aid adaptation of wild clonal plant populations.

Introduction

Given their sessile nature, plants must quickly adjust their phenotype to cope with changing environments *in situ*. One of the mechanisms allowing plants to phenotypically adapt to changing environments include epigenetic alterations of gene expression, for example via DNA methylation (Riggs and Porter 1996). In plants, DNA methylation can occur in three DNA sequence contexts: CG, CHG and CHH (where H is A, C or T) (Finnegan *et al.,* 1998), which have different respective functions (*e.g.* reviewed in: Niederhuth and Schmitz 2017). DNA methylation in all the sequence contexts represses transposon (TE) mobilization (reviewed in: Zemach and Zilberman 2010). Furthermore, CG methylation alone usually represses gene expression when present in gene promoters, while CHH methylation can occur close to active genes (X. Li *et al.,* 2012; Rajkumar *et al.,* 2020; Wang *et al.,* 2020). Importantly, DNA methylation patterns are generally mitotically and/or meiotically heritable, thus also affecting the offspring phenotype (reviewed in: Niederhuth and Schmitz 2014). Given that phenotype is the ultimate target of natural selection (Darwin and Wallace 1858), one might wonder whether epigenetic variation (especially the one induced by the environment) can alter the adaptive trajectories of natural plant populations (reviewed in: Jablonka and Raz 2009; Ashe, Colot, and Oldroyd 2021). Exploring the role of environmentally induced epigenetic variation in adaptation is a complex task especially under natural conditions, since epigenetic variation may have multiple sources, namely stochastic, genetic and environmental (Zhang *et al.,* 2013; 2018; Dubin *et al.,* 2015; Johannes and Schmitz 2019, Díez Rodriguez et al 2022, Galanti et al 2022). Genetically induced DNA methylation variants depend on genetic modifications (reviewed in: Richards 2006), which can be *cis*-acting (*e.g.* when a TE inserted upstream a gene promoter drives the methylation of the promoter itself) (*e.g.* Martin *et al.,* 2009), or *trans*-acting (*e.g.* when genetic mutations in genes involved in the DNA methylation machinery induce overall changes in DNA methylation patterns) (Dubin *et al.,* 2015; Baduel *et al.,* 2021). Instead, environmentally induced DNA methylation variants are under the sole control of environmental cues (reviewed in: Richards 2006), and can arise quickly in response to environmental stimuli (*e.g.* Zhang *et al.,* 2013; reviewed in: Thiebaut, Hemerly, and Ferreira 2019). Importantly, while genetically induced epigenetic variation reflects inheritance and selection of DNA sequence variations, environmentally induced epigenetic variation heritable across generations represents the *truly* adaptive role of epigenetic variation to changing environment. However, most studies on natural plant populations were not able or did not attempt to detect the source of epigenetic variation and thus its adaptive potential. The inability to disentangle between genetically and

environmentally related epigenetic variation can be ascribed either to low-resolution molecular methods that are not able to detect genome-wide methylation and/or to limitations due to the experimental design (such as very low number of populations) (*e.g.* Zoldoš *et al.,* 2018; Medrano *et al.,* 2020; Miryeganeh *et al.,* 2022). It is evident that the ability to discriminate particularly between genetically and environmentally induced heritable epigenetic variation is essential for improving our knowledge on the role of epigenetic variation in ecology and evolution of plants.

In particular, epigenetic variation might be especially important for the success and survival of clonal species, since it could compensate for their often low standing genetic variation (Latzel and Klimešová 2010; Verhoeven and Preite 2014; Dodd and Douhovnikoff 2016; Latzel, Rendina González, and Rosenthal 2016; Rendina González *et al.,* 2018; Münzbergová *et al.,* 2019; Shi *et al.,* 2019; Sammarco, Münzbergová, and Latzel 2022). Furthermore, in clonal species, epigenetic variation may play even higher role in heritable phenotypic variation than in plants reproducing only sexually. In fact, the inheritance of environmentally induced DNA methylation variation seems to be particularly prominent across clonal generations, which lack meiosis and the associated genome-wide epigenetic resetting leading to erasure of most environmentally induced epigenetic variation (reviewed in: Feng, Jacobsen, and Reik 2010, and Anastasiadi *et al.,* 2021). However, studies assessing the role of the environment in inducing heritable epigenetic variation in natural clonal plant populations are still scarce (*e.g.* C. L. Richards, Schrey, and Pigliucci 2012; De Kort *et al.,* 2020; 2022) (Díez Rodriguez et al 2022), although clonal propagation is the main reproductive mode in many ecosystems (Klimeš *et al.,* 1997).

To improve our understanding of the role of heritable, environmentally induced epigenetic variation in ecology and evolution of wild plant populations, we need to study DNA methylation variation in plants grown in their natural environments and the extent of heritability of such variation and its effects on gene expression. This epigenetic variation can in fact ultimately alter phenotypes and thus plant fitness.

To explore epigenetic variation and its heritability, we analyzed the methylomes of 231 plants from 21 natural European populations of a clonal species, the wild strawberry (*Fragaria vesca*). Specifically, we determined DNA methylation patterns of plant samples collected in environmentally different natural habitats in three European countries, and samples of their clonal offspring (ramets) of at least the third generation grown in a common garden. We also analyzed the functional role of inherited environmentally induced DNA methylation in gene expression. We asked the following questions:

1. Do environmental conditions such as temperature and precipitation induce DNA methylation variation in natural populations of *F. vesca*?
2. If so, is environmentally induced epigenetic variation inherited across clonal generations and does it modulate gene expression?

Answering these questions would provide evidence on whether environmental conditions can affect methylomes independent of genetic variation, and whether DNA methylation variation might affect gene expression, thus having an evolutionary potential.

Results

We collected 231 individuals of 21 natural populations of *F. vesca* following a temperature gradient in each of the three European countries. We performed Whole Genome Bisulfite Sequencing (WGBS) for leaf samples collected in the plants' natural conditions (hereafter referred to as *field conditions*) (total field N = 84), as well as for their clonally propagated offspring of at least the third generation (*i.e.* third order clonal ramet) grown for one year in a common garden (hereafter referred to as *garden conditions*). We also transplanted extra plants from the field to the garden conditions (total garden N = 147), but for these we did not collect samples from the parental generation grown in the field. We used all the available garden samples only for the GWA analysis (see later), while for the other analysis we used only the garden samples that had been clonally propagated from the field ones (*i.e.* for which we had samples from both field and garden conditions) (N = 84). From a subset of the garden individuals (N = 63), we also performed RNA-Sequencing, and for all the individuals we inferred Single Nucleotide Polymorphisms (SNPs) from the WGBS data.

(Epi)genetic variation

In order to assess whether the analyzed populations showed overall DNA methylation and genetic variation, we performed a Principal Component Analysis (PCA) for genetic variants (SNPs) and DNA methylation of plants from both field and garden conditions (Fig. 1). All the PCAs showed a clear clustering of the plants according to the country of their origin, irrespective of the growing condition. For DNA methylation, the PCAs clustered the individuals according to the country of origin of the plants in the CG and CHG contexts (Fig. 1b, 1c), while it showed an overlap among all the countries in the CHH context (Fig. 1d). We then performed Redundancy analysis (RDA) to assess the proportion of variance explained by country of origin and growing condition on genetic

and DNA methylation variation (Supplementary material Tab. S1). Country accounted for the highest proportion of variance in the CG context, followed by CHG, SNPs and CHH (respectively, 11.9%, 9.4%, 2.5% and 8.7% for CG, CHG, CHH and SNPs), while growing condition accounted for the highest proportion of variance in the CHG and CHH contexts, followed by CG and SNPs (respectively, 0.4%, 1.6%, 1.2% and 0.1% for CG, CHG, CHH and SNPs).

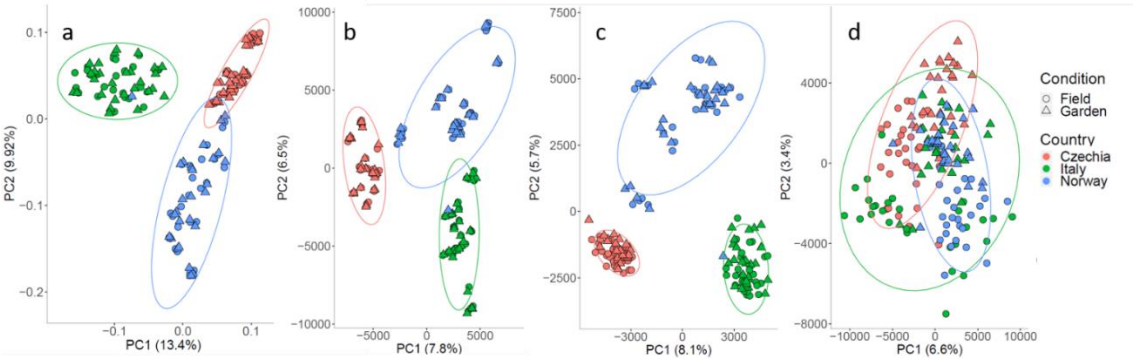


Figure 1: Principal Component Analysis (PCA) for genetic variants (SNPs) and DNA methylation. SNPs **(a)** and CG **(b)**, CHG **(c)** and CHH **(d)** methylated positions**.** Field (plants = 84), garden (plants = 84).

To assess the extent of methylation differences at the genomic-region level, we identified Differentially Methylated Regions (DMRs) separately for plants from all the populations from the field and from the garden conditions. By summing all the DMRs identified for each pairwise comparison, we identified more DMRs in the field conditions (CG = 2 536 619, CHG = 1 215 943, CHH = 6 738 040) than in the garden conditions (CG = 2 130 669, CHG = 1 019 222, CHH = 2 171 928), with the highest difference in CHH. Irrespective of the growing condition, the majority of CG-DMRs overlapped with gene bodies, while CHG- and CHH-DMRs with gene promoters and Transposable Elements (TEs) (Fig. 2). The number of DMRs in CHH-gene promoters and TEs was much higher in the field than in the garden conditions.

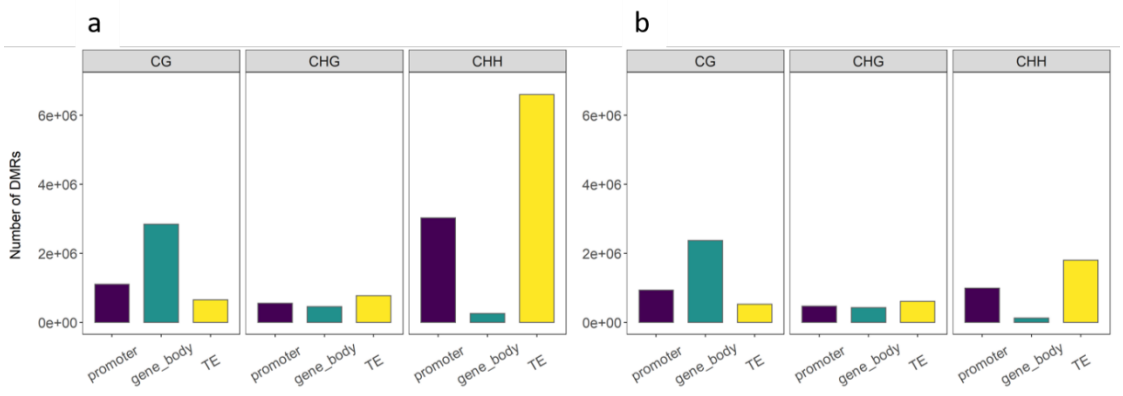


Figure 2: total number of Differentially Methylated Regions (DMRs) identified in the field and garden conditions, across different genomic features. (a) Field (plants = 84). **(b)** Garden (plants = 84).

Contribution of genetic and climatic variation to epigenetic variation: DMR variance decomposition analysis

In order to assess the relative contribution of genetic and climatic variation to epigenetic variation, we performed a DMR decomposition analysis using cis-genetic, trans-genetic and climatic data as predictors. For each DMR, we analyzed three independent mixed models including a distance matrix derived from cis-genetic variants, trans-genetic variants or climatic distances as predictors, as in Galanti (2022). Since each model was independent of the others, we assigned each DMR to the predictor explaining the highest DMR variance (hereafter referred to as *strongest predictor*).

We found that for all the sequence contexts and for both field and garden conditions, the highest proportion of DMRs was assigned to trans-genetic variation, followed by climatic variation and cis-genetic variation (Fig. 3). The amount of trans-predicted DMRs gradually decreased from CG to CHG and CHH, while the amount of climate-predicted DMRs gradually increased from CG to CHG and CHH. The only exception was the CHH context in garden conditions, which showed less climate-predicted DMRs than the CHG context, but many more unexplained DMRs (DMRs where all the three predictors failed to explain >10% of the variance) (as in Galanti *et al.,* 2022). Since many cis-, trans- and climate-predicted DMRs overlapped with genic regions (Supplementary material, Fig. S1), we then performed a Gene Ontology (GO) enrichment analysis to functionally characterize the DMR-related gene promoters assigned to the three different predictors. We performed the GO analysis separately for cis-, trans-, climate- and unexplained-predicted DMRs. We found no enriched GO terms for CG, while we found several for CHG and CHH. For cis-predicted DMR-related gene promoters, we found enrichment for

several GO terms, including two Cellular Component (CC) terms related to chromatin and chromatin remodeling activity (nucleosome; Ino8 complex), and several Biological Process (BP) terms (telomere maintenance; protein metabolic process; phloem development; glucose metabolic process) (Fig. 4). For both trans-predicted and climate-predicted DMR-related gene promoters, we found enrichment for few terms, including some common between the two (RNA-DNA hybrid ribonuclease activity (MF); Retrotransposon nucleocapsid (CC)). We found Retrotransposon nucleocapsid (CC) also for the unexplained-predicted DMR-related gene promoters.

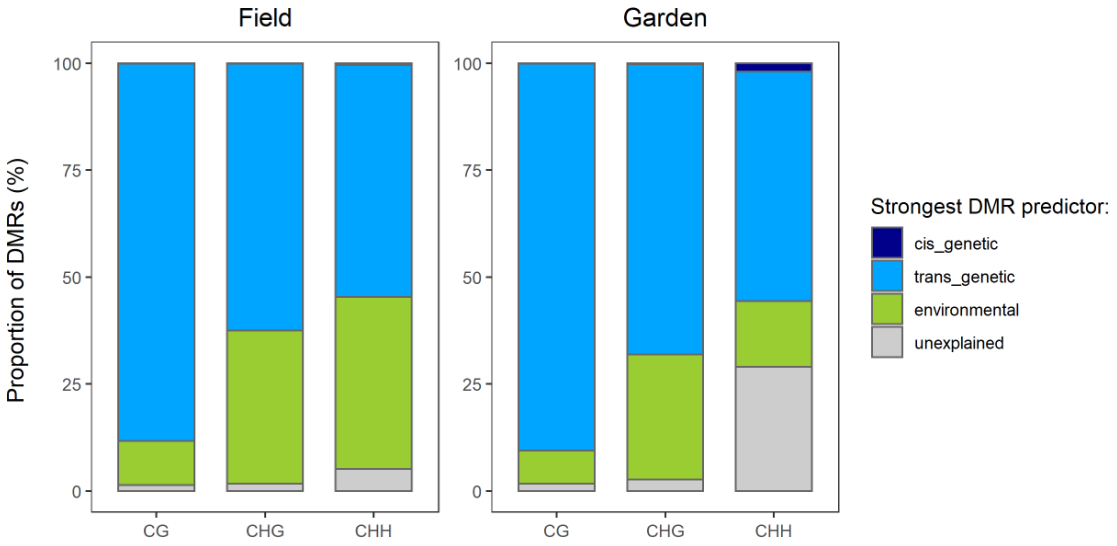


Figure 3: Strongest DMR predictor. Proportion of DMRs with strongest predictor assigned to cis-genetic, trans-genetic or climatic variation. DMRs where all the three predictors failed to explain >10% of the variance are classified as “unexplained”. **(a)** Field, total DMRs: 82 546 CG, 49 459 CHG and 211 363 CHH (plants = 84). **(b)** Garden, total DMRs: 71 856 CG, 37 795 CHG and 138 807 CHH (plants = 84).

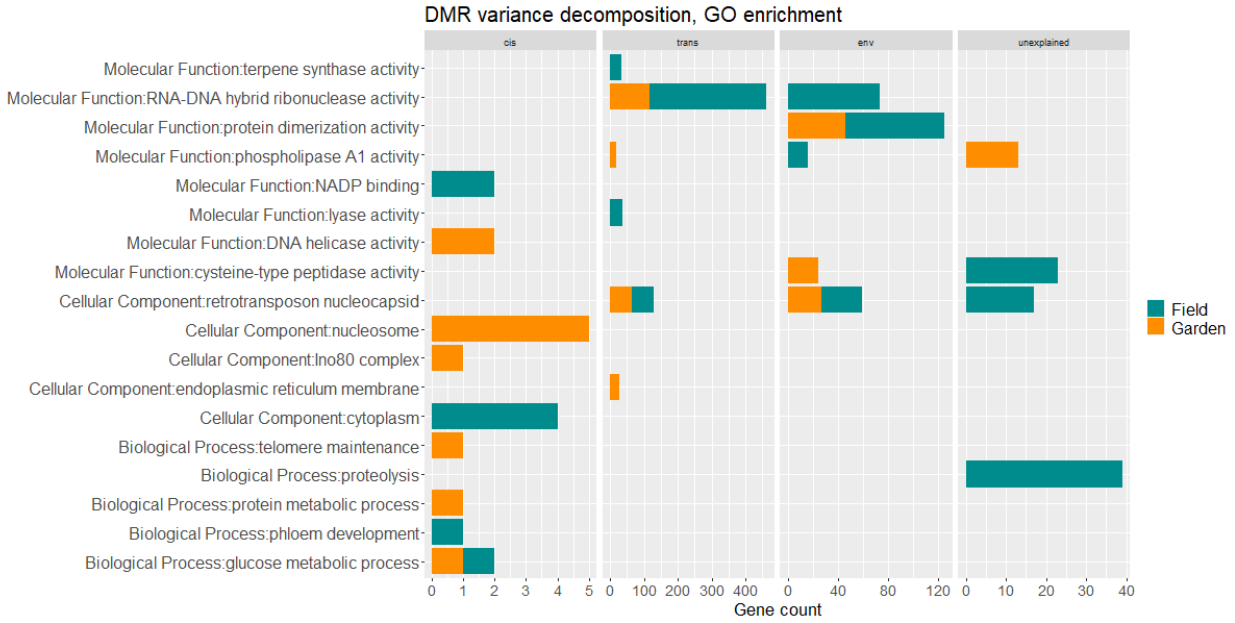


Figure 4: Gene Ontology (GO) enrichment analysis for cis-, trans-, climate- and unexplained-predicted DMRs in CHG and CHH (no enrichment for CG). Gene count represents the number of genes assigned to each GO category. Only GO terms with an adjusted P value < 0.05 are shown.

Correlation of climate-predicted DMRs with gene expression

We envisioned that the climate-predicted DMRs could be either under *direct* or *indirect* control of the climate of origin of the populations. In the first case, the climate-predicted DMRs would depend only on environmental factors, and they would be classified as *directly* environmentally induced DMRs. In the second case, the climate-predicted DMRs would be under the control of genetic variants selected by environmental factors, and they would be thus classified as *indirectly* environmentally induced DMRs. In this study, since we were testing for the adaptive role of epigenetic variation, we were mainly interested in the direct effect of environmental factors on DMRs. Since the DMR variance decomposition analysis was not able to distinguish among the direct or indirect effect of climate on DNA methylation, we performed a genome-wide association (GWA) analysis to assess the proportion of climate-predicted DMRs showing a genetic basis, in order to exclude them from further analysis. For the GWA analysis, we used only the samples grown in garden conditions to ensure that the selected climate-predicted DMRs were heritable and thus of evolutionary potential. To increase the statistical power of the analysis, we included all the samples available from the garden (147 total plants). Out of 2 439 CG-, 3 448 CHG- and 9 508 CHH-climate-predicted DMRs overlapping gene promoters, we performed GWA and correlation analysis (see later) for a random set of

100 DMRs per context, using individual DMR methylation as a phenotype (and SNPs as predictor). We repeated the same analysis three times to assess the variation among analyses and thus whether the amount of DMRs chosen as significantly related to climate was representative for the whole dataset. By averaging the three independent analyses, we found a significant GWA hit for 36.3% (± 3.5, SD) of the climate-predicted DMRs in CG, and 22% in both CHG and CHH (respectively, ± 2.6 and 3.2, SD). We thus classified these DMRs as *indirectly* environmentally induced, as opposed to those *directly* induced by the environment for which we found no significant hit in the GWA analysis (mean ± SD, CG = 63.7% ± 3.5, CHG = 79.0% ± 2.6, CHH = 81.7% ± 3.2).

In order to assess whether the *directly* environmentally induced DMRs had a functional role, we then tested whether the methylation level of each individual DMR was correlated with the expression of the overlapping gene. We found a statistically significant correlation for 9.3% of the cases in CG (± 3.2, SD), 5.7% in CHG (± 1.5, SD) and 7.7% in CHH (± 1.5, SD). In particular, for both CG and CHH, we found slightly more genes with a positive correlation between promoter methylation and gene expression than with negative correlation (mean ± SD, positive correlation: CG = 5.2% ± 1.0, CHH = 4.6% ± 1.7; negative: CG = 4.2% ± 2.9, CHH = 3.3% ± 0.4). For CHG we found more genes with a negative correlation than with a positive one (mean ± SD, positive correlation: 1.8% ± 0.6; negative: 4.2% ± 2.1) (Fig. 5).

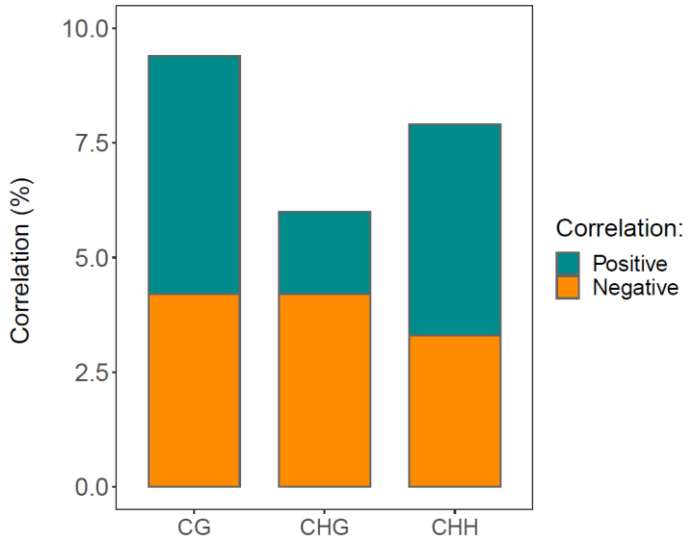


Figure 5: percentage of genes showing significant positive or negative correlation between DMR-promoter methylation and gene expression. Plants = 63.

Discussion

Growing evidence suggests that environmental variation can induce DNA methylation variation that may be inherited across generations (Raj *et al.,* 2011; Thiebaut, Hemerly, and Ferreira 2019). Such environmentally induced DNA methylation variation might be affecting the phenotypes of the offspring generations and therefore also evolutionary trajectories of plant populations (Jablonka and Raz 2009; Ashe, Colot, and Oldroyd 2021). It has also been proposed that the inheritance of environmentally induced DNA methylation variation might be particularly prominent across clonal generations, due to the lack of meiosis in soma-clonal reproduction (Feng, Jacobsen, and Reik 2010); (Latzel and Klimešová 2010; Simonetti *et al.,* 2019); (Dodd and Douhovnikoff 2016; Latzel, Rendina González, and Rosenthal 2016; Münzbergová *et al.,* 2019; Shi *et al.,* 2019). However, clear evidence is still scarce, particularly from natural conditions. In this study, we tested whether variation in environmental conditions can trigger DNA methylation variation in natural plant populations of the clonal herb *Fragaria vesca*, whether such DNA methylation variation is stable across clonal generations and whether it alters gene expression. Specifically, we studied (epi)genetic variation of 21 natural populations of the wild strawberry in their natural habitats and of their clonal offspring propagated in a common garden.

DNA methylation variation of wild populations of F. vesca is strongly genotype dependent but reflects also natural climatic conditions (Q1)

The populations analyzed presented both a significant genetic and epigenetic geographic structure (Fig. 1). In the CG and CHG contexts, country of origin accounted for higher percentage of variance than for genetic variants (respectively, 11.9%, 9.4% and 8.7% for CG, CHG and SNPs; Supplementary material Tab. S1), suggesting that part of epigenetic variation was not under genetic control. Furthermore, in addition to the differentially methylated regions (DMRs) related to trans- and cis-genetic variation, we identified also DMRs directly related to climatic variation (Fig. 3a). Interestingly, climate-predicted DMRs gradually increased from CG to CHG and CHH context (respectively, 10.36%, 35.86% and 40.16%), suggesting that non-CG methylation was particularly sensitive to climatic conditions.

We then assessed the putative functional role of DMR-related gene promoters assigned to cis-, trans-, climate or unexplained variation (Fig. 4). While we did not identify enriched GO terms for CG, we found enrichment for several GO terms in CHG and CHH. Cis-genetic variants induced mainly DNA methylation variants in genes related to chromatin

and chromatin remodeling activity, telomere maintenance and metabolic processes, trans-genetic and climatic variation altered methylation in genes related to RNA-DNA hybrid ribonuclease activity and retrotransposon nucleocapsid. Interestingly, RNA-DNA hybrids and retrotransposon nucleocapsid can be related to retrotransposon mobilization (Todd *et al.,* 2020), suggesting that both trans-genetic and climatic variation modulate transposition and that the environment might control TE mobilization in wild conditions (Rey *et al.,* 2016; Baduel *et al.,* 2021).

We found some overlapping GO enriched terms between trans-predicted and climate-predicted DMR-related gene promoters. This evidence thus suggests that both trans-genetic and climatic variation can affect the methylation status of a similar set of genes. It is however important to note that only 54% of genes present a GO annotation in *F. vesca* (Li *et al.,* 2019), suggesting that some important gene functions may remain hidden to us.

Environmentally induced DNA methylation variation is stable across clonal generations and might directly affect gene expression (Q2)

The comparison of epigenetic patterns of plants from the field with their clonal offspring propagated in a common garden allowed us to assess the rate of inheritance of environmentally induced DNA methylation during clonal reproduction and its role in gene regulation. The epigenetic structure was significantly different between field and garden conditions only in the CHG and CHH contexts (RDA with growing condition as predictor: $P < 0.001$; Supplementary material Tab. S1), suggesting that methylation in these contexts might be particularly sensitive to environmental variation. However, we identified a similar amount of gene- and TE-related DMRs in the field and garden conditions in both the CG and CHG contexts (percentage difference %, CG: promoters 16.8, gene bodies 18.2, TEs 21.8; CHG: promoters 15.8, gene bodies 6.9, TEs 23.4) (Fig. 2), suggesting that the majority of methylation found in these contexts on functional regions was highly heritable and likely genotype dependent. The biggest difference in DMRs between plants from the field and garden conditions was in the CHH context, which showed a much higher number of gene promoters- and TEs-DMRs in the field than in the garden (percentage difference %, promoters 101.5, gene bodies 70.1, TEs 114.1). CHH methylation in plants plays crucial role in TE silencing, in the establishment of heterochromatin (Slotkin and Martienssen 2007; Fultz, Choudury, and Slotkin 2015), and recent in gene regulation in euchromatic regions (Zemach *et al.,* 2010; Gent *et al.,* 2013; Martin *et al.,* 2021). The high number of gene promoters- and TEs-DMRs found in

the field in the CHH context thus suggests that natural environmental conditions likely play a role in regulating gene expression and TE mobilization, which is in accordance with our previous finding that the environment might control TE mobilization in wild conditions (see above). This process might have a strong evolutionary relevance, since TE mobilization can have profound consequences for the evolution of plant genomes and thus speciation (Schmidt and Anderson 2006; Oliver and Greene 2009).

In order to distinguish whether the inherited DNA methylation variation was related to environmental induction or genetic variation, we performed the DMR variance decomposition analysis also for the plants grown in common garden conditions (Fig. 3b). Our assumption was that the climate-predicted DMRs found in common environment conditions could be considered with high fidelity to be transgenerationally inherited, since they correspond to the climate of the original field conditions. Interestingly, we found a similar amount of climate-predicted DMRs between field and garden, especially in the CG and CHG contexts. In contrast, the CHH context in garden conditions showed a great increase in unexplained variation at the expense of climatic variation. In agreement with what discussed above for the epigenetic clustering and number of DMRs, we speculate that CHH methylation is the least stable across clonal generations and/or is the most responsive to short-term environmental changes. We therefore hypothesize that the unexplained DMRs found in CHH might be due to the environmental conditions attributable to the common garden. On the other hand, the similar CG- and CHG-DMR variation for both field and garden conditions due to climatic variation suggests that the climate of origin of the populations induced DMRs heritable across clonal generations in these sequence contexts.

A similar extent of climate-predicted DMRs was observed in natural accessions of the field pennycress *Thlaspi arvense* grown in greenhouse conditions (no field data are available for this study) (Galanti *et al.,* 2022), which provide some evidence for generalization in non-model plant species. The authors found that the contribution of climatic variation to DMR variation gradually increases from CG to CHG and CHH, and that among all the sequence contexts CHH presents the highest amount of unexplained variance. Like our results, they also found that trans-genetic variation explains the highest DMR variation, but that cis-genetic variation explains higher DMR variation than we found in our study (~5-14% in *T. arvense*, 0.2-2% in *F. vesca*). This might be due to the higher standing genetic variation of sexually reproducing *T. arvense* (Frels *et al.,* 2019) in comparison to mostly clonally reproducing *F. vesca* in natural conditions (Schulze *et al.,* 2012). Similar findings showing a high contribution of genetic variation to

DNA methylation variation were also reported in studies on *F. vesca* by De Kort and her colleagues (De Kort *et al.,* 2020, 2022).

In the next step, the GWA analysis revealed that the majority of climate-predicted DMRs was not linked to the genetic data and could thus be expected to be *directly* induced by the environment. Our findings are in contrast with those observed in natural *Arabidopsis thaliana* accessions grown under controlled conditions at two different temperatures (Dubin *et al.,* 2015). The authors found a strong association of CHH climate-predicted DMRs with both cis- and trans-genetic variants, thus suggesting that the climate-predicted DMRs were actually induced by the environment *indirectly* via selection of particular genetic patterns. We speculate that the contrasting results might be due to species-specific characteristics, such as different lifestyles and reproductive modes (perennial clonal vs annual sexual reproductive strategy) (Warwick *et al.,* 2011; Schulze *et al.,* 2012) and/or differences in the experimental design.

Finally, part of the climate-predicted DMRs identified in the uniform garden conditions (*i.e.* inherited) were significantly correlated with gene expression. Although the correlation was significant only for a subset of the inherited DMRs, this is still an intriguing result considering that many other studies have failed to find any significant correlation between promoter methylation and gene expression (Ganguly *et al.,* 2017). Interestingly, the correlation analysis revealed both positive and negative roles of promoter methylation on gene expression. Although promoter methylation is usually negatively correlated with gene expression (X. Li *et al.,* 2012), some studies report also a positive effect of promoter methylation on gene expression (Lang *et al.,* 2017), especially in the CHH context (Gent *et al.,* 2013; Xu *et al.,* 2018; Wang *et al.,* 2020; Rajkumar *et al.,* 2020). Therefore, our detection of a positive correlation between promoter methylation and gene expression in CG and CHG context is a rather surprising and unique novel finding. We provide two possible explanations of the unexpected result. First, by testing the correlation between DNA methylation and expression of the same gene across 63 samples, we captured more variation than studies tested such a correlation among different genes in the same individual (Wang *et al.,* 2020). Second, we performed this analysis only for climate-predicted DMRs, which might have different effects on gene expression than other DMRs, since they might behave differently than other DMRs (*e.g.* by being directly induced by environmental conditions, these DMRs could have specific effects on gene expression). However, such a speculation remains to be tested by further studies.

Conclusion

To conclude, the results of our study suggest that inherited-environmentally induced DNA methylation variation can have an impact on gene expression, and thus ultimately on phenotype, and can thus play an adaptive role in wild *F. vesca* populations. It is worth noting that the heritable epigenetic variation found in our study may also be due to natural selection acting on stochastic changes in DNA methylation (*i.e.* spontaneous epimutations) (Richards *et al.,* 2017b). However, we could not assess the rate of spontaneous epimutations in our dataset. That being said, few other studies also found an association between environmentally induced epigenetic changes and gene expression (Secco *et al.,* 2015; Wibowo *et al.,* 2016). These studies, however, were performed under conditions, which limit their generalization on the adaptive potential of epigenetic changes in natural environments. By working with natural plant populations, the epigenetic changes identified in our study have high ecological and evolutionary relevance. Moreover, by experimental alteration of DNA methylation in a complementary study using a subset of the same populations, we showed that DNA methylation can play a role in local adaptation, thus suggesting that the changes in gene expression found in this study might actually affect plant fitness (Sammarco *et al.,* 2022). Our findings thus suggest that environmentally induced heritable epigenetic variation might affect phenotypic variation on which natural selection might act, and might thus be ecologically and evolutionarily relevant (Bossdorf, Richards, and Pigliucci 2007). Inheritance of epigenetic variation might be even more relevant for clonal species, in which they might create an additional layer of variation that might compensate the low-standing genetic variation often found in these species, thus contributing to the ecological success of clonal species (Latzel & Klimešová, 2010b; Verhoeven & Preite, 2014b; Dodd & Douhovnikoff, 2016b).

Materials and methods

Study Species

Fragaria vesca L., Rosaceae, is an herbaceous perennial species with wide geographic distribution (Europe, northern Asia, North America, and northern Africa) (Darrow, 1966). It reproduces both clonally through stolons and sexually through seeds, but its sexual reproduction is very rare in natural conditions (Schulze *et al.,* 2012).

Plant collection and growth

Between May and July 2018, we selected 21 natural populations of *F. vesca* from three European countries, Italy, Czechia and Norway (see Supplementary material Tab. S2 for geographic locations and climatic characteristics of the selected populations). We chose these countries as they represented the southern limit (Italy), the core (Czechia) and the northern limit (Norway) of the native range of *F. vesca* distribution in Europe. To increase the environmental difference among the populations' sites, we sampled the populations following a climatic (mostly corresponding to altitudinal) gradient within each country.

For each population, we collected mature, fully developed leaf of 4 individuals directly from the field conditions (N = 84), we dried them in silica gel and used them for Whole Genome Bisulfite Sequencing (WGBS) analysis, see later. We then dug up the same ramets plus additional three (N = 147) and planted them individually following a random block design in 70 × 40 × 20 cm pots filled with a commercial mixture of compost and sand located in the common garden of the Institute of Botany of the Czech Academy of Sciences in Průhonice, Czechia (49.994°N, 14.566°E) one to ten days after their collection (see Tab. S2 for the climatic characteristics of the common garden). Plants were grown under a shading coverage reducing 50% of the light to simulate natural light levels at most of the localities. We let the plants propagate clonally for one year. Then, we selected the biggest offspring ramet of at least the third generation from every clone and we collected mature, fully developed leaf samples and froze them immediately in liquid nitrogen. These samples were later used for WGBS (N = 147). From a subset of 3 plants per population (N = 63), we also collected mature leaf samples for RNA-Sequencing in the same way as samples for WGBS.

WGBS library preparation and sequencing

We extracted genomic DNA from individual plants using the Qiagen DNeasy Plant Mini Kit, following the manufacturer's instructions with minor modifications. To improve DNA quality and yield from *F. vesca*, a known recalcitrant species, we used an increased amount of buffer AP1 (600 µl) together with 100 µl of EDTA (0.5 M, pH=8) and PVPP (polyvinilpolypyrrolidone), and an increased amount of buffer P3 (260 µl).

We then prepared libraries for WGBS using the NEBNext® Ultra™ II DNA Library Prep Kit and EZ-96 DNA Methylation-Gold™ MagPrep (ZYMO). Briefly, we sonicated 200-300 ng of genomic DNA to a mean fragment size of ~350 bp using the Covaris instrument. We then performed end repair and 3' adenylation of sonicated DNA fragments, ligated the NEBNext adaptors, performed size selection with AMPure XP Beads (Beckman Coulter, Brea, CA), we treated the DNA with bisulfite, we performed PCR enrichment and index ligation using Kapa HiFi Hot Start Uracil+ Ready Mix (Agilent) (14 cycles). Finally, we sequenced paired-end reads on HiSeq X Ten (Illumina, San Diego, CA), using a sequencing coverage per sample of 30x.

All the sequencing data (WGBS and RNA-Sequencing) can be found in the European Nucleotide Archive (ENA, www.ebi.ac.uk/ena/), under the project PRJEB51609.

Methylation and DMR calling

We used the EpiDiverse WGBS pipeline for bisulfite reads mapping and methylation calling (https://github.com/EpiDiverse/wgbs), which was specifically designed for non-model plant species (Nunn *et al.,* 2021). The pipeline performed quality control (FastQC), base quality and adaptor trimming (cutadapt), bisulfite mapping (erne-bs5) and bisulfite non-conversion rate, duplicates detection (Picard MarkDuplicates), alignment statistics and methylation calling (Methyldackel). In the mapping step, we used the most recent version of the genome of *F. vesca* v4.0.a2 (Edger *et al.,* 2018; Li *et al.,* 2019). In average, the sequencing produced 97 142 389 reads per sample (see Supplementary material Tab. S3 for detailed information), of which 96% mapped successfully to the genome after retaining only uniquely-mapping reads. We calculated the bisulfite non-conversion rate using the chloroplast genome, which is naturally unmethylated (Fojtová *et al.,* 2001), and we found an average non-bisulfite conversion rate among samples of 0.5% (see Supplementary material Tab. S3). We obtained individual bedGraph files of methylated positions for each sample and sequence context.

For PCA and RDA analyses, we then combined the individual bedGraph files from both field and garden conditions in a multisample unionbed file using custom scripts and bedtools (Quinlan & Hall, 2010). In order to compare the field with the garden conditions, we used only the samples for which we had WGBS data for both conditions (N = 84 per condition). We retained all the cytosines having coverage ≥ 5 in at least 80% of the samples (total methylated positions: 1 644 729, 2 574 494 and 12 335 916, respectively for CG, CHG and CHH). We then performed PCA with custom scripts using the R function prcomp (Sigg & Buhmann, 2008), and RDA with the R function vegan (Oksanen Jari *et al.,* 2020). For DNA methylation, we performed two separate RDA analyses, using in all of them Hellinger-transformed methylated positions as independent variable and in 1) country of origin as predictor and growing condition (field, garden) as a covariate to account for the effect of growing condition on the plants' methylomes; 2) growing condition as predictor and country as a covariate. We tested the statistical significance of the RDA analyses with one-way analysis of variance (ANOVA).

We called DMRs using the EpiDiverse DMR pipeline (https://github.com/EpiDiverse/dmr) (Nunn *et al.,* 2021) and using the DMR caller metilene (Jühling *et al.,* 2016). We used populations as groups, and we called DMRs separately for all the pairwise comparisons between the populations from the field, and the populations from the garden (*i.e.* we never compared a field population with a garden population). We used as input individual bedGraph files filtered for cytosine coverage ≥ 5. Separately for field and garden conditions, we then combined the output bed files ($p < 0.05$) in a multisample bed file using custom scripts and bedtools (v2.27.1) (Quinlan & Hall, 2010). To assess the number of DMRs overlapping with gene promoters, gene bodies and TEs, we then intersected the multisample bed files with gene and TE annotations. For genes, we used the gene annotations v4.0.a2 downloaded from the Genome Database for Rosaceae (GDR) (https://www.rosaceae.org/species/fragaria_vesca/genome_v4.0.a2) (Jung *et al.,* 2019), while for TEs we used an annotation carried out using the EDTA annotation pipeline v1.9.6 (Ou *et al.,* 2019) on the substituted genome using default parameters, kindly provided by (López *et al.,* 2022).

SNP calling

We inferred Single Nucleotide Polymorphisms (SNPs) from WGBS data using the EpiDiverse SNP pipeline with default parameters (https://github.com/epidiverse/snp) ((Nunn *et al.,* 2021; 2022). For the DMR variance decomposition analysis, separately for field and garden conditions, we then combined the output individual VCF files into

multisample VCF files using BCFtools (v1.9) (Danecek *et al.,* 2021). As above, we used only the samples for which we had WGBS data for both conditions (N = 84 per condition). Using VCFtools (v0.1.16) (Danecek *et al.,* 2021), we filtered the variants successfully genotyped in 80% of individuals, with a minimum quality score of 30 and a minimum mean depth of 3.

For PCA and RDA analyses, we combined the individual VCF files from both field and garden (N = 84 per condition) into a multisample VCF file and performed the same filtering steps as above. We also filtered for Minor Allele Frequency (MAF) ≥ 0.05 and ≤ 0.95, and pruned for Linkage Disequilibrium (LD) with an LD threshold (r^2) of 0.2 for SNP pairs in a sliding window of 50 SNPs, sliding by 5. After filtering, we were able to retain 76 669 SNPs. We then calculated the PCA eigenvalues and eigenvectors with PLINK (v1.90b6.12) (http://pngu.mgh.harvard.edu/purcell/plink/) ((Purcell *et al.,* 2007) and we plotted the PCAs with custom scripts in R. We performed RDA analysis similar to methylation (see above), but using Hellinger-transformed SNPs as independent variable.

DMR variance decomposition analysis and GO enrichment

To assess the amount of methylation variance explained by cis-variants, trans-variants and climatic variation, we ran three mixed models for each individual DMR for both field and garden conditions, as in Galanti *et al.,* (2022). Briefly, for cis-variants, we used an IBS matrix generated with PLINK (v1.90b6.12) (http://pngu.mgh.harvard.edu/purcell/plink/) ((Purcell *et al.,* 2007) using variants within 50kb from the DMR middle point. We retrieved 103 839 and 95 410 SNPs for field and garden conditions, respectively. For trans-variants, we used an IBS matrix obtained from variants filtered for Minor Allele Frequency (MAF) ≥ 0.01 and pruned for Linkage Disequilibrium (LD) with an LD threshold (r^2) of 0.2 for SNP pairs in a sliding window of 50 SNPs, sliding by 5. For climatic variation, for both field and garden conditions, we calculated a Euclidean distance matrix between climatic data from all the field sites, which we reversed and normalized to obtain a similarity matrix in a 0 to 1 range. The climatic data included mean, maximum and minimum temperature, and precipitation, all averaged over 7 years before the sampling year (2011-2018). We sourced climatic data at a horizontal resolution of 0.1 × 0.1° (v20.0e) from the European gridded dataset E-OBS, available through the C3S Climate Data Store (CDS) website (https://cds.climate.copernicus.eu/cdsapp#!/home) ((Cornes *et al.,* 2018). For methylation variants, we merged the DMRs obtained from all the pairwise comparisons with bedtools, separately for field and garden. As above, we used only the samples for

which we had WGBS data for both conditions (N = 84 per condition). We obtained 82 546 CG-DMRs, 49 459 CHG-DMRs and 211 363 CHH-DMRs for field, and 71 856 CG-DMRs, 37 795 CHG-DMRs and 138 807 CHH-DMRs for garden. We extracted average methylation of the resulting DMRs from all the samples with the function *regionCounts* from the R package methylKit (v1.16.1) (Akalin *et al.,* 2012), using a minimum cytosine coverage of 5.

We then ran a GO enrichment analysis for cis-, trans-, climate- and unexplained-predicted DMRs, separately for each sequence context and for field and garden conditions. We extracted DMR-related gene promoters with bedtools, and performed a GO enrichment analysis using the clusterProfiler package for R (v3.18.1) with an FDR-adjusted P value < 0.05 (Yu *et al.,* 2012).

Genome-wide association (GWA) analysis

To assess the putative genetic basis of the climate-predicted DMRs, we ran GWA analysis for the garden conditions and including all the available samples (7 samples per population, N = 147) to increase the statistical power of the analysis. We ran GWA analysis as described in Galanti *et al.,* (2022). For genetic variants, we imputed the missing genotype calls with BEAGLE 5.2 (Browning *et al.,* 2018), filtered for MAF > 0.04 and pruned for Linkage Disequilibrium (LD) with an LD threshold (r^2) of 0.8 for SNP pairs in a sliding window of 50 SNPs, sliding by 5. After filtering, we were able to retain 83 095 SNPs. We used individual average DMR methylation for each sequence context as phenotype, calculated with the *regionCounts* methylKit function (v1.16.1) (Akalin *et al.,* 2012), using a minimum cytosine coverage of 5. Since running the analysis for all the climate-predicted DMRs would be too computationally intensive and time demanding, we selected three random sets of 100 climate-predicted DMRs, and we ran GWA analysis for each of them. To be sure to identify all the potential GWA hits, we chose the less stringent threshold of significance level, and we calculated the average and standard deviation of the significant GWA hits among the three sets of DMRs.

RNA-Sequencing and correlation of climate-predicted DMRs with gene expression

We collected mature leaf samples from 3 randomly selected plants per population from garden condition (total samples = 63), and we froze them immediately in liquid nitrogen. We extracted mRNA using the Nucleospin RNA Plus kit (Macherey Nagel), following the manufacturer's instructions with minor modifications. To improve RNA quality and yield

from *F. vesca*, we used an increased amount of lysis buffer (500 µl) together with 100 µl of EDTA (0.5 M, pH=8) and PVPP (polyvinylpolypyrrolidone). The cDNA library and sequencing (PE150, 6 Gb per sample of raw data) were performed by Novogene Co., Ltd, Cambridge, using an Illumina NovaSeq 6000 platform. On average, we obtained 22252025 raw reads. We trimmed adaptors with cutadapt (v1.16) and assessed sequencing quality with MultiQC (v1.10.1) (Ewels *et al.,* 2016). We aligned the reads to the *Fragaria vesca* genome (v4.0.a2) using STAR (Spliced Transcripts Alignment to a Reference) (v2.7.1a) (Dobin *et al.,* 2013), assembled them into transcripts and quantified using StringTie (v2.1.5) (Kovaka *et al.,* 2019). For each sample, we normalized the Fragments Per Kilobase of transcript per Million mapped reads (FPKM) values and extracted the genes adjacent to the *directly* environmentally induced DMR-promoters. For each of the three 100 random sets of DMR-promoters, we combined DMR-promoter methylation, expression of the adjacent gene, sample ID and gene ID in the same file, and for each individual gene we performed Spearman correlation ($p < 0.05$) between methylation and expression. We calculated the average and standard deviation of the significant correlations found in the three random sets of DMR-promoters.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

The study was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 764965, the Czech Science Foundation (GACR 20-00871S) and partly by institutional research project RVO 67985939.

ACKNOWLEDGMENTS

We thank members of the EpiDiverse consortium (www.epidiverse.eu) for valuable inputs during preparation and execution of the study.

References

Akalin, Altuna, Matthias Kormaksson, Sheng Li, Francine E. Garrett-Bakelman, Maria E. Figueroa, Ari Melnick, and Christopher E. Mason. 2012. "MethylKit: A Comprehensive R Package for the Analysis of Genome-Wide DNA Methylation Profiles." *Genome Biology* 13 (10): 1–9.

Anastasiadi, Dafni, Clare J. Venney, Louis Bernatchez, and Maren Wellenreuther. 2021. "Epigenetic Inheritance and Reproductive Mode in Plants and Animals." *Trends in Ecology & Evolution* 36 (12): 1124–40.

Ashe, Alyson, Vincent Colot, and Benjamin P. Oldroyd. 2021. "How Does Epigenetics Influence the Course of Evolution?" *Philosophical Transactions of the Royal Society B* 376 (1826).

Baduel, Pierre, Basile Leduque, Amandine Ignace, Isabelle Gy, José Gil, Olivier Loudet, Vincent Colot, and Leandro Quadrana. 2021. "Genetic and Environmental Modulation of Transposition Shapes the Evolutionary Potential of Arabidopsis Thaliana." *Genome Biology* 22 (1): 1–26.

Bossdorf, Oliver, Christina L. Richards, and Massimo Pigliucci. 2007. "Epigenetics for Ecologists." *Ecology Letters* 0 (0): 071117033013002-???

Cornes, Richard C., Gerard van der Schrier, Else J. M. van den Besselaar, and Philip D. Jones. 2018. "An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets." *Journal of Geophysical Research: Atmospheres* 123 (17): 9391–9409.

Danecek, Petr, James K. Bonfield, Jennifer Liddle, John Marshall, Valeriu Ohan, Martin O. Pollard, Andrew Whitwham, *et al.,* 2021. "Twelve Years of SAMtools and BCFtools." *GigaScience* 10 (2): 1–4.

Darrow, G. M. 1966. "The Strawberry. History, Breeding and Physiology." *The Strawberry. History, Breeding and Physiology.*

Darwin, Charles, and Alfred Wallace. 1858. "On the Tendency of Species to Form Varieties; and on the Perpetuation of Varieties and Species by Natural Means of Selection." *Journal of the Proceedings of the Linnean Society of London. Zoology* 3 (9): 45–62.

Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. 2013. "STAR: Ultrafast Universal RNA-Seq Aligner." *Bioinformatics (Oxford, England)* 29 (1): 15–21.

Dodd, Richard S., and Vladimir Douhovnikoff. 2016. "Adjusting to Global Change through Clonal Growth and Epigenetic Variation." *Frontiers in Ecology and Evolution* 0 (JUL): 86.

Dubin, Manu J., Pei Zhang, Dazhe Meng, Marie Stanislas Remigereau, Edward J. Osborne, Francesco Paolo Casale, Philipp Drewe, *et al.,* 2015. "DNA Methylation in Arabidopsis Has a Genetic Basis and Shows Evidence of Local Adaptation." *ELife* 4 (MAY).

Edger, Patrick P., Robert VanBuren, Marivi Colle, Thomas J. Poorten, Ching Man Wai, Chad E. Niederhuth, Elizabeth I. Alger, *et al.,* 2018. "Single-Molecule Sequencing and Optical Mapping Yields an Improved Genome of Woodland Strawberry (*Fragaria Vesca*) with Chromosome-Scale Contiguity." *GigaScience* 7 (2).

Ewels, Philip, Måns Magnusson, Sverker Lundin, and Max Käller. 2016. "MultiQC: Summarize Analysis Results for Multiple Tools and Samples in a Single Report." *Bioinformatics (Oxford, England)* 32 (19): 3047–48.

Feng, Suhua, Steven E. Jacobsen, and Wolf Reik. 2010. "Epigenetic Reprogramming in Plant and Animal Development." *Science (New York, N.Y.)* 330 (6004): 622–27.

Finnegan, E. J., R. K. Genger, W. J. Peacock, and E. S. Dennis. 1998. "DNA METHYLATION IN PLANTS." *Annual Review of Plant Physiology and Plant Molecular Biology* 49 (1): 223–47.

Fojtová, Miloslava, Aleš Kovařík, and Roman Matyášek. 2001. "Cytosine Methylation of Plastid Genome in Higher Plants. Fact or Artefact?" *Plant Science : An International Journal of Experimental Plant Biology* 160 (4): 585–93.

Frels, Katherine, Ratan Chopra, Kevin M. Dorn, Donald L. Wyse, M. David Marks, and James A. Anderson. 2019. "Genetic Diversity of Field Pennycress (Thlaspi Arvense) Reveals Untapped Variability and Paths Toward Selection for Domestication." *Agronomy 2019, Vol. 9, Page 302* 9 (6): 302.

Fultz, Dalen, Sarah G. Choudury, and R. Keith Slotkin. 2015. "Silencing of Active Transposable Elements in Plants." *Current Opinion in Plant Biology* 27 (October): 67–76.

Ganguly, Diep R., Peter A. Crisp, Steven R. Eichten, and Barry J. Pogson. 2017. "The Arabidopsis DNA Methylome Is Stable under Transgenerational Drought Stress." *Plant Physiology* 175 (4): 1893–1912.

Gent, Jonathan I., Nathanael A. Ellis, Lin Guo, Alex E. Harkess, Yingyin Yao, Xiaoyu Zhang, and R. Kelly Dawe. 2013. "CHH Islands: De Novo DNA Methylation in near-Gene Chromatin Regulation in Maize." *Genome Research* 23 (4): 628–37.

Jablonka, E. V.A., and G. A.L. Raz. 2009. "Transgenerational Epigenetic Inheritance: Prevalence, Mechanisms, and Implications for the Study of Heredity and Evolution." *Quarterly Review of Biology* 84 (2): 131–76.

Johannes, Frank, and Robert J. Schmitz. 2019. "Spontaneous Epimutations in Plants." *New Phytologist* 221 (3): 1253–59.

Jühling, Frank, Helene Kretzmer, Stephan H. Bernhart, Christian Otto, Peter F. Stadler, and Steve Hoffmann. 2016. "Metilene: Fast and Sensitive Calling of Differentially Methylated Regions from Bisulfite Sequencing Data." *Genome Research* 26 (2): 256–62.

Jung, Sook, Taein Lee, Chun Huai Cheng, Katheryn Buble, Ping Zheng, Jing Yu, Jodi Humann, *et al.,* 2019. "15 Years of GDR: New Data and Functionality in the Genome Database for Rosaceae." *Nucleic Acids Research* 47 (Database issue): D1137.

Klimeš, L., J. Klimešová, Rob J. J. Hendriks, and J. Groenendael. 1997. "Clonal Plant Architecture A Comparative Analysis of Form and Function." *The Ecology and Evolution of Clonal Plants*.

Kort, Hanne De, Bart Panis, Kenny Helsen, Rolland Douzet, Steven B. Janssens, and Olivier Honnay. 2020. "Pre-Adaptation to Climate Change through Topography-Driven Phenotypic Plasticity." *Journal of Ecology* 108 (4): 1465–74.

Kort, Hanne De, Tuomas Toivainen, Filip Van Nieuwerburgh, Javier Andrés, Timo P. Hytönen, and Olivier Honnay. 2022. "Signatures of Polygenic Adaptation Align with Genome-Wide Methylation Patterns in Wild Strawberry Plants." *New Phytologist*, August..

Kovaka, Sam, Aleksey V. Zimin, Geo M. Pertea, Roham Razaghi, Steven L. Salzberg, and Mihaela Pertea. 2019. "Transcriptome Assembly from Long-Read RNA-Seq Alignments with StringTie2." *Genome Biology* 20 (1): 1–13.

Lang, Zhaobo, Yihai Wang, Kai Tang, Dengguo Tang, Tatsiana Datsenka, Jingfei Cheng, Yijing Zhang, Avtar K. Handa, and Jian Kang Zhu. 2017. "Critical Roles of DNA Demethylation in the Activation of Ripening-Induced Genes and Inhibition of Ripening-Repressed Genes in Tomato Fruit." *Proceedings of the National Academy of Sciences of the United States of America* 114 (22): E4511–19.

Latzel, Vít, and Jitka Klimešová. 2010. "Transgenerational Plasticity in Clonal Plants." *Evolutionary Ecology* 24 (6): 1537–43.

Latzel, Vít, Alejandra P. Rendina González, and Jonathan Rosenthal. 2016. "Epigenetic Memory as a Basis for Intelligent Behavior in Clonal Plants." *Frontiers in Plant Science* 7 (AUG2016): 1354.

Li, Xin, Jingde Zhu, Fengyi Hu, Song Ge, Mingzhi Ye, Hui Xiang, Guojie Zhang, *et al.,* 2012. "Single-Base Resolution Maps of Cultivated and Wild Rice Methylomes and Regulatory Roles of DNA Methylation in Plant Gene Expression." *BMC Genomics* 13 (1): 1–15.

Li, Yongping, Mengting Pi, Qi Gao, Zhongchi Liu, and Chunying Kang. 2019. "Updated Annotation of the Wild Strawberry *Fragaria Vesca* V4 Genome." *Horticulture Research* 6 (1).

López, María-Estefanía, David Roquis, Claude Becker, Béatrice Denoyes, and Etienne Bucher. 2022. "DNA Methylation Dynamics during Stress Response in Woodland Strawberry (*Fragaria Vesca*)." *Horticulture Research* 9 (January).

Martin, Antoine, Christelle Troadec, Adnane Boualem, Mazen Rajab, Ronan Fernandez, Halima Morin, Michel Pitrat, Catherine Dogimont, and Abdelhafid Bendahmane. 2009. "A Transposon-Induced Epigenetic Change Leads to Sex Determination in Melon." *Nature* 461 (7267): 1135–38.

Martin, Galen T., Danelle K. Seymour, and Brandon S. Gaut. 2021. "CHH Methylation Islands: A Nonconserved Feature of Grass Genomes That Is Positively Associated with Transposable Elements but Negatively Associated with Gene-Body Methylation." *Genome Biology and Evolution* 13 (8).

Medrano, Mónica, Conchita Alonso, Pilar Bazaga, Esmeralda López, and Carlos M. Herrera. 2020. "Comparative Genetic and Epigenetic Diversity in Pairs of Sympatric, Closely Related Plants with Contrasting Distribution Ranges in South-Eastern Iberian Mountains." *AoB Plants* 12 (3).

Miryeganeh, Matin, Ferdinand Marlétaz, Daria Gavriouchkina, and Hidetoshi Saze. 2022. "De Novo Genome Assembly and in Natura Epigenomics Reveal Salinity-Induced DNA Methylation in the Mangrove Tree Bruguiera Gymnorhiza." *The New Phytologist* 233 (5): 2094–2110.

Münzbergová, Zuzana, Vít Latzel, Maria Šurinová, and Věroslava Hadincová. 2019. "DNA Methylation as a Possible Mechanism Affecting Ability of Natural Populations to Adapt to Changing Climate." *Oikos* 128 (1): 124–34.

Niederhuth, Chad E., and Schmitz, Robert J. 2014. "Covering Your Bases: Inheritance of DNA Methylation in Plant Genomes." *Molecular Plant* 7 (3): 472–80.

Niederhuth, Chad E., and Schmitz, Robert J. 2017. "Putting DNA Methylation in Context: From Genomes to Gene Expression in Plants." *Biochimica et Biophysica Acta* 1860 (1): 149.

Nunn, Adam, Sultan Nilay Can, Christian Otto, Mario Fasold, Bárbara Diéz Rodríguez, Noé Fernández-Pozo, Stefan A. Rensing, Peter F. Stadler, and David Langenberger. 2021. "EpiDiverse Toolkit: A Pipeline Suite for the Analysis of Bisulfite Sequencing Data in Ecological Plant Epigenetics." *NAR Genomics and Bioinformatics* 3 (4).

Nunn, Adam, Christian Otto, Mario Fasold, Peter F. Stadler, and David Langenberger. 2022. "Manipulating Base Quality Scores Enables Variant Calling from Bisulfite Sequencing Alignments Using Conventional Bayesian Approaches." *BioRxiv*, March, 2021.01.11.425926..

Oksanen Jari, Blanchet F. Guillaume, Friendly Michael, Kindt Roeland, Legendre Pierre, McGlinn Dan, Minchin Peter R., *et al.,* 2020. "CRAN - Package Vegan." 2020. https://cran.r-project.org/web/packages/vegan/index.html.

Oliver, Keith R., and Wayne K. Greene. 2009. "Transposable Elements: Powerful Facilitators of Evolution." *BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology* 31 (7): 703–14.

Ou, Shujun, Weija Su, Yi Liao, Kapeel Chougule, Jireh R.A. Agda, Adam J. Hellinga, Carlos Santiago Blanco Lugo, *et al.,* 2019. "Benchmarking Transposable Element Annotation Methods for Creation of a Streamlined, Comprehensive Pipeline." *Genome Biology* 20 (1).

Purcell, Shaun, Benjamin Neale, Kathe Todd-Brown, Lori Thomas, Manuel A.R. Ferreira, David Bender, Julian Maller, *et al.,* 2007. "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses." *American Journal of Human Genetics* 81 (3): 559–75.

Quinlan, Aaron R., and Ira M. Hall. 2010. "BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features." *Bioinformatics (Oxford, England)* 26 (6): 841–42.

Raj, Sherosha, Katharina Bräutigam, Erin T. Hamanishi, Olivia Wilkins, Barb R. Thomas, William Schroeder, Shawn D. Mansfield, Aine L. Plant, and Malcolm M. Campbell. 2011. "Clone History Shapes Populus Drought Responses." *Proceedings of the National Academy of Sciences of the United States of America* 108 (30): 12521–26.

Rajkumar, Mohan Singh, Rama Shankar, Rohini Garg, and Mukesh Jain. 2020. "Bisulphite Sequencing Reveals Dynamic DNA Methylation under Desiccation and Salinity Stresses in Rice Cultivars." *Genomics* 112 (5): 3537–48.

Rendina González, Alejandra Pilar, Veronica Preite, Koen J. F. Verhoeven, and Vít Latzel. 2018. "Transgenerational Effects and Epigenetic Memory in the Clonal Plant Trifolium Repens." *Frontiers in Plant Science* 9 (November): 1677.

Rey, Olivier, Etienne Danchin, Marie Mirouze, Céline Loot, and Simon Blanchet. 2016. "Adaptation to Global Change: A Transposable Element-Epigenetics Perspective." *Trends in Ecology and Evolution*. Elsevier Ltd.

Richards, Christina L., Conchita Alonso, Claude Becker, Oliver Bossdorf, Etienne Bucher, Maria Colomé-Tatché, Walter Durka, *et al.,* 2017. "Ecological Plant Epigenetics: Evidence from Model and Non-Model Species, and the Way Forward." *Ecology Letters* 20 (12): 1576–90.

Richards, Christina L., Aaron W. Schrey, and Massimo Pigliucci. 2012. "Invasion of Diverse Habitats by Few Japanese Knotweed Genotypes Is Correlated with Epigenetic Differentiation." Edited by Mark Vellend. *Ecology Letters* 15 (9): 1016–25.

Richards, Eric J. 2006. "Inherited Epigenetic Variation — Revisiting Soft Inheritance." *Nature Reviews Genetics 2006 7:5* 7 (5): 395–401..

Riggs, Arthur D., and Thomas N. Porter. 1996. "Overview of Epigenetic Mechanisms." *Cold Spring Harbor Monograph Archive* 32 (0): 29–45.

Sammarco, Iris, Zuzana Münzbergová, and Vít Latzel. 2022. "DNA Methylation Can Mediate Local Adaptation and Response to Climate Change in the Clonal Plant *Fragaria Vesca*: Evidence From a European-Scale Reciprocal Transplant Experiment." *Frontiers in Plant Science* 13 (February): 435.

Schmidt, Adele L., and Lucy M. Anderson. 2006. "Repetitive DNA Elements as Mediators of Genomic Change in Response to Environmental Cues." *Biological Reviews of the Cambridge Philosophical Society* 81 (4): 531–43.

Schulze, Juerg, Rita Rufener, Andreas Erhardt, and Peter Stoll. 2012. "The Relative Importance of Sexual and Clonal Reproduction for Population Growth in the Perennial Herb *Fragaria Vesca*." *Population Ecology 2012 54:3* 54 (3): 369–80.

Secco, David, Chuang Wang, Huixia Shou, Matthew D Schultz, Serge Chiarenza, Laurent Nussaume, Joseph R Ecker, James Whelan, and Ryan Lister. 2015. "Stress Induced Gene Expression Drives Transient DNA Methylation Changes at Adjacent Repetitive Elements." *ELife* 4 (July).

Shi, Wen, Xiaojie Chen, Lexuan Gao, Cheng-Yuan Xu, Xiaokun Ou, Oliver Bossdorf, Ji Yang, and Yupeng Geng. 2019. "Transient Stability of Epigenetic Population Differentiation in a Clonal Invader." *Frontiers in Plant Science* 9 (March): 1851.

Sigg, Christian D., and Joachim M. Buhmann. 2008. "Expectation-Maximization for Sparse and Non-Negative PCA." *Proceedings of the 25th International Conference on Machine Learning*, 960–67.

Simonetti, G., N. Pucci, E. Brasili, A. Valletta, I. Sammarco, E. Carnevale, G. Pasqua, and S. Loreti. 2019. "In Vitro Antimicrobial Activity of Plant Extracts against Pseudomonas Syringae Pv. Actinidiae Causal Agent of Bacterial Canker in Kiwifruit." *Plant Biosystems*.

Slotkin, R. Keith, and Robert Martienssen. 2007. "Transposable Elements and the Epigenetic Regulation of the Genome." *Nature Reviews Genetics*. Nature Publishing Group..

Thiebaut, Flávia, Adriana Silva Hemerly, and Paulo Cavalcanti Gomes Ferreira. 2019. "A Role for Epigenetic Regulation in the Adaptation and Stress Responses of Non-Model Plants." *Frontiers in Plant Science*. Frontiers Media S.A.

Todd, Lauren A., Amanda C. Hall, Violena Pietrobon, Janet N.Y. Chan, Guillaume Laflamme, and Karim Mekhail. 2020. "RNA-CDNA Hybrids Mediate Transposition via Different Mechanisms." *Scientific Reports 2020 10:1* 10 (1): 1–12.

Verhoeven, Koen J.F., and Veronica Preite. 2014. "Epigenetic Variation in Asexually Reproducing Organisms." Evolution 68 (3): 644–55.

Wang, Guanqun, Haoxuan Li, Shuan Meng, Jianchang Yang, Nenghui Ye, and Jianhua Zhang. 2020. "Analysis of Global Methylome and Gene Expression during Carbon Reserve Mobilization in Stems under Soil Drying." *Plant Physiology* 183 (4): 1809–24.

Warwick, S. I., A. Francis, and D. J. Susko. 2011. "The Biology of Canadian Weeds. 9. Thlaspi Arvense L. (Updated)." *Https://Doi.Org/10.4141/P01-159* 82 (4): 803–23.

Wibowo, Anjar, Claude Becker, Gianpiero Marconi, Julius Durr, Jonathan Price, Jorg Hagmann, Ranjith Papareddy, *et al.,* 2016. "Hyperosmotic Stress Memory in Arabidopsis Is

Mediated by Distinct Epigenetically Labile Sites in the Genome and Is Restricted in the Male Germline by DNA Glycosylase Activity." *ELife* 5 (MAY2016).

Xu, Jidi, Shasha Zhou, Xiaoqing Gong, Yi Song, Steve van Nocker, Fengwang Ma, and Qingmei Guan. 2018. "Single-Base Methylome Analysis Reveals Dynamic Epigenomic Differences Associated with Water Deficit in Apple." *Plant Biotechnology Journal* 16 (2): 672–87.

Yu, Guangchuang, Li Gen Wang, Yanyan Han, and Qing Yu He. 2012. "ClusterProfiler: An R Package for Comparing Biological Themes among Gene Clusters." *Omics : A Journal of Integrative Biology* 16 (5): 284–87.

Zemach, Assaf, Ivy E. McDaniel, Pedro Silva, and Daniel Zilberman. 2010. "Genome-Wide Evolutionary Analysis of Eukaryotic DNA Methylation." *Science (New York, N.Y.)* 328 (5980): 916–19.

Zemach, Assaf, and Daniel Zilberman. 2010. "Evolution of Eukaryotic DNA Methylation and the Pursuit of Safer Sex." *Current Biology* 20 (17): R780–85.

Zhang, Yuan-Ye, Markus Fischer, Vincent Colot, and Oliver Bossdorf. 2013. "Epigenetic Variation Creates Potential for Evolution of Plant Phenotypic Plasticity." *New Phytologist* 197

Zhang, Yuan-Ye, Vit Latzel, Markus Fischer, and Oliver Bossdorf. 2018. "Understanding the Evolutionary Potential of Epigenetic Variation: A Comparison of Heritable Phenotypic Variation in EpiRILs, RILs, and Natural Ecotypes of Arabidopsis Thaliana." *Heredity* 121 (3): 257–65.

Zoldoš, Vlatka, Ivan Biruš, Edina Muratović, Zlatko Šatović, Aleksandar Vojta, Odile Robin, Fatima Pustahija, Faruk Bogunić, Vedrana Vičić Bočkor, and Sonja Siljak-Yakovlev. 2018. "Epigenetic Differentiation of Natural Populations of Lilium Bosniacum Associated with Contrasting Habitat Conditions." *Genome Biology and Evolution* 10 (1): 291–303.

Supplementary material

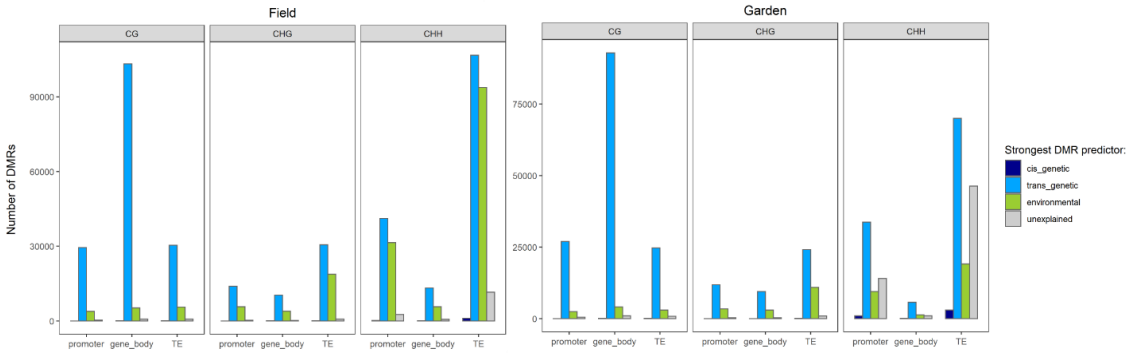


Figure S1: Number of cis-, trans-, climate-predicted and unexplained DMRs overlapping gene promoters, gene bodies or transposable elements (TEs). (a) Field plants: N = 84. (b) Garden plants: N = 84.

Table S1: redundancy analysis (RDA). Percentage of epigenetic (CG, CHG and CHH) and genetic (SNPs) variance explained by country of origin and growing conditions. Significant values (P ≤ 0.05) are shown in bold. Field plants: N = 84, garden plants: N = 84.

	Country	*P*	Growing Condition	*P*
CG	**11.9**	**0.001**	0.4	0.986
CHG	**9.4**	**0.001**	**1.6**	**0.001**
CHH	**2.5**	**0.001**	**1.2**	**0.001**
SNPs	**8.7**	**0.001**	0.1	1

Table S2: Characteristics of the sites of origin of the populations of *F. vesca* used for this study, and of the common garden where the plants were cultivated. Country, population ID, population size, geographic coordinates, and climatic variables (mean, maximum, minimum temperature, and precipitation) averaged over 7 years before the sampling year (2011-2018). For the common garden, the climatic variables refer to the year of cultivation of the plants in such conditions (2018-2019). For this condition, we do not report precipitation as these plants were watered regularly.

COUNTRY	POPULATION ID / GARDEN	ELEVATION (M)	POPULATION SIZE (M2)	LATITUDE (°N)	LONGITUDE (°E)	MEAN TEMP (°C)	MAX TEMP (°C)	MIN TEMP (°C)	PRECIPITATION (MM/DAY)
CZECHIA	FV_CZ_01	201	45	50.399	14.412	10.62	15.72	5.89	1.4
CZECHIA	FV_CZ_02	217	81	50.269	14.757	10.35	15.41	5.85	1.5
CZECHIA	FV_CZ_03	306	156	50.459	14.785	9.49	14.55	5.07	1.64
CZECHIA	FV_CZ_04	454	36	50.646	15.340	8.06	12.65	4.04	2.27
CZECHIA	FV_CZ_05	765	66	50.751	15.345	6.51	11	2.7	2.83
CZECHIA	FV_CZ_06	875	42	50.811	15.359	5.59	9.84	2.03	3.06
CZECHIA	FV_CZ_07	848	216	50.783	15.348	6.51	11	2.7	2.83
ITALY	FV_IT_01	352	80	46.037	11.097	10.15	15.71	5.33	2.73
ITALY	FV_IT_02	1450	330	46.051	11.363	8.58	13.97	3.36	2.9
ITALY	FV_IT_03	468	140	46.471	11.343	10.34	16.21	5.9	2.33
ITALY	FV_IT_04	606	100	46.645	11.133	11	17.11	6.01	2.25
ITALY	FV_IT_05	1436	800	46.726	11.429	2.46	6.37	-1.32	2.76

Table S2. Continuation

COUNTRY	POPULATION ID / GARDEN	ELEVATION (M)	POPULATION SIZE (M2)	LATITUDE (°N)	LONGITUDE (°E)	MEAN TEMP (°C)	MAX TEMP (°C)	MIN TEMP (°C)	PRECIPITATION (MM/DAY)
ITALY	FV_IT_06	1905	400	46.320	11.737	3.55	7.22	-0.12	3.11
ITALY	FV_IT_07	1590	1500	46.595	10.779	2.57	7.19	-1.79	2.26
NORWAY	FV_NO_01	131	20	60.290	10.176	5.45	10.28	1.53	1.93
NORWAY	FV_NO_02	597	160	61.036	9.079	2.88	7.45	-1.18	1.95
NORWAY	FV_NO_03	818	12	60.821	8.706	1.5	5.93	-2.6	2.12
NORWAY	FV_NO_04	386	105	60.881	7.166	2.98	6.38	-0.15	4.16
NORWAY	FV_NO_05	60	63	60.840	7.123	2.98	6.38	-0.15	4.16
NORWAY	FV_NO_06	244	60	60.885	7.339	1.25	4.43	-1.86	3.58
NORWAY	FV_NO_07	510	100	60.896	7.353	1.25	4.43	-1.86	3.58
CZECHIA	COMMON GARDEN		/	50.663	14.951	9.62	14.94	4.92	/

CHAPTER VI

Tailored tools for ecological epigenetics

Adam Nunn, Sultan Nilay Can, Christian Otto, Mario Fasold,
Bárbara Díez Rodríguez, Noé Fernández-Pozo, Stefan A Rensing,
Peter F Stadler and David Langenberger

This chapter has been published in *NAR Genomics and Bioinformatics*, Volume 3, Issue 4 (December 2021)

Published online 12 November 2021

NAR Genomics and Bioinformatics, 2021, Vol. 3, No. 4 **1**
https://doi.org/10.1093/nargab/lqab106

EpiDiverse Toolkit: a pipeline suite for the analysis of bisulfite sequencing data in ecological plant epigenetics

Adam Nunn [1,2], Sultan Nilay Can [3], Christian Otto [1], Mario Fasold [1], Bárbara Díez Rodríguez [3], Noé Fernández-Pozo [3,4], Stefan A. Rensing [3,5,6], Peter F. Stadler [2,7,8,9,10,11] and David Langenberger [1,*]

[1]ecSeq Bioinformatics GmbH, Leipzig 04103, Germany, [2]Institute for Computer Science, University of Leipzig, Leipzig 04107, Germany, [3]Deparment of Biology, University of Marburg, Marburg 35043, Germany, [4]Institute for Mediterranean and Subtropical Horticulture, Málaga 29010, Spain, [5]Center for Biological Signaling Studies (BIOSS), University of Freiburg, Freiburg 79104, Germany, [6]Center for Synthetic Microbiology (SYNMIKRO), University of Marburg, Marburg 35043, Germany, [7]Interdisciplinary Center for Bioinformatics; German Center for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig; Competence Center for Scalable Data Services and Solutions; Leipzig Research Center for Civilization Diseases; and Leipzig Research Center for Civilization Diseases (LIFE), University of Leipzig, Leipzig 04109, Germany, [8]Max Planck Institute for Mathematics in the Sciences, Leipzig 04103, Germany, [9]Institute for Theoretical Chemistry, University of Vienna, Vienna 1090, Austria, [10]Facultad de Ciencias, Universidad National de Colombia, Sede Bogotá, Colombia and [11]Santa Fe Institute, Santa Fe, USA

Received May 14, 2021; Revised September 29, 2021; Editorial Decision October 20, 2021; Accepted October 22, 2021

ABSTRACT

The expanding scope and scale of next generation sequencing experiments in ecological plant epigenetics brings new challenges for computational analysis. Existing tools built for model data may not address the needs of users looking to apply these techniques to non-model species, particularly on a population or community level. Here we present a toolkit suitable for plant ecologists working with whole genome bisulfite sequencing; it includes pipelines for mapping, the calling of methylation values and differential methylation between groups, epigenome-wide association studies, and a novel implementation for both variant calling and discriminating between genetic and epigenetic variation.

INTRODUCTION

Model organisms such as *Arabidopsis thaliana* have helped lay the foundation for our understanding of plant epigenetics (1–3), often proceeding DNA methylation profiling techniques such as whole genome bisulfite sequencing (WGBS) to study the DNA methylome at a nucleotide-level resolution. Historically, this practice has been considered by many as the 'gold-standard' for DNA methylation analysis, but can also be prohibitively expensive beyond a focus on model species (4). Cost-effective alternatives, such as affinity-based enrichment (e.g. MeDIP-seq, MDB-seq) or restriction-enzyme digestion (e.g. RRBS, MSCC), necessitate narrower hypotheses and risk spurious findings by neglecting the broader relationships detectable by more comprehensive methods. Now, the increasingly competitive costs of next generation sequencing (NGS) have opened the door for plant ecologists to apply previous lessons from WGBS on the population and community level, to gain more specific insight into non-model species (5). The EpiDiverse Toolkit addresses the challenges of expanding scope and scale for existing computational techniques, with a suite of pipelines to streamline the analysis of DNA methylation from bisulfite sequencing (bs-seq; methylC-seq) data under FAIR principles (Findability, Accessibility, Interoperability and Reusability). The aim is to provide a flexible and standardised approach when implementing 'gold-standard' DNA methylation analyses for non-model species in plant ecology, which additionally offers some minor improvements to further cut cost and improve computational efficiency.

The basis of bisulfite sequencing is to differentiate methylated and unmethylated cytosine nucleotides. During NGS library preparation, sodium bisulfite treatment facilitates the conversion of unmethylated cytosine to uracil while leaving 5-methylcytosine (5mC) positions intact (6). This necessitates specialised or adapted tools to carry out conventional downstream procedures such as mapping (7) and variant calling (8). For non-model plant species this is fur-

*To whom correspondence should be addressed. Tel: +49 341 425 891 99; Fax: +49 341 33187 962; Email: david.langenberger@ecseq.com

ther confounded by poor quality reference genomes, with additional difficulties due in part to a high tolerance for polyploidy and high rates of heterozygosity (9). All of these aspects present difficulties in terms of running time and the optimisation of computational resources. Finally, DNA methylation can occur in additional sequence contexts (CHG, CHH) which in contrast to CG are not prevalent in mammalian data (10).

The tools presented herein (Figure 1) are implemented with Nextflow (11), building on best-practice concepts outlined by nf-core (12). They are intended to be efficient, intuitive for novice users, optimisable for laptop, HPC cluster or the cloud, and scalable from small lab studies to field trials with large populations. A list of individual pipeline processes alongside the default, recommended resource configurations are provided in Supplementary Table S1. Each resource allocation is fully customisable under the Nextflow framework to suit integration under different systems and scheduling software. Dependencies are as simple as installing Nextflow alongside one of either Bioconda (13), Docker (14) or Singularity (15) on a POSIX compatible system, facilitating a high level of flexibility and reproducibility through the use of portable software containers and environments. Each pipeline is fully self-contained and can be easily transferred from one system to another without need for specific, manual installation of the component software. The toolkit is open-source and publicly available on GitHub, allowing users to fork and modify the pipelines at their own discretion including access to the entire change history. The toolkit represents a starting point for the standardisation of DNA methylation profiling in ecological plant epigenetics, and will be actively maintained and expanded upon as additional tools are developed in the future. All pipeline output is streamlined to standard, recognised formats to facilitate interoperability with external software and help create flexible analyses for a wide range of possible experiments, for example when intersecting methylation bedGraph files with gene or transposable element (TE) annotations using BEDTools (16).

MATERIALS AND METHODS

Test data

In order to demonstrate selected features from the toolkit, a subset of 23 independent, whole genome bisulfite sequencing libraries (150 bp long paired-end reads) of the deciduous tree species *Populus nigra* were selected from the repository hosted by the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB44879 (https://www.ebi.ac.uk/ena/browser/view/PRJEB44879). The libraries were sequenced under the broader initiative of the EpiDiverse consortium according to the procedures outlined by Díez Rodríguez *et al.* (manuscript in prep.). This subset represents two clone populations (Supplementary Table S2) derived from cuttings originating from field sites in Germany and Lithuania and cultivated together under common garden conditions. Measurements of leaf flavonol content from the parent generation were derived from observations taken in the field by Díez Rodríguez *et al.* (manuscript in prep.). The reference genome was obtained from the repository hosted by the ENA at EMBL-EBI under accession number PRJEB44889 (https://www.ebi.ac.uk/ena/browser/view/PRJEB44889).

Whole genome bisulfite sequencing (WGBS)

The EpiDiverse WGBS pipeline derives sequence alignments in BAM format from input NGS reads in FASTQ format and a provided reference genome in FASTA format, which are taken forward to estimate the methylation level over each position under the given methylation context(s), in bedGraph format. The reference genome is optionally indexed by the pipeline itself, or provided alongside the relevant index files to begin with. Mapping of bisulfite sequencing data can be carried out either in 'high-throughput mode', with a low memory footprint and a runtime suitable for rapid analysis of population data, or 'high-sensitivity mode', with a demonstrable improvement in precision-recall and downstream methylation analysis for non-model plant species, as selected according to corresponding benchmarks (17). Multiple samples can be processed in parallel, and quality control (QC) is performed with a combination of published tools and in-house scripting. Basic visualisation of alignment statistics is performed with samtools stats and the corresponding plot-bamstats tool (18). Methylation values based on coverage are called with MethylDackel (https://github.com/dpryan79/MethylDackel), which also provides QC for M-bias analysis and overlapping paired-end reads.

Variant calling and sample clustering (SNP)

The EpiDiverse SNP pipeline performs a novel masking procedure which compares individual nucleotides from the bisulfite sequencing alignments, obtained from the EpiDiverse WGBS pipeline, to the reference genome. Joint variant calling is then performed on the masked BAM files, to provide single nucleotide polymorphisms (SNPs) in standard VCF/BCF format which are filtered by the pipeline according to customisable parameters. As SNPs in a cytosine-to-thymine context are obscured in bisulfite data (8), neither variant calling nor sample methylation clustering can be resolved using conventional methods. A simple post-processing procedure for *in silico* manipulation of both base qualities and base nucleotides in bisulfite contexts, following alignment, has been shown to facilitate conventional SNP calling on WGBS data which outperforms equivalent, specialised software (19). This heuristic method has been implemented herein and enables a) downstream analysis with tools that are already well-established for DNA-seq such as Freebayes (20), and b) sample clustering with kWIP (21) which uses k-mer diversity to estimate a distance matrix. Variant calling in this manner can eliminate the need for conventional DNA-seq data alongside bisufite sequencing data, thus reducing sequencing costs for plant ecologists. Basic visualisation of variant statistics is also carried out using bcftools stats and the corresponding plot-vcfstats tool (22).

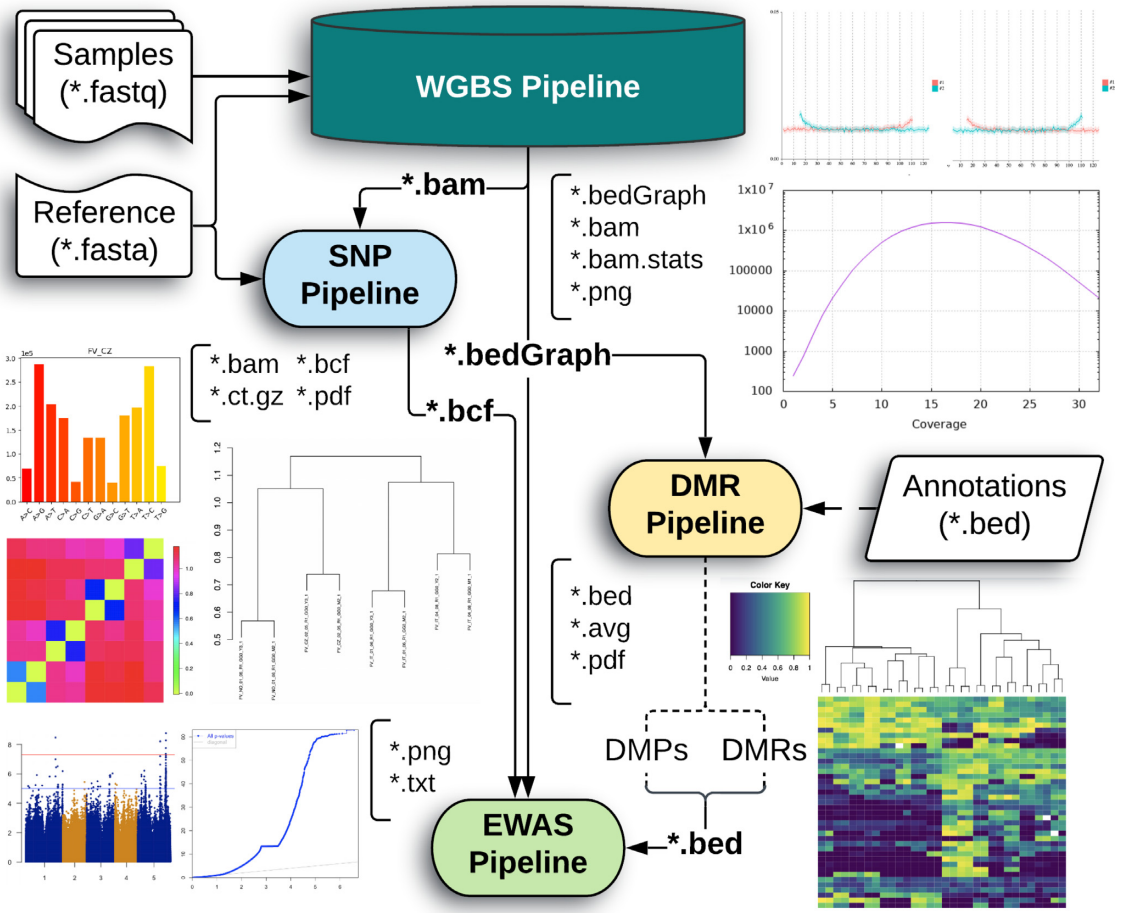


Figure 1. Overview of the EpiDiverse Toolkit. The WGBS data forms the foundation of the analysis, and each downstream pipeline is built to work either in cooperation with one another or, optionally, with independently-generated input data. All pipelines output runtime metadata, tracing and further visualisation in addition to what is shown here. The full output is described for each pipeline in the documentation on Github.

Differential methylation (DMR)

The EpiDiverse DMR pipeline analyses statistically significant differential methylation from a collection of sample-specific methylation files in bedGraph format obtained from the EpiDiverse WGBS pipeline, providing the output in a custom BED format. A recent benchmark demonstrated a higher sensitivity for finding DMRs with metilene in comparison to other tools (23). Pairwise comparisons of methylation profiles between groups are therefore made with metilene (24), to derive either differentially methylated regions (DMRs) or positions (DMPs) while also correcting for multiple comparisons. Any annotations in BED format can also be provided by the user, as pre-selected regions for comparison, as an alternative to the default boundary estimation based on the methylation signal. Due to the non-parametric statistical test, each methylation context (CG, CHG, CHH) can be analysed independently (or combined) without any *a priori* assumptions about the underlying distribution of methylation values. Significant DMRs in terms of hyper- and hypo-methylation are visualised using custom Rscripts to provide density plots and heatmaps.

Epigenome-wide association studies (EWAS)

For a given population of samples, the output derived from previous aspects of the toolkit (i.e. methylation files in bedGraph format, SNP variants in VCF format, annotations such as DMPs/DMRs in BED format) can be combined and processed using the EpiDiverse EWAS pipeline (25) for analysis using the GEM suite (26), in order to study the association between epigenetics, genetics, and environmental metadata through the identification of quantitative trait loci (QTL). These QTLs can be discovered either by taking the full set of methylated positions, in any methylation context, or by first subsetting according to provided annotations (e.g. DMPs/DMRs), or even by taking the provided annotations themselves in place of methylated positions for use as genomic markers, whereby the pipeline will calculate the average methylation level in each case by intersecting the methylated positions. The confounding genetic component can be resolved in each case by providing the SNPs derived in the first place from the same bisulfite data, without the need for conventional whole genome sequencing data alongside.

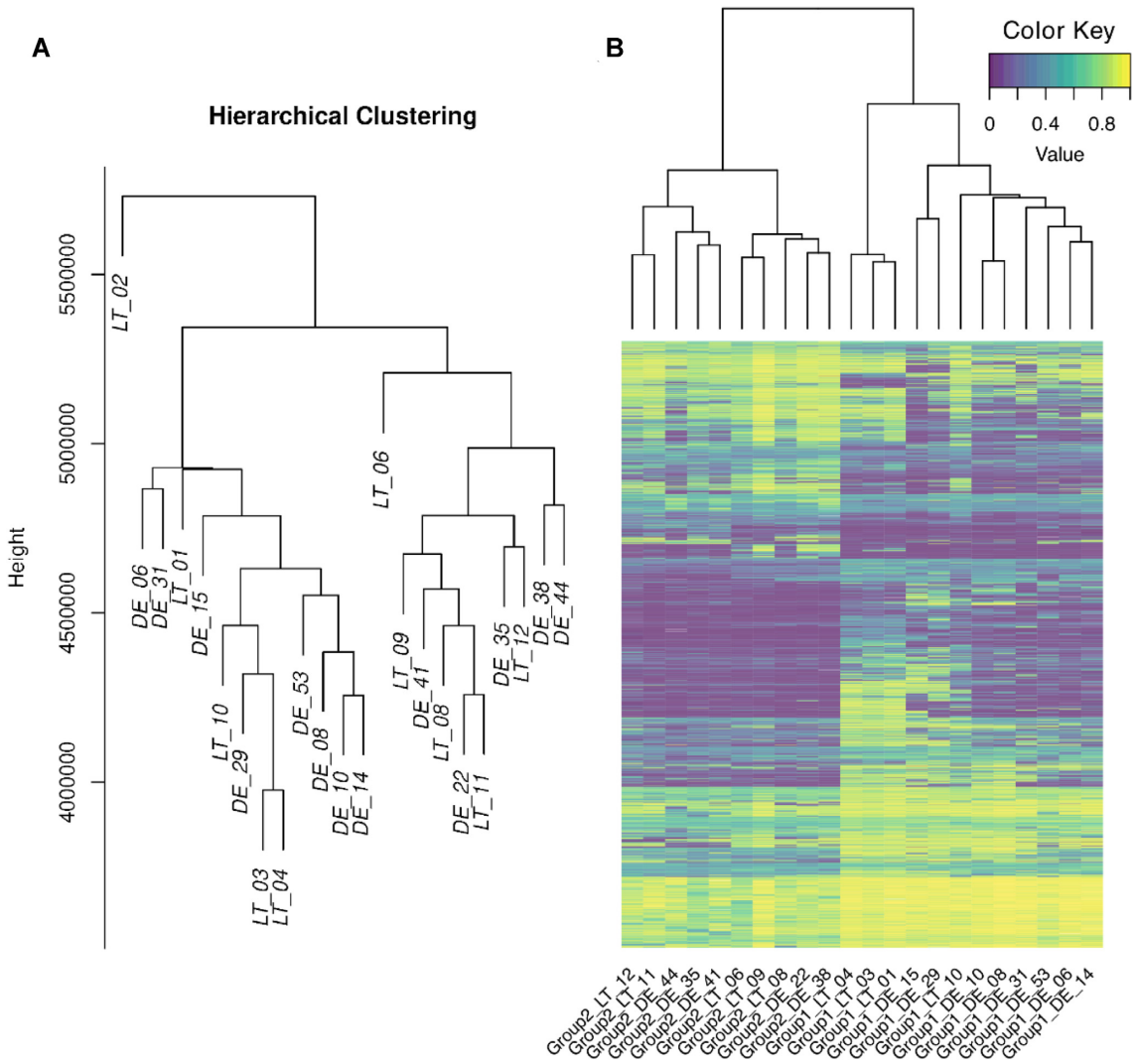


Figure 2. (**A**) Hierarchical clustering of methylated sites (all contexts) derived from the cohort of *P. nigra* samples from populations in Germany and Lithuania, and (**B**) the resulting heatmap of significant DMRs ($q < 0.05$) obtained after cutting the hierarchical tree at 5.25×10^6 to form two discrete groups (leaving LT_02 as outlier). Either plot can be obtained using the EpiDiverse toolkit.

RESULTS AND DISCUSSION

The 23 independent WGBS libraries were first mapped in 'high-throughput mode' with the EpiDiverse WGBS pipeline, resulting in mapping rates ranging from 78.38% to 80.44% under default parameter settings (Supplementary Table S3). The global methylation level in all contexts is reported in Supplementary Figure S1, alongside a principal component analysis demonstrating the unsupervised grouping of all samples based on the variation in shared methylated sites.

Following alignment, variant calling was performed with the EpiDiverse SNP pipeline to identify SNPs from bisulfite-treated data based on sequence masking and base quality manipulation (19). The total number of variants in each sample are summarised in Supplementary Table S4. Alternatively, the pipeline can attempt to mask short variants and normalise the genetic diversity between samples. As studies on population epigenetics tend to centre around species with low genetic diversity (cf. hierarchical clustering tree on genetic information in Supplementary Figure S2a), a hierarchical clustering based on sequence k-mer diversity (21) after masking short variants can instead give an indication of grouping based on DNA methylation patterns (Supplementary Figure S2b). Such an analysis can facilitate the identification of discrete groups prior to calling differentially methylated positions / regions, without limiting the analysis to only those methylated positions that are shared across all samples by a minimum threshold on sequencing depth. Otherwise, the distance matrix can instead be estimated from the methylation values in the conventional approach following per-sample methylation calling.

Appropriate groupings of samples are dependent on the specific experimental design of each study. Once identified, they can be subsequently evaluated for differential methylation with the EpiDiverse DMR pipeline, which analyses either all possible pairwise comparisons of groups or each group in relation to a designated control group. Conventionally, groups are identified based on *a priori* knowledge

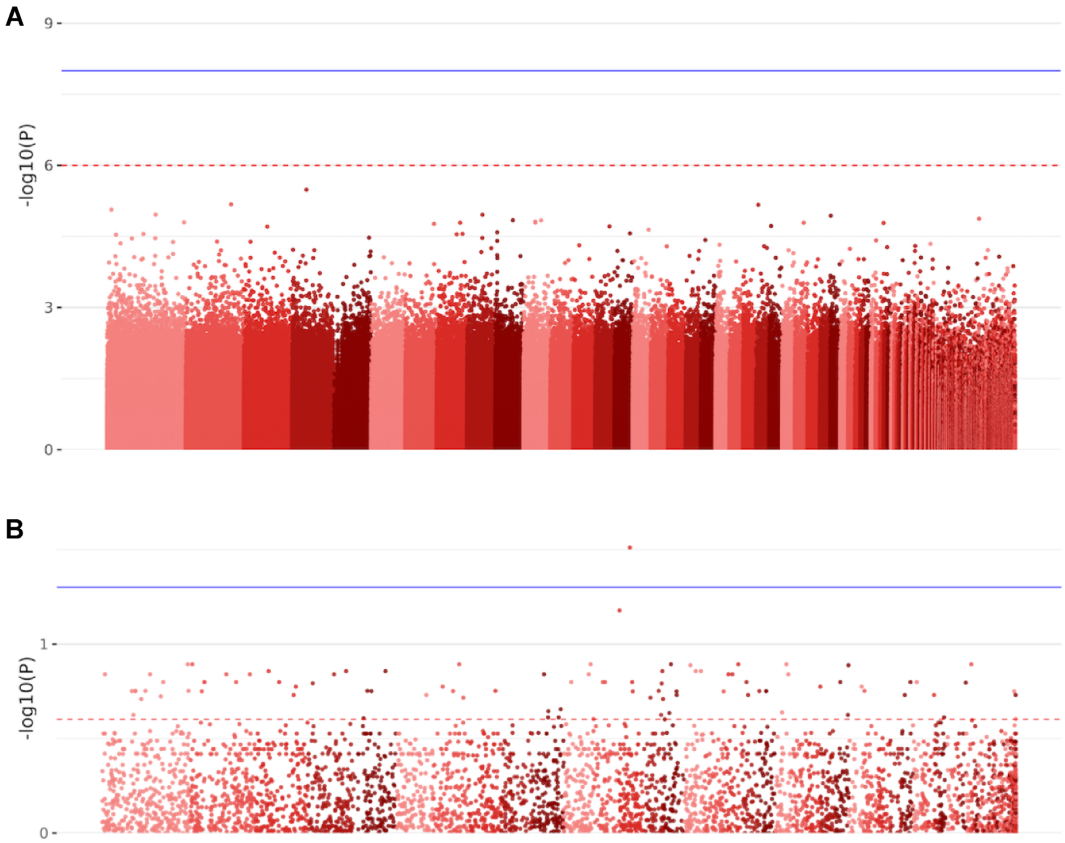


Figure 3. Manhattan plots demonstrating (**A**) the total number of tested positions during EWAS, from the cohort of *P. nigra* samples obtained from populations in Germany and Lithuania, and (**B**) the same analysis performed using significant DMRs instead. At the position-level, none were found to be significant ($P < 1 \times 10^{-8}$) or even suggestive ($P < 1 \times 10^{-6}$) based on common thresholds selected to account for the burden of multiple testing. At the region-level it becomes feasible to use Benjamini–Hochberg adjusted *P*-values (q), where 92 tests were found below a significance threshold of $q < 0.25$ and one even at $q < 0.05$. The plots are obtained automatically from the EWAS pipeline output (E-model).

or a global clustering of methylated sites. When grouping in this manner, however, local differences which are perhaps biologically relevant to the study question may be obscured by the global methylation profile and thus not revealed in the subsequent differential methylation analysis. Here, methylated sites (all contexts) obtained from the cohort of German and Lithuanian populations of *P. nigra* were subject to hierarchical clustering and the resulting tree cut at $\sim 5.25 \times 10^6$ to form two discrete groups and one outlier (Figure 2A). The total number of significant DMRs ($q < 0.05$) resulting from the pairwise comparison of these groups are given in Supplementary Table S5, and the corresponding heatmap showing the differential methylation level across the range of selected samples is shown in Figure 2B. Interestingly, the heatmap in some instances shows greater congruency with the clustering based on kWIP in Supplementary Figure S2b (e.g. LT_10, DE_41, DE_44, and a distinct clade with LT_01, LT_03, LT_04), indicating the potential utility as an alternative approach. While still a global clustering analysis, the local information inferred from sequence k-mers may be more robust in identifying groups based on regional differences in comparison to the site-by-site approach.

Finally, the accumulation of results from the WGBS and DMR pipelines were combined into a small analysis with EpiDiverse EWAS, based on the methylated sites in CG context and subset according to the significant DMRs discovered in the same context, using leaf flavonol content measured in the parent generation as a phenotypic trait. In the case of *P. nigra* the resulting manhattan plot (E-model) in Figure 3 reveals initially no significant QTLs below the common significance threshold of $P < 1 \times 10^{-8}$, or even below the suggestive significance threshold of $P < 1 \times 10^{-6}$, based on the global analysis of all methylated sites. The same analysis when conducted however at the region-level revealed a total of 92 significant QTLs ($q < 0.25$) which could be taken forward for further investigation (Supplementary Table S6). A brief inspection of these regions intersected with functional annotations in the *P. nigra* genome returned some features potentially relevant to flavonol content, including genes with homology to *ascorbate-specific transmembrane electron transporter 1*, *caspase family protein* and *mechanosensitive ion channel protein 3* alongside also *methyltransferases PMT2/PMT24*. Regions of hyper- or hypo- methylation may convey a more consistent association among the population of samples and can be more indicative of a mechanism which interacts for example with gene expression. This approach can be therefore more robust than the study of individual methylated sites, depending on the extent of stochastic variation in the DNA methy-

lation signal, but true associations may be missed in regions where DMRs were not identified as a result of local methylation differences which were obscured by global clustering techniques. Furthermore, the incorporation of SNP data into the G-model aspect of the EWAS pipeline can help to resolve any underlying genetic component which may be driving such associations with epigenetic markers.

A typical drawback of any (epi)genome-wide association study is the high burden of multiple testing, necessitating the use of a controlling procedure which can often be excessively conservative due to the high number of negative tests, thus obscuring many genuine biological findings which may be present within the dataset. The common significance threshold of $P < 1 \times 10^{-8}$ is based on a Bonferroni adjustment limited to a maximum of 1 million tests, regardless of the true number of tests. It is often argued with genetic data that a lack of true sample independence owing to linkage disequilibrium between SNPs can facilitate the use of this more heuristic variant of the Bonferroni adjustment, but statistically speaking this may be less than ideal. A more robust solution would be to reduce the total number of tests in the first place based on *a priori* knowledge. The EWAS pipeline therefore provides a mechanism to subset data based on any such regions provided by the end-user, for example here with DMRs obtained from the DMR pipeline, with the aim to reduce the majority of negative tests while still capturing the majority of positive tests. True positives may still be missed, depending largely on the selection criteria of such regions, though often more can be gained relative to the global analysis of all methylated positions.

CONCLUSION

The EpiDiverse Toolkit provides a suite of software pipelines for the analysis of ecological plant epigenetics, which adheres to the principles of 'FAIR' (Findability, Accessibility, Interoperability, and Reusability). The toolkit combines common procedures, such as mapping and methylation calling, with novel implementations for short variant calling and combining all results within a robust variation of EWAS, with each aspect benchmarked specifically for non-model plant species. This provides a consistent, repeatable framework which not only streamlines computational analyses within-species, but also facilitates more general comparisons between different organisms which may have evolved very different mechanisms involving DNA methylation.

DATA AVAILABILITY

All pipelines are open-source and publicly available through the https://github.com/EpiDiverse domain. The data used for analysis was generated by the European Training Network "EpiDiverse" to be published in the European Nucleotide Archive, and is otherwise available upon reasonable request to the authors.

SUPPLEMENTARY DATA

Supplementary Data are available at NARGAB Online.

ACKNOWLEDGEMENTS

We would like to thank all the members of the EpiDiverse Consortium for their active and invaluable support in discussing, developing and testing these tools.

FUNDING

The European Training Network "EpiDiverse" received funding from the EU Horizon 2020 program under Marie Skłodowska-Curie grant agreement No 764965.
Conflict of interest statement. None declared.

REFERENCES

1. Bossdorf,O., Arcuri,D., Richards,C.L. and Pigliucci,M. (2010) Experimental alteration of DNA methylation affects the phenotypic plasticity of ecologically relevant traits in Arabidopsis thaliana. *Evol. Ecol.*, **24**, 541–553.
2. Boyko,A. and Kovalchuk,I. (2010) Transgenerational response to stress in Arabidopsis thaliana. *Plant Signal. Behav.*, **5**, 995–998.
3. Cokus,S.J., Feng,S., Zhang,X., Chen,Z., Merriman,B., Haudenschild,C.D., Pradhan,S., Nelson,S.F., Pellegrini,M. and Jacobsen,S.E. (2008) Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature*, **452**, 215–219.
4. Ortega-Recalde,O., Peat,J.R., Bond,D.M. and Hore,T.A. (2021) Estimating Global Methylation and Erasure Using Low-Coverage Whole-Genome Bisulfite Sequencing (WGBS). In: Bogdanovic,O. and Vermeulen,M. (eds.) *TET Proteins and DNA Demethylation*. Methods in Molecular Biology, Humana Press, New York, Vol. **2272**, pp. 29–44.
5. Richards,C.L., Alonso,C., Becker,C., Bossdorf,O., Bucher,E., Colomé-Tatché,M., Durka,W., Engelhardt,J., Gaspar,B., Gogol-Döring *et al.* (2017) Ecological plant epigenetics: Evidence from model and non-model species, and the way forward. *Ecology letters*, **20**, 1576–1590.
6. Frommer,M., McDonald,L.E., Millar,D.S., Collis,C.M., Watt,F., Grigg,G.W., Molloy,P.L. and Paul,C.L. (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl. Acad. Sci. U.S.A.*, **89**, 1827–1831.
7. Tran,H., Porter,J., Sun,M.A., Xie,H. and Zhang,L. (2014) Objective and comprehensive evaluation of bisulfite short read mapping tools. *Adv. Bioinformatics*, **2014**, 1–11.
8. Liu,Y., Siegmund,K.D., Laird,P.W. and Berman,B.P. (2012) Bis-SNP: combined DNA methylation and SNP calling for Bisulfite-seq data. *Genome Biol.*, **13**, R61.
9. Schatz,M.C., Witkowski,J. and McCombie,W.R. (2012) Current challenges in de novo plant genome sequencing and assembly. *Genome Biol.*, **13**, 243.
10. Feng,S., Cokus,S.J., Zhang,X., Chen,P.Y., Bostick,M., Goll,M.G., Hetzel,J., Jain,J., Strauss,S.H., Halpern,M.E. *et al.* (2010) Conservation and divergence of methylation patterning in plants and animals. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 8689–8694.
11. di Tommaso,P., Chatzou,M., Floden,E.W., Barja,P.P., Palumbo,E. and Notredame,C. (2017) Nextflow enables reproducible computational workflows. *Nat. Biotechnol.*, **35**, 316–319.
12. Ewels,P.A., Peltzer,A., Fillinger,S., Patel,H., Alneberg,J., Wilm,A., Garcia,M.U., di Tomasso,P. and Nahnsen,A. (2020) The nf-core framework for community-curated bioinformatics pipelines. *Nat. Biotechnol.*, **38**, 276–278.
13. Grüning,B., Dale,R., Sjödin,A., Chapman,B.A., Rowe,J., Tomkins-Tinch,C.H., Valieris,R., Köster,J. and The Bioconda Team (2018) Bioconda: sustainable and comprehensive software distribution for the life sciences. *Nat. Methods*, **15**, 475–476.
14. Merkel,D. (2014) Docker: lightweight linux containers for consistent development and deployment. *Linux journal*, **2014**, 1–5.
15. Kurtzer,G.M., Sochat,V. and Bauer,M.W. (2017) Singularity: Scientific containers for mobility of compute. *PLoS One*, **12**, e0177459.

16. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.
17. Nunn,A., Otto,C., Stadler,P.F. and Langenberger,D. (2021) Comprehensive benchmarking of software for mapping whole genome bisulfite data: from read alignment to DNA methylation analysis. *Brief. Bioinformatics*, **22**, bbab021.
18. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G., Durbin,R. and 1000 Genome Project Data Processing,Subgroup (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
19. Nunn,A., Otto,C., Stadler,P.F. and Langenberger,D. (2021) Manipulating base quality scores enables variant calling from bisulfite sequencing alignments using conventional Bayesian approaches. bioRxiv doi: https://www.biorxiv.org/content/10.1101/2021.01.11.425926, 11 January 2021, preprint: not peer reviewed.
20. Garrison,E. and Marth,G. (2012) Haplotype-based variant detection from short-read sequencing. bioRxiv doi: https://arxiv.org/abs/1207.3907, 17 July 2012, preprint: not peer reviewed.
21. Murray,K.D., Webers,C., Ong,C.S., Borevitz,J. and Warthmann,N. (2017) kWIP: The k-mer weighted inner product, a de novo estimator of genetic similarity. *PLoS Comput. Biol.*, **13**, e1005727.
22. Li,H. (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, **27**, 2987–2993.
23. Kreutz,C., Can,N.S., Bruening,R.S., Meyberg,R., Mérai,Z., Fernandez-Pozo,N. and Rensing,S.A. (2020) A blind and independent benchmark study for detecting differentially methylated regions in plants. *Bioinformatics*, **36**, 3314–3321.
24. Jühling,F., Kretzmer,H., Bernhart,S.H., Otto,C., Stadler,P.F. and Hoffmann,S. (2016) metilene: Fast and sensitive calling of differentially methylated regions from bisulfite sequencing data. *Genome Res.*, **26**, 256–262.
25. Can,S.N., Nunn,A., Galanti,D., Langenberger,D., Becker,C., Volmer,K., Heer,K., Opgenoorth,L., Fernandez-Pozo,N. and Rensing,S.A. (2021) The EpiDiverse plant epigenome-wide association studies (EWAS) pipeline. *Epigenomes*, **5**, 12.
26. Pan,H., Holbrook,J.D., Karnani,N. and Kwoh,C.K. (2016) Gene, environment and methylation (GEM): a tool suite to efficiently navigate large scale epigenome wide association studies and integrate genotype and interaction between genotype and environment. *BMC Bioinformatics*, **17**, 299.

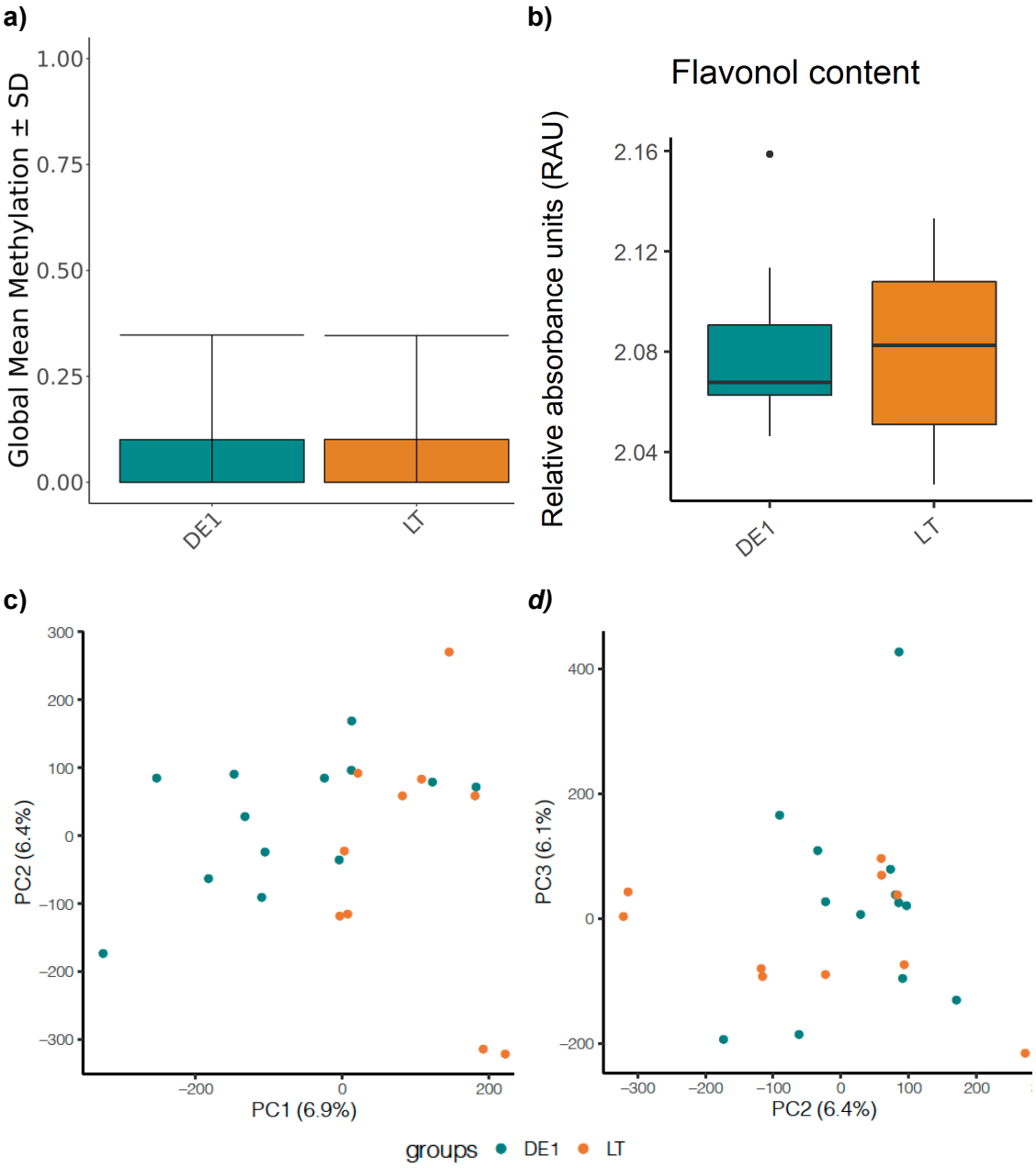
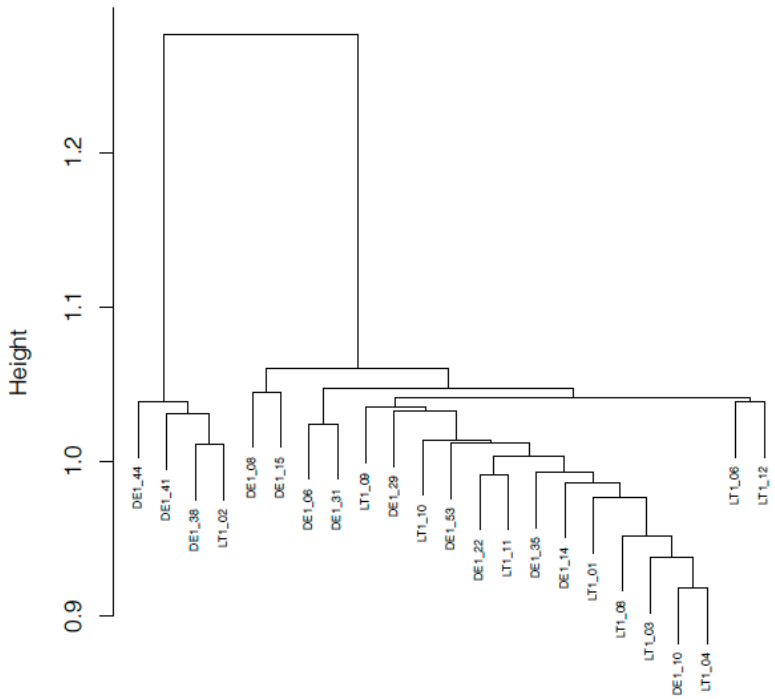


Figure S1. Descriptive statistics of the German (DE1) and Lithuanian (LT) populations of *Populus nigra* samples demonstrating **a)** global methylation levels in all contexts following cultivation under common garden conditions, **b)** leaf flavonol content as measured from samples collected in the field, and Principal Component Analysis (PCA) based on the per-site methylation levels in all contexts between **c)** component 1 and 2, and **d)** component 2 and 3.

a)



b)

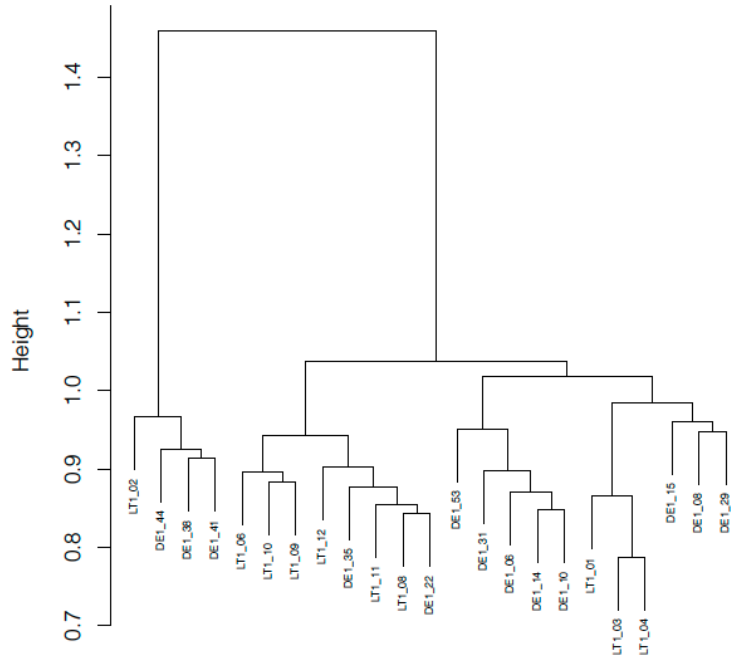


Figure S2. Hierarchical clustering of mappable FASTQ reads by k-mer diversity, using kWIP, following either **a)** bisulfite masking, or **b)** masking short variants to normalise genetic diversity.

CHAPTER VII

Synthesis

Bárbara Díez Rodríguez

It is predicted that in the next decades extreme weather events will increase in frequency and severity. These extreme events have devastating consequences and are causing mass mortality in many tree species. It is thus crucial to understand how plants respond to such events and other environmental cues. To cope with unpredicted environmental conditions, plants have evolved the ability to produce different phenotypes from one single genotype (Callahan *et al.,* 1997; Hughes *et al.,* 2008). In recent years, advances in molecular biology and genomics have shown that genetic variation is not the only cause of phenotypic variation among individuals (Rapp & Wendel, 2005). Mounting evidence suggests that epigenetic variation can play a role in shaping phenotypic plasticity (Richards *et al.,* 2017). Epigenetic mechanisms, such as DNA methylation, can in fact induce phenotypic changes (Niederhuth & Schmitz, 2017; Lloyd & Lister, 2022). Furthermore, these changes can be transmitted across clonal generations, suggesting that DNA methylation can in turn induce heritable phenotypic variation and potentially contribute to local adaptation in plants. However, the link between epigenetic variation and phenotypic variation remains poorly understood, and it is still unclear whether epigenetic variation can truly have an effect on plant adaptation.

In this thesis, I studied natural patterns of epigenetic variation and the link between epigenetic variation and phenotypic variation. Specifically, I used a clonally propagated tree cultivar (*Populus nigra* cv. Italica') to investigate patterns of DNA methylation variation, and its association with environmental conditions. Further, I assessed whether these patterns can be transmitted to the clonal offspring, thus potentially contributing to local adaptation of plant species. Below, I summarize the main results of my research concerning i) parental effects on phenotypic variation, ii) the effects of environmental conditions (i.e. biotic and abiotic stress) on DNA methylation, iii) natural patterns of DNA methylation and iv) the persistence of methylation patterns across clonal generations.

Effects of epigenetic variation on phenotypic variation

In **Chapter II** I investigated the effect of previous environments on the clonal offspring of Lombardy poplar ramets growing under common garden conditions. This is one of the few studies that explores the suitability of common garden experiments for studies on natural epigenetic variation at the landscape level. In this study, I surveyed several functional traits related to plant growth, ecophysiology, biotic stress, and leaf chemistry. Using linear mixed models, I studied the relationship between phenotypic variation and geographic origin. I found that phenotypic differences between genetically identical ramets can partially be attributed to shared environmental conditions and might be

transmitted as part of the epigenetic memory. These results are particularly relevant to the study of natural epigenetic variation, because common garden experiments have been proposed as a simple approach to exclude environmental noise and study transgenerational epigenetic variation (Latzel & Klimešová, 2010a; Richards *et al.*, 2017; Heer *et al.*, 2018). First, clones are characterized by low to zero genetic diversity, so effects of epigenetic variation on trait variation will not be confounded by effects of genetic variation. Second, epigenetic marks are more stable across clonal generations (Verhoeven & Preite, 2014). The poplar genotyping analysis showed three main genetic clusters, but a major fraction of the individuals sampled across Europe belonged to the same genotype, indicating that genetic variation among the clones was indeed considerably low (**Chapter II)**. Thus, within single poplar genotypes, in absence of genetic effects on trait variation, phenotypic variation among plants can also arise based on transmission of parental environmental effects, potentially mediated by epigenetic mechanisms (Raj *et al.*, 2011).

However, I also found that uneven microenvironmental conditions in the common garden had a significant effect on the observed phenotypic variation, possibly outweighing parental effects and thus allowing for short term acclimation to new environmental conditions. On the other hand, multiple studies have suggested that high phenotypic plasticity levels associated with epigenetic diversity might contribute to the successful establishment of clonal (and often invasive) plant species (Davidson *et al.*, 2011; Richards *et al.*, 2012; Mounger *et al.*, 2020). Specifically, changes in CHH methylation levels are less stable than those in CG or CHG (Secco *et al.*, 2017; Wibowo *et al.*, 2018). In poplar, methylation levels in CHH were shown to be associated with abiotic stress conditions (e.g. drought and heat stress, **Chapter III**) and with climatic variables (e.g. temperature and precipitation bioclimatic variables, **Chapter IV**). These results suggest that methylation patterns in CHH might be highly dynamic, acting rapidly in response to unpredicted environmental conditions. Thus, it is possible that any epigenetic effects associated with parental origin might be overwritten by the necessity to acclimate to new and/or stressful environmental conditions.

Effects of environmental cues on DNA methylation

The effects of environmental cues on DNA methylation have been extensively reviewed (Niederhuth *et al.*, 2016; Lloyd & Lister, 2021). However, many of the studies were constrained to short-lived annual species or used low-resolution molecular methods, leading to low statistical power for detecting DNA methylation variants (Hagmann *et al.*,

2015). Moreover, this has also led to the study of single environmental factors at a time, restricting the identified responsive genomic regions to specific experimental conditions. **Chapter III** presented a high-resolution analysis of the responses of the poplar methylome to different biotic (fungal infection, herbivory, salicylic acid treatment) and abiotic (drought, heat and cold) stress conditions. It provided further evidence that DNA methylation patterns vary in response to different environmental stimuli. In poplar, drought, heat and infection with a fungal pathogen can result in DNA hypermethylation events. On the other hand, cold and treatment with SA resulted in DNA hypomethylation. These responses seemed to be context specific, and might reflect differences in the duration of stress events (e.g. cold periods vs brief episodes of drought or heat). Interestingly, although some stress-specific regions were identified in this study, a considerable number of regions (especially in the CG and CHG contexts) showed a similar response irrespective of treatment, suggesting that the response to stress conditions in poplar is generic, and that many of these regions might have a regulatory role in the response to these conditions.

Additionally, the response induced by abiotic factors was stronger when compared to the response induced by biotic factors. From the poplar point of view, biotic stress conditions (like herbivory) occur every growing season, but some abiotic stress events (like extreme heat events) might not happen. We could therefore speculate that poplars may rely on plastic responses against expected conditions, but on the methylome to respond to unpredicted stress events. Methylation responses to drought stress in *Populus nigra* are particularly interesting from an ecological point of view. *Populus nigra* is a riparian species that depends on river flooding regimes for successful seed and cutting dispersal (Smulders *et al.*, 2008). In **Chapter IV**, I showed that global mCG levels in coding regions were negatively correlated with precipitation of the wettest quarter and month, and mCG and mCHG levels in gene promoters were correlated with precipitation in the warmest quarter. Because methylation in CG and CHG is mitotically stable, it is possible that certain degree of environmental information regarding water availability might be encoded as “epigenetic memory” and transmitted to the clonal offspring. This is further supported by the gene ontology enrichment analysis conducted in **Chapter III**, were poplar genes associated with drought CHH-DMRs seemed enriched in general responses to drought, such as ABA signaling, protein kinase activity, and response to water deprivation. The findings in **Chapters III** and **IV** support the notion that environmental cues can induce changes in the poplar methylome and may play a role at least in acclimation to new environmental conditions, if not in local adaptation.

Natural patterns of DNA methylation

Similarly to studies on the effects of environmental variation on DNA methylation, studies on natural patterns of methylation have been limited by the elevated cost of high-resolution sequencing techniques. Thus, these studies have often been performed under controlled greenhouse conditions or limited to small-scale geographic gradients (Lira-Medeiros *et al.*, 2010; Nicotra *et al.*, 2015; Avramidou *et al.*, 2015; Gugger *et al.*, 2016; Herrera *et al.*, 2017; Dewan *et al.*, 2018; Gáspár *et al.*, 2019). Using two clonal species, *P. nigra* and *F. vesca*, **Chapters IV** and **V** investigated DNA methylation patterns in natural populations at the landscape level. In **Chapter IV**, I used the poplar individuals established in the common garden described in **Chapter II** to study natural methylation patterns across Europe. I found that variation in DNA methylation did not have any evident geographic structure, but average methylation levels in gene promoters, coding regions and transposable elements were associated with bioclimatic variables, especially in the CHH context. Furthermore, variation in the methylome was not correlated with genetic variation. In this clonal system, the accumulation of methylation variation can be attributed mostly to spontaneous and environmentally induced variation (see also **Chapter III**). Somatic mutations have been shown to accumulate during vegetative propagation in other *Populus* species, affecting genetic variation within single genotypes (Tuskan *et al.*, 1996; Ally *et al.*, 2008; Chenault *et al.*, 2011). In a similar fashion, several studies have reported spontaneous methylation changes resulting from tissue culture (Kaeppler & Phillips, 1993; Fraga *et al.*, 2002; Dubrovina & Kiselev, 2016; Han *et al.*, 2018). These methylation changes can be stable and accumulate across somatic development and appear at higher rate than genetic mutations (Stroud *et al.*, 2013; Hofmeister *et al.*, 2020). Considering that the 'Italica' cultivar has been artificially propagated for the last two centuries, spontaneous epimutations might have accumulated across several clonal generations, creating a layer of "background noise" in the methylome that might be confounding a stronger environmental signal. Nevertheless, **Chapter III** identified many CpG/CHG-DMRs among parental individuals, which are thought to be mitotically stable and hence clonally transmissible, and a large proportion of stress-induced DMRs that overlapped with ortet-DMRs. This seems to indicate that at least part of the natural methylation variation in poplar is induced by changing environments, and not by spontaneous epimutations. Consequently, environment-induced methylation variants in CpG/CHG contexts could be fixed and appear as natural epialleles detectable across the tree lifespan and maybe next clonal generations.

Chapter V found strong evidence of genetic and epigenetic geographic structure in *F. vesca*, suggesting that epigenetic structure might be associated with genetic structure. However, in the CG and CHG contexts, country of origin accounted for higher percentage of variance than for genetic variants (respectively, 11.9%, 9.4% and 8.7% for CG, CHG and SNPs), indicating that part of the epigenetic variation was not under genetic control. Furthermore, they identified climate-predicted DMRs, and the number of these DMRs increased from CG to CHG and CHH, indicating that methylation in CHH might be particularly sensitive to climatic conditions. Similarly, Galanti and colleagues (Galanti *et al.*, 2022) showed that genetic variation in *Thlaspi arvense* strongly contributed to epigenetic variation, but also identified a similar number of climate-predicted DMRs. They also showed that average methylation levels in CHH were correlated with bioclimatic variables, especially with temperature related ones (see also **Chapter IV**). The differences in the amount of epigenetic variation explained by genetic variation seem to indicate that natural methylation patterns might reflect life history traits and differ between sexual and asexual species. *T. arvense* is a sexually reproducing annual species and *F. vesca* can reproduce both sexually and clonally. *Populus nigra*, however, is a dioecious species that has both male and female individuals. The Italica cultivar is an exclusively male clonal cultivar, because the characteristic columnar growth would disappear in the sexual offspring. Furthermore, the cultivar has been mostly artificially propagated by humans, so it lacks any genetic population structure (*sensu stricto,* see **Chapter II**). The lack of epigenetic population structure could then reflect this unique origin, if most of the individuals found across Europe are the clonal offspring of one (or maybe a few) original clones. In practical terms, this could mean that methylation in the CHH context is plastic and involved in plastic responses, while methylation in CG and CHG can be transmitted through multiple clonal generations. Despite the mentioned differences among species, methylation patterns in the CHH context seem to be associated with climatic variables in all three cases, providing further evidence that methylation in this context might be dynamic and play a role in plant phenotypic plasticity.

Persistence of DNA methylation patterns: implications for plant adaptation

The question of whether epigenetic mechanisms, and DNA methylation in particular, can have an adaptive potential and can ultimately act as an evolutionary driver has been the focus of considerable interest in the field of epigenetics (Bossdorf *et al.*, 2008; Richards *et al.*, 2017; Heer *et al.*, 2018; Boquete *et al.*, 2021). If DNA methylation can mediate phenotypic plastic responses in the short term, and be under selection processes, it

might also act as a bridge to local adaptation. For this to occur, both epigenetic variation and its associated phenotype must be heritable and independent from genetic variation. DNA methylation can indeed induce phenotypic variation (reviewed in Lloyd and Lister, 2022), and epigenetically induced phenotypic responses can be transmitted to the offspring (e.g., Verhoeven et al., 2010). However, to what extent these phenotypic responses are transmitted to the offspring and how much they contribute to plant adaptation is still being debated (Platt *et al.*, 2015; Schmid *et al.*, 2018; Münzbergová *et al.*, 2019; Muyle *et al.*, 2021). **Chapters IV** and **V** further explore the persistence of DNA methylation patterns in natural populations in *Populus* and *Fragaria*, respectively. It has been proposed that the inheritance of environmentally induced DNA methylation variation might be particularly prominent across clonal generations, due to the lack of meiosis in soma-clonal reproduction (Feng, Jacobsen, and Reik 2010; Latzel & Klimešová, 2010b; Simonetti et al., 2019). In poplar, in the CG context (symmetrical and mostly stable), individuals with the same geographic origin clustered together regardless of stress treatment, indicating similar methylation profiles (**Chapter III)**. In **Chapter IV**, I compared parental (ortets) and offspring (ramets) poplar individuals. I showed that, also in the CG context, the ortets and their ramets clustered together. The cuttings used in the stress experiments of **Chapter III** all belonged to the same original ortet. This suggests that the grouping observed in both chapters is the result of clonally transmitted methylation patterns. If methylation patterns in the CG context are faithfully transmitted to the next clonal generation, this would partially explain the patterns observed in *F. vesca* **(Chapter V)**, where ortet and ramets also clustered together. Furthermore, a fraction of DMRs is present both among the poplar ortets and among the ramets in the CG and CHG contexts, suggesting that some DMRs might be stable across clonal generations, especially in symmetrical contexts (**Chapter IV)**. From a practical perspective, the fact that the same DMR was found among poplar ortets and among ramets implicates that those specific regions were differentially methylated both in the ortets and in the ramets. It is also possible that some regions might have the exact same methylation levels. If this was the case our pipelines would fail to identify these regions, but they would still be methylated regions that were transmitted to the clonal offspring. In **Chapter IV**, I also called DMRs between ortets and ramets. Interestingly, in the case of some comparisons, I found a very low number of DMRs (e.g. Spain, Italy 2 and Norway). This could mean two things: 1) as mentioned above, the methylation patterns were faithfully transmitted to the ramets, or 2) the environmental conditions in the field and in the common garden were similar, and DMRs acting as regulatory regions were responding in a similar way. However, in the CHH context the number of DMRs found between ortets and ramets sampled in Italy was considerably high, supporting the first

assumption rather than the second. Ultimately, the patterns observed both in *Populus* and *Fragaria* provide further evidence that DNA methylation patterns can be transmitted across clonal generations, but the link between these patterns and phenotypic variation needs to be further explored.

References

Ally D, Ritland K, Otto SP. **2008**. Can clone size serve as a proxy for clone age: An exploration using microsatellite divergence in *Populus tremuloides*. *Molecular Ecology* **17**: 4897–4911.

Avramidou E v., Ganopoulos I v., Doulis AG, Tsaftaris AS, Aravanopoulos FA. **2015**. Beyond population genetics: natural epigenetic variation in wild cherry (Prunus avium). *Tree Genetics and Genomes* **11**: 1–9.

Boquete MT, Muyle A, Alonso C. **2021**. Plant epigenetics: phenotypic and functional diversity beyond the DNA sequence. *American Journal of Botany* **108**: 553–558.

Bossdorf O, Richards CL, Pigliucci M. **2008**. Epigenetics for ecologists. *Ecology Letters* **11**: 106–115.

Chenault N, Arnaud-Haond S, Juteau M, Valade R, Almeida JL, Villar M, Bastien C, Dowkiw A. **2011**. SSR-based analysis of clonality, spatial genetic structure and introgression from the Lombardy poplar into a natural population of *Populus nigra* L. along the Loire River. *Tree Genetics & Genomes 2011 7:6* **7**: 1249–1262.

Davidson AM, Jennions M, Nicotra AB. **2011**. Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta-analysis. *Ecology Letters* **14**: 419–431.

Dewan S, de Frenne P, vanden Broeck A, Steenackers M, vander Mijnsbrugge K, Verheyen K. **2018**. Transgenerational effects in asexually reproduced offspring of *Populus* (D Gomory, Ed.). *PLOS ONE* **13**: e0208591.

Dubrovina AS, Kiselev K v. 2016. Age-associated alterations in the somatic mutation and DNA methylation levels in plants. *Plant Biology* **18**: 185–196.

Feng S, Jacobsen SE, Reik W. **2010**. Epigenetic reprogramming in plant and animal development. *Science (New York, N.Y.)* **330**: 622–627.

Fraga MF, Rodriguez R, Canal MJ. **2002**. Genomic DNA methylation-demethylation during aging and reinvigoration of Pinus radiata. *Tree Physiology* **22**: 813–816.

Galanti D, Ramos-Cruz D, Nunn A, Rodríguez-Arévalo I, Scheepens JF, Becker C, Bossdorf O. **2022**. Genetic and environmental drivers of large-scale epigenetic variation in Thlaspi arvense (NM Springer, Ed.). *PLOS Genetics* **18**: e1010452.

Gáspár B, Bossdorf O, Durka W. **2019**. Structure, stability and ecological significance of natural epigenetic variation: a large-scale survey in Plantago lanceolata. *New Phytologist* **221**: 1585–1596.

Gugger PF, Fitz-Gibbon S, Pellegrini M, Sork VL. **2016**. Species-wide patterns of DNA methylation variation in Quercus lobata and their association with climate gradients. *Molecular Ecology* **25**: 1665–1680.

Hagmann J, Becker C, Müller J, Stegle O, Meyer RC, Wang G, Schneeberger K, Fitz J, Altmann T, Bergelson J, *et al.* 2015. Century-scale Methylome Stability in a Recently Diverged Arabidopsis thaliana Lineage. *PLOS Genetics* **11**: e1004920.

Han Z, Crisp PA, Stelpflug S, Kaeppler SM, Li Q, Springer NM. **2018**. Heritable epigenomic changes to the maize methylome resulting from tissue culture.

Heer K, Mounger J, Boquete MT, Richards CL, Opgenoorth L. **2018**. The diversifying field of plant epigenetics. *New Phytologist* **217**: 988–992.

Herrera CM, Medrano M, Bazaga P. **2017**. Comparative epigenetic and genetic spatial structure of the perennial herb *Helleborus foetidus*: Isolation by environment, isolation by distance, and functional trait divergence. *American Journal of Botany* **104**: 1195–1204.

Hofmeister BT, Denkena J, Colomé-Tatché M, Shahryary Y, Hazarika R, Grimwood J, Mamidi S, Jenkins J, Grabowski PP, Sreedasyam A, *et al.* 2020. A genome assembly and the somatic genetic and epigenetic mutation rate in a wild long-lived perennial *Populus trichocarpa*. *Genome Biology* **21**: 1–27.

Kaeppler SM, Phillips RL. **1993**. DNA methylation and tissue culture-induced variation in plants. *In Vitro Cellular & Developmental Biology - Plant 1993 29:3* **29**: 125–130.

Latzel V, Klimešová J. **2010a**. Transgenerational plasticity in clonal plants. *Evolutionary Ecology* **24**: 1537–1543.

Latzel V, Klimešová J. **2010b**. Transgenerational plasticity in clonal plants. *Evolutionary Ecology* **24**: 1537–1543.

Lira-Medeiros CF, Parisod C, Fernandes RA, Mata CS, Cardoso MA, Ferreira PCG. **2010**. Epigenetic Variation in Mangrove Plants Occurring in Contrasting Natural Environment. *PLOS ONE* **5**: e10326.

Lloyd JPB, Lister R. **2021**. Epigenome plasticity in plants. *Nature Reviews Genetics 2021 23:1* **23**: 55–68.

Mounger J, Ainouche M, Bossdorf O, Cavé-Radet A, Li B, Parepa M, Salmon A, Yang J, Richards C. **2020**. Epigenetics and the success of invasive plants.

Münzbergová Z, Latzel V, Šurinová M, Hadincová V. **2019**. DNA methylation as a possible mechanism affecting ability of natural populations to adapt to changing climate. *Oikos* **128**: 124–134.

Muyle A, Ross-Ibarra J, Seymour DK, Gaut BS. **2021**. Gene body methylation is under selection in Arabidopsis thaliana. *Genetics* **218**.

Nicotra AB, Segal DL, Hoyle GL, Schrey AW, Verhoeven KJF, Richards CL. **2015**. Adaptive plasticity and epigenetic variation in response to warming in an Alpine plant. *Ecology and Evolution* **5**: 634–647.

Niederhuth CE, Bewick AJ, Ji L, Alabady MS, Kim K do, Li Q, Rohr NA, Rambani A, Burke JM, Udall JA, *et al.* 2016. Widespread natural variation of DNA methylation within angiosperms. *Genome Biology* **17**.

Platt A, Gugger PF, Pellegrini M, Sork VL. **2015**. Genome-wide signature of local adaptation linked to variable CpG methylation in oak populations. *Molecular Ecology* **24**: 3823–3830.

Raj S, Bräutigam K, Hamanishi ET, Wilkins O, Thomas BR, Schroeder W, Mansfield SD, Plant AL, Campbell MM. **2011**. Clone history shapes *Populus* drought responses. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 12521–12526.

Richards CL, Alonso C, Becker C, Bossdorf O, Bucher E, Colomé-Tatché M, Durka W, Engelhardt J, Gaspar B, Gogol-Döring A, *et al.* 2017. Ecological plant epigenetics: Evidence from model and non-model species, and the way forward (T Coulson, Ed.). *Ecology Letters* **20**: 1576–1590.

Richards CL, Schrey AW, Pigliucci M. **2012**. Invasion of diverse habitats by few Japanese knotweed genotypes is correlated with epigenetic differentiation (M Vellend, Ed.). *Ecology Letters* **15**: 1016–1025.

Schmid MW, Heichinger C, Coman Schmid D, Guthörl D, Gagliardini V, Bruggmann R, Aluri S, Aquino C, Schmid B, Turnbull LA, *et al.* 2018. Contribution of epigenetic variation to adaptation in Arabidopsis. *Nature Communications 2018 9:1* **9**: 1–12.

Secco D, Whelan J, Rouached H, Lister R. **2017**. Nutrient stress-induced chromatin changes in plants. *Current Opinion in Plant Biology* **39**: 1–7.

Simonetti G, Pucci N, Brasili E, Valletta A, Sammarco I, Carnevale E, Pasqua G, Loreti S. **2019**. In vitro antimicrobial activity of plant extracts against Pseudomonas syringae pv. actinidiae causal agent of bacterial canker in kiwifruit. *Plant Biosystems*.

Smulders MJM, Cottrell JE, Lefèvre F, van der Schoot J, Arens P, Vosman B, Tabbener HE, Grassi F, Fossati T, Castiglione S, *et al.* 2008. Structure of the genetic diversity in black poplar

(*Populus nigra* L.) populations across European river systems: Consequences for conservation and restoration. *Forest Ecology and Management* **255**: 1388–1399.

Stroud H, Greenberg MVC, Feng S, Bernatavichute Y v., Jacobsen SE. **2013**. Comprehensive Analysis of Silencing Mutants Reveals Complex Regulation of the Arabidopsis Methylome. *Cell* **152**: 352–364.

Tuskan GA, Francis KE, Russ SL, Romme WH, Turner MG. **1996**. RAPD markers reveal diversity within and among clonal and seedling stands of aspen in Yellowstone National Park, U.S.A. *Canadian Journal of Forest Research* **26**: 2088–2098.

Verhoeven KJF, Jansen JJ, van Dijk PJ, Biere A. **2010**. Stress-induced DNA methylation changes and their heritability in asexual dandelions. *New Phytologist* **185**: 1108–1118.

Verhoeven KJF, Preite V. **2014**. Epigenetic variation in asexxually reproducing organisms. *Evolution* **68**: 644–655.

Wibowo A, Becker C, Durr J, Price J, Spaepen S, Hilton S, Putra H, Papareddy R, Saintain Q, Harvey S, *et al.* 2018. Incomplete reprogramming of cell-specific epigenetic marks during asexual reproduction leads to heritable phenotypic variation in plants. *bioRxiv*: 267955.

CHAPTER VIII

The road goes ever on: Current challenges and future perspectives

Bárbara Díez Rodríguez

My work and recent reviews (Niederhuth *et al.*, 2016; Zhang *et al.*, 2018; Lloyd & Lister, 2021; Hannan Parker *et al.*, 2022) stress that plant epigenetic variation is far more complex than previously assumed, and that our knowledge of the role of epigenetic mechanisms in phenotypic variation and plant adaptation is still extremely limited. In this chapter, I would like to focus on current challenges and future perspectives in the field of plant epigenetics.

The study of ecological epigenetics has been hindered not only by the limited economical resources, but also by computational ones. While computational analyses of genome-wide data have become routine in model species, not all approaches are easily transferred to ecological epigenetics and non-model species. The lack of high-quality reference genomes, transcriptomes, and other resources has considerably impeded advances in the field. However, with costs of sequencing decreasing, this will probably change in the near future. This in practice means that, with limited computational resources, the sheer amount of data generated can also be a major bottleneck. While parallelization of processes can help with the optimal use of computer resources, there is still room for improvement. Furthermore, the very high number of comparisons which are often necessary in the statistical analysis of such datasets create a multiple testing burden. This in turn means that a post-hoc correction is performed, often leaving few significant results despite clear differences being observed.

From a conceptual perspective, downstream analyses of population-level data bring their own challenges. For example, how much does epigenetic variation depend on genetic variation? To what extent do stochastic epimutations contribute to epigenetic variation? Can they be considered “epigenetic noise”, or do they also contribute to the effects of epigenetic variation in phenotypic plasticity? How do we separate these stochastic epimutations from other environmentally induced epimutations, and do we need to? DMRs are considered to be more biologically relevant from a functional point of view, because they can affect longer DNA regions. But this poses a problem on its own: to this day, there is no standard definition of what a DMR is. Because there is also no standard bioinformatic practices, and each researcher decides what is the most appropriate method for the questions each experiment is trying to address, the definition of DMR varies considerably between studies (Kreutz et al., 2020; Robinson et al., 2014). Furthermore, in a lot of cases, calling DMRs between groups implies defining the groups before the analysis (e.g. each population is a group) which by default introduces a bias, because population-level DMRs might not follow a predefined structure (Akalin *et al.*, 2012; Jühling *et al.*, 2016; Hüther *et al.*, 2022)

These challenges are further merged to create a unicorn: Epigenome Wide Association Studies (EWAS) (Rakyan *et al.*, 2011). Like genome-wide association studies (GWAS), EWAS makes use of linear models in population-scale data to reveal significant epigenotype-phenotype associations. However, given how dynamic epigenetic changes can be, it is difficult to discern a significant relationship between phenotype and epigenetic variation. This remains the most challenging aspect in any study on the effects of natural epigenetic variation on phenotypic variation. Thus, available methods to perform EWAS remain scarce, particularly in the context of plant ecology. In order to extend its functionality in the context of plant ecology, the EpiDiverse consortium has adapted the R package Gene, Environment and Methylation (GEM) (Pan *et al.*, 2016) in a pipeline for EWAS analysis (as part of the EpiDiverse Toolkit, see **Chapter VI,** Nunn et al., 2021). As part of my research, I used the data generated in **Chapters II** and **IV** to find significant associations between phenotypic variation and methylation variants. However, the available computational resources at the time were insufficient for the kind of testing that the pipeline requires. Furthermore, the number of tests carried out by the pipeline, and the necessity to correct for multiple testing, resulted in a lack of significant methylation-phenotype associations, rendering the analysis inconclusive. In the future, this pipeline will most probably be improved, thus opening new possibilities for ecological epigenetics.

On a more positive note, recent advances in the field of ecological epigenetics have potential applications for nature conservation policies. In 2019, Rey and colleagues (Rey *et al.*, 2020) reviewed current knowledge about the importance of epigenetic mechanisms in (a) orchestrating fundamental development alternatives in organisms, (b) enabling individuals to respond rapidly to selection pressures and (c) improving ecosystem stability and functioning. They concluded that an *epigenetic conservation* perspective will provide environmental managers with the possibility to refine evolutionary significant units (ESUs), to set conservation plans taking into account the capacity of organisms to rapidly cope with environmental changes, and hence to improve the conservation of wild populations. In addition, epigenetic regulation of plant responses to adverse environmental conditions can potentially provide a tunable mechanism to optimize plant growth, adaptation and ultimately yield, in a world were food security and sustainable agriculture are facing increasing challenges (Springer & Schmitz, 2017).

To conclude, the plant epigenome appears to be highly plastic and highly dynamic. Not only further research on more ecologically relevant plant species is needed, but also novel approaches in experimental design. For example, though many studies have investigated the effects of specific stress conditions on DNA methylation, only the study

presented in Chapter III has explored a multi-stress global response (Peña-Ponton *et al.*, 2022). Similarly, although many studies have dealt with the inheritance of methylation patterns, only a few have addressed seasonal changes of DNA and differences among plant tissues (Daccord et al., 2017; Ito et al., 2019; Perrin et al., 2020; D'Amico-Willman et al., 2022). Ultimately, further research is crucially needed in order to advance our knowledge on how epigenetic mechanisms affect ecological processes.

References

Akalin A, Kormaksson M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, Mason CE. **2012**. MethylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biology* **13**.

Daccord N, Celton JM, Linsmith G, Becker C, Choisne N, Schijlen E, van de Geest H, Bianco L, Micheletti D, Velasco R, *et al.* 2017. High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. *Nature Genetics 2017 49:7* **49**: 1099–1106.

D'Amico-Willman KM, Ouma WZ, Meulia T, Sideli GM, Gradziel TM, Fresnedo-Ramírez J. **2022**. Whole-genome sequence and methylome profiling of the almond [Prunus dulcis (Mill.) D.A. Webb] cultivar 'Nonpareil'. *G3 Genes|Genomes|Genetics* **12**.

Hannan Parker A, Wilkinson SW, Ton J. **2022**. Epigenetics: a catalyst of plant immunity against pathogens. *New Phytologist* **233**: 66–83.

Hüther P, Hagmann J, Nunn A, Kakoulidou I, Pisupati R, Langenberger D, Weigel D, Johannes F, Schultheiss SJ, Becker C. **2022**. MethylScore, a pipeline for accurate and context-aware identification of differentially methylated regions from population-scale plant whole-genome bisulfite sequencing data. *Quantitative Plant Biology* **3**: e19.

Ito, Nishio, Tarutani, Emura, Honjo, Toyoda, Fujiyama, Kakutani, Kudoh. **2019**. Seasonal Stability and Dynamics of DNA Methylation in Plants in a Natural Environment. *Genes* **10**: 544.

Jühling F, Kretzmer H, Bernhart SH, Otto C, Stadler PF, Hoffmann S. **2016**. metilene: fast and sensitive calling of differentially methylated regions from bisulfite sequencing data. *Genome Research* **26**: 256–262.

Kreutz C, Can NS, Bruening RS, Meyberg R, Mérai Z, Fernandez-Pozo N, Rensing SA. **2020**. A blind and independent benchmark study for detecting differentially methylated regions in plants. *Bioinformatics* **36**: 3314–3321.

Lloyd JPB, Lister R. **2021**. Epigenome plasticity in plants. *Nature Reviews Genetics 2021 23:1* **23**: 55–68.

Niederhuth CE, Bewick AJ, Ji L, Alabady M, Kim KD, Page JT, Li Q, Rohr NA, Rambani A, Burke JM, *et al.* 2016. Widespread natural variation of DNA methylation within angiosperms. *bioRxiv*.

Nunn A, Can SN, Otto C, Fasold M, Diéz Rodríguez B, Fernández-Pozo N, Rensing SA, Stadler PF, Langenberger D. **2021**. EpiDiverse Toolkit: a pipeline suite for the analysis of bisulfite sequencing data in ecological plant epigenetics. *NAR Genomics and Bioinformatics* **3**.

Pan H, Holbrook JD, Karnani N, Kwoh CK. **2016**. Gene, Environment and Methylation (GEM): A tool suite to efficiently navigate large scale epigenome wide association studies and integrate genotype and interaction between genotype and environment. *BMC Bioinformatics* **17**: 1–8.

Peña-Ponton C, Diez-Rodriguez B, Perez-Bello P, Becker C, McIntyre LM, Putten W van der, Paoli E de, Heer K, Opgenoorth L, Verhoeven KJF. **2022**. High-resolution methylome analysis in the clonal *Populus nigra* cv. 'Italica' reveals environmentally sensitive hotspots and drought-responsive TE superfamilies. *bioRxiv*: 2022.10.18.512698.

Perrin A, Daccord N, Roquis D, Celton JM, Vergne E, Bucher E. **2020**. Divergent DNA Methylation Signatures of Juvenile Seedlings, Grafts and Adult Apple Trees. *Epigenomes 2020, Vol. 4, Page 4* **4**: 4.

Rakyan VK, Down TA, Balding DJ, Beck S. **2011**. Epigenome-wide association studies for common human diseases. *Nature Reviews Genetics 2011 12:8* **12**: 529–541.

Rey O, Eizaguirre C, Angers B, Baltazar-Soares M, Sagonas K, Prunier JG, Blanchet S. **2020**. Linking epigenetics and biological conservation: Towards a conservation epigenetics perspective. *Functional Ecology* **34**: 414–427.

Robinson MD, Kahraman A, Law CW, Lindsay H, Nowicka M, Weber LM, Zhou X. **2014**. Statistical methods for detecting differentially methylated loci and regions. *Frontiers in Genetics* **5**: 324.

Springer NM, Schmitz RJ. **2017**. Exploiting induced and natural epigenetic variation for crop improvement. *Nature Reviews Genetics 2017 18:9* **18**: 563–575.

Zhang H, Lang Z, Zhu JK. **2018**. Dynamics and function of DNA methylation in plants. *Nature Reviews Molecular Cell Biology 2018 19:8* **19**: 489–506.